# A Benchmark Dataset for Segmenting Liver, Vasculature and Lesions from Large-scale Computed Tomography Data

Bo Wang[1,3], Qingsen Yan[2,3], Zhengqing Xu[1],Jingyang Ai[3], Shuo Jin[1], Wei Xu[1], Wei Zhao[3], Liang Zhang[4], Zheng You[1]
[1]Tsinghua University, China, [2]The University of Adelaide, Australia, [3]Beijing Jingzhen Technology Ltd. [4]Xidian University

ICPR 2020
25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
Milan, Italy 10 | 15 January 2021

## Introduction

How to build a high-performance liver-related computer assisted diagnosis system is an open question of great interest. However, the performance of the state-of-art algorithm is always limited by the amount of data and the quality of the label. To address this problem, we propose the biggest treatment-oriented liver cancer dataset for liver surgery and treatment planning. This dataset provides 216 cases (total about 268K frames) scanned images in contrast-enhanced computed tomography (CT). We labeled all the CT images with the liver, liver vasculature, and liver tumor segmentation ground truth for train and tune segmentation algorithms in advance. Based on that, we evaluate several recent and state-of-the-art segmentation algorithms, including 7 deep learning methods, on CT sequences. All results are compared to reference segmentations five error metrics that highlight different aspects of segmentation accuracy.

## Contributions

The main contributions of this work are summarized as follows.

**Core contributions:**

- We propose the biggest treatment-oriented liver cancer dataset for liver surgery and treatment planning. To the best of our knowledge, this is the first attempt to do this in liver computer assisted diagnosis.
- We summarize datasets of liver imaging and point the characters (e.g. imaging method and parameters) of each dataset.
- We design three tasks to construct a benchmark that compares several state-of-theart methods on the proposed dataset.

## Related Works

**Dataset:**

**3D-IRCADb-01**: There are 20 3D CT volumes in this dataset. 75% of these images have hepatic tumors. The liver edges, liver vessels and liver tumors are labeled out for segmentation tasks. Besides, these image characteristics, such as the liver size and the location of tumors, are listed in a table on the web page.

**Sliver07**: This dataset is provided by the Segmentation of the Liver Competition 2007. 30 CT cases are available for the public. 20 cases are randomly chosen as the training set, and the remaining 10 cases for testing. Only the segmentation reference of the training set is downloadable.

**ImageCLEF**: This dataset belongs to the ImageCLEF Liver CT Image Annotation Challenge. It includes 50 training cases and 10 testing cases. Each case data consists of the raw image, the liver lesion bounding box and the liver mask. In addition, the CT features of the liver, vessels and lesions are listed.

**LiTs**: This dataset is from the Liver Tumor Segmentation Benchmark. It contains 131 CT cases in the training set and 30 CT cases in the test set. Only these images in the training set are labeled with liver contour and liver tumor area.

**Jiang** et al.: This dataset is used in [15]. In this dataset, 83 DCE-MRI images are scanned from 83 patients with liver cancer. The training set and the test set contain 62 and 21 cases, respectively.

**Ouhmich** et al.: The authors used an in-house database (Db) containing ground truth 2D annotations for three classes, namely the healthy parenchyma, the active and the necrotic parts of the lesions. This database is composed of data from seven patients suffering from hepatocellular carcinoma (HCC) that underwent from one to three contrast-enhanced computed tomography (CECT) examinations, resulting in a total number of 13 CT sequences.

**Segmentation Algorithms:**

1) With cheap hand-crafted features, hand-crafted feature based methods have gained attention due to their computational efficiency and competitive performance.

2) Many researchers followed this trend and proposed to utilize various CNNs for learning feature representations in the application of liver and lesion segmentation.

## Experiments

**Dataset:**

TABLE I
COMPARING THE PROPOSED DATASET WITH OTHER LIVER IMAGING DATASETS.

| Dataset | D-IRCADb [11] | Sliver07 [12] | ImageCLEF [13] | LiTs [14] | Jiang [15] | Ouhmich [16] | **Our** |
|---|---|---|---|---|---|---|---|
| Slices(/case) | 2,823(70-260) | 6,427(64-502) | 41-588 | 58,638(42-1026) | 4,283(42-84) | 104(8) | **268,431(205-1,207)** |
| Cases | 20 | 30 | 60 | 201 | 117 | 13 | **216** |
| Thickness(mm) | 1.25-4mm | 1-3mm | 0.399-2.5mm | 0.45-6mm | 1.25-5.0mm | 0.7-1.25mm | **0.62-1.25mm** |
| Resolution | 0.56-0.87mm | 0.55-0.8mm | 0.674-1.007mm | 0.56-1mm | 0.71-1.17mm | 0.66-0.97mm | **0.55-0.76mm** |
| Liver label | ✓ | ✓ | ✓ | ✓ | None | None | ✓ |
| Vessel label | ✓ | None | None | None | None | None | ✓ |
| Tumor label | ✓ | None | (bounding box) | ✓ | ✓ | ✓ | ✓ |
| Diagnose report | None | None | None | None | None | None | Li-Rads |

TABLE II
COMPARING DIFFERENT METHODS WITH THE PROPOSED DATASET ON THE LIVER SEGMENTATION TASK.

| Metrics | FCN | UNet | UNet++ | Attention UNet | VNet | 3D UNet | 3D Attention |
|---|---|---|---|---|---|---|---|
| Precision | 0.9612 | 0.9670 | 0.9652 | 0.9667 | 0.9661 | 0.9662 | 0.9667 |
| Recall | 0.9631 | 0.9762 | 0.9759 | 0.9798 | 0.9661 | 0.9745 | 0.9772 |
| Accuracy | 0.9702 | 0.9824 | 0.9865 | 0.9873 | 0.9868 | 0.9887 | 0.9894 |
| Specificity | 0.9658 | 0.9684 | 0.9697 | 0.9735 | 0.9728 | 0.9737 | 0.9798 |
| DICE | 0.9756 | 0.9824 | 0.9853 | 0.9871 | 0.9862 | 0.9876 | 0.9899 |

TABLE III
COMPARING DIFFERENT METHODS WITH THE PROPOSED DATASET ON THE VESSEL SEGMENTATION TASK.

| Metrics | FCN | UNet | UNet++ | Attention UNet | VNet | 3D UNet | 3D Attention |
|---|---|---|---|---|---|---|---|
| Precision | 0.8031 | 0.9257 | 0.9576 | 0.9601 | 0.9637 | 0.9677 | 0.9724 |
| Recall | 0.7671 | 0.8637 | 0.8646 | 0.8651 | 0.8111 | 0.8922 | 0.9021 |
| Accuracy | 0.7564 | 0.8033 | 0.8557 | 0.8629 | 0.8517 | 0.8723 | 0.8779 |
| Specificity | 0.7351 | 0.7990 | 0.8207 | 0.8619 | 0.8147 | 0.8787 | 0.9001 |
| DICE | 0.6243 | 0.7522 | 0.8166 | 0.6887 | 0.8453 | 0.8665 | 0.8907 |

TABLE IV
COMPARING DIFFERENT METHODS WITH THE PROPOSED DATASET ON THE LIVER TUMOR SEGMENTATION TASK.

| Metrics | FCN | UNet | UNet++ | Attention UNet | VNet | 3D UNet | 3D Attention |
|---|---|---|---|---|---|---|---|
| Precision | 0.3824 | 0.4425 | 0.4006 | 0.4356 | 0.4531 | 0.3787 | 0.3784 |
| Recall | 0.3092 | 0.3513 | 0.3423 | 0.3698 | 0.3471 | 0.3661 | 0.3545 |
| Accuracy | 0.3612 | 0.3670 | 0.3652 | 0.3667 | 0.3663 | 0.3652 | 0.3627 |
| Specificity | 0.3579 | 0.3607 | 0.3618 | 0.3629 | 0.3606 | 0.3787 | 0.3801 |
| DICE | 0.3097 | 0.3265 | 0.3693 | 0.3896 | 0.3370 | 0.4021 | 0.4133 |