Deep Supervision for Pancreatic Cyst Segmentation in Abdominal CT Scans

A research work funded by the FELIX project

Yuyin Zhou[1], Lingxi Xie[1], Elliot K. Fishman[2], Alan L. Yuille[1]

[1]The Johns Hopkins University [2]The Johns Hopkins School of Medicine

Scan & Bookmark!

ABSTRACT

Automatic segmentation of an organ and its cystic region is a prerequisite of computer-aided diagnosis. In this paper, we focus on pancreatic cyst segmentation in abdominal CT scan. This task is important and very useful in clinical practice yet challenging due to the low contrast in boundary, the variability in location, shape and the different stages of the pancreatic cancer. Inspired by the high relevance between the location of a pancreas and its cystic region, we introduce extra deep supervision into the segmentation network, so that cyst segmentation can be improved with the help of relatively easier pancreas segmentation. Under a reasonable transformation function, our approach can be factorized into two stages, and each stage can be efficiently optimized via gradient back-propagation throughout the deep networks. We collect a new dataset with 131 pathological samples, which, to the best of our knowledge, is the largest set for pancreatic cyst segmentation. Without human assistance, our approach reports a 63.44% average accuracy, measured by the Dice-Sørensen coefficient (DSC), which is higher than the number (60.46%) without deep supervision.

CONTRIBUTION

We study the problem of automatically detecting the pancreatic cyst from abdominal CT scan. This task is very challenging, and some previous work is based on the old-fashioned models such as GraphCut. We verify that the state-of-the-art deep learning techniques can be used for this purpose. Note that segmenting the cyst is different than segmenting the pancreas itself. Our technical contributions can be summarized as:

- We propose a framework which first finds the rough position of the pancreas, and then segments the cyst based on the pancreas. With a reasonable approximation, we decompose it into two stages which can be trained and evaluated individually.

Without human assistance, our framework obtained promising segmentation accuracy, *i.e.*, 63.44% over 131 pathological samples, with only 8 cases completely mis-detected.

We will try our best to release the pathological dataset for research purposes in the future.

REFERENCES

Key references are numbered as in the paper.

[6] **J. Long** *et.al.*, Fully Convolutional Networks for Semantic Segmentation, *CVPR*, 2015.

[9] **Y. Zhou** *et.al.*, Pancreas Segmentation in Abdominal CT Scan: A Coarse-to-Fine Approach, *arXiv preprints 1612.08230*, 2016.

ACKN.

This work was supported by the Lustgarten Foundation for Pancreatic Cancer Research. We thank Dr. Seyoun Park for enormous help.

THE PROPOSED APPROACH

The FELIX Project

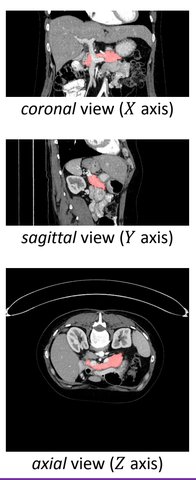
The pancreatic cancer is a major killer to humans. As the symptom is very difficult to detect at an early stage, the cancer has often spread to other organs at the time of diagnosis, causing a very high death rate (5-year survival rate is merely 7.7%).

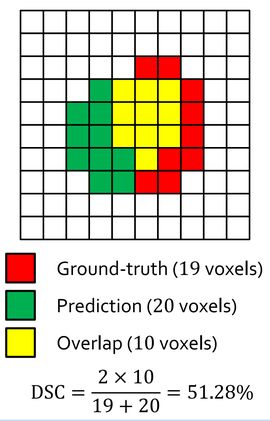
The FELIX project is aimed at applying modern approaches in computer vision to assist doctors in diagnosis at an early stage. This is challenging, as both the pancreas and the cystic region are small targets with blurring boundary, and the shape can vary a lot from case to case. To the best of our knowledge, there are **no** existing methods for automatic pancreatic cyst discovery.

Data and Evaluation

Data: a pathological dataset collected by ourselves

- 131 samples covering a set of unhealthy people
- Labeling both the pancreas and the cystic region
- Evaluation: the DSC score



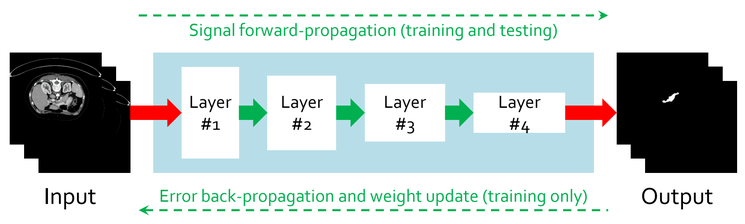


$$\mathrm{DSC} = \frac{2 \times 10}{19 + 20} = 51.28\%$$

Note: evaluation in 3D volumes!

Deep Learning Basics

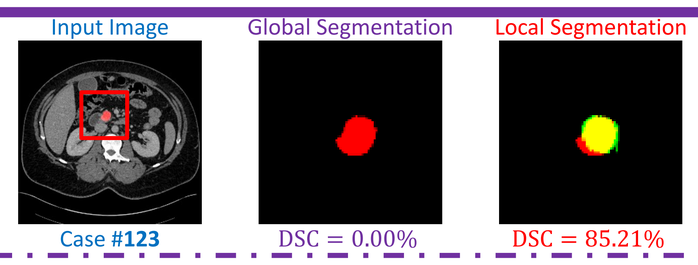
Deep learning is the state-of-the-art solution for a wide range of image-based applications. It is based on the idea that a deep network can capture very complicated distribution in image space. The basic unit of a deep network is a **neuron**, *i.e.*, a mathematical function for a specified purpose. Neurons with the same function form a **layer**, and a deep network is a *hierarchical* structure with many layers. Training a deep network involves propagating neural responses back and forth and updating network weights.



Motivation & Approach

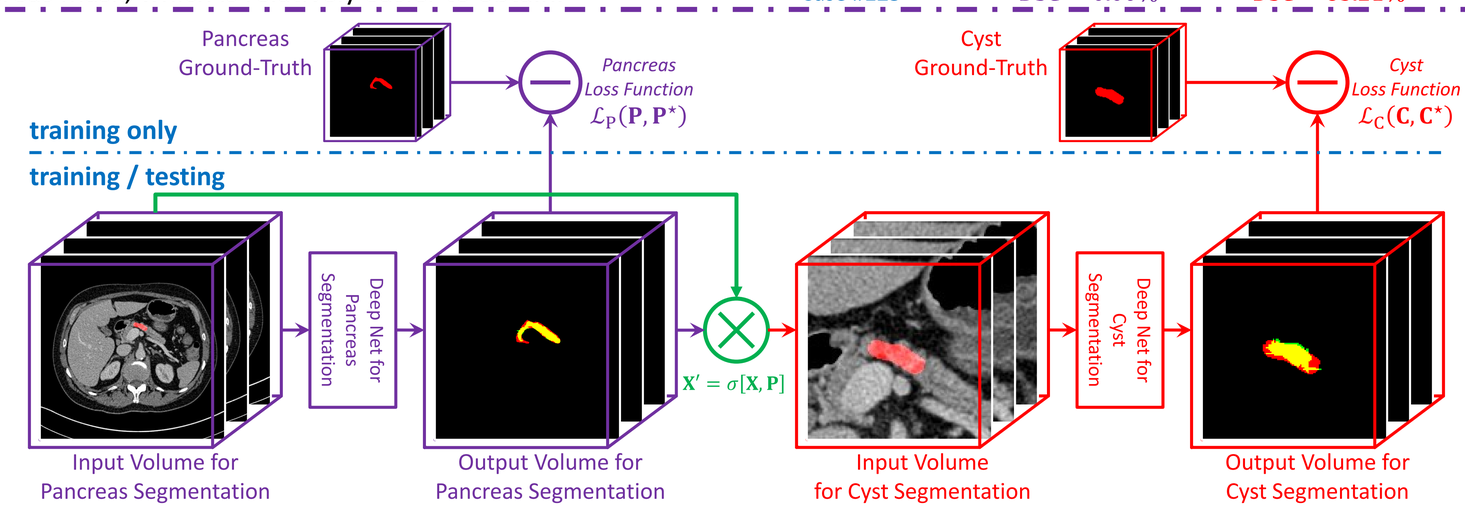
The cystic region can appear almost anywhere in the pancreas, and in very strange shape. The size of the cyst can vary from less than 1% to over 70%, with respect to the pancreas. Deep neural networks are often less accurate on such small and variable targets.

Our idea is to shrink the input region for better segmentation results (see the right figure). The reference is set to be the pancreas mask, which is relatively easier to obtain.



Technical Details

- We apply the coarse-to-fine approach [9] to both pancreas and cyst segmentation. The baseline segmentation model is the fully-convolutional network [6] with a 16-layer **VGGNet**. We believe that a stronger baseline leads to better performance.
- As the cyst is very small and easy to be mis-detected, we use 9 networks (snapshots) for individual segmentation, and then fuse these volumes by computing the union of the predicted voxel sets.
- The threshold t (for the transformation function) is set to be 15.



The two-stage framework is illustrated in the figure. The pancreas segmentation mask is introduced as extra supervision, named *deep supervision* in our paper. Two loss terms are computed, *i.e.*, $\mathcal{L}_{\mathrm{P}}(\mathbf{P}, \mathbf{P}^{\star})$ and $\mathcal{L}_{\mathrm{C}}(\mathbf{C}, \mathbf{C}^{\star})$, for pancreas and cyst segmentation, respectively. The overall loss function is therefore formulated as $\mathcal{L} = \lambda \mathcal{L}_{\mathrm{P}}(\mathbf{P}, \mathbf{P}^{\star}) + (1 - \lambda)\mathcal{L}_{\mathrm{C}}(\mathbf{C}, \mathbf{C}^{\star})$. The input for cyst segmentation $\mathbf{X}'$ is related to both the original input image $\mathbf{X}$ and the pancreas segmentation mask $\mathbf{P}$. In practice, we define the transformation function to be $\mathbf{X}' = \sigma[\mathbf{X}, \mathbf{P}]$, in which a pixel is considered in cyst segmentation if and only if it is close enough (within a threshold of t) to the pancreas segmentation mask. Under this formulation, the gradient of $\mathbf{X}'$ over $\mathbf{P}$ is 0 almost everywhere, and the two stages in our framework can be dealt with individually in both training and testing processes.

EXPERIMENTAL RESULTS

Quantitative Results

We evaluate our approach on the pathological dataset collected by ourselves. Over 131 cases, the framework with deep supervision significantly outperforms that without deep supervision, in both average accuracy and the number of mis-detected cases.

Cyst Segmentation Accuracy (%)	Mean DSC	# Miss.	Max/Min DSC
Segmentation w/ GT cyst b-box *	77.92 ± 12.83	0	96.14/24.69
Segmentation w/o Deep Supervision	60.46 ± 31.37	16	95.67/00.00
Segmentation w/ Deep Supervision	63.44 ± 27.71	8	95.55/00.00

* This is not a fair comparison, yet providing the GT b-box requires heavy human labor.

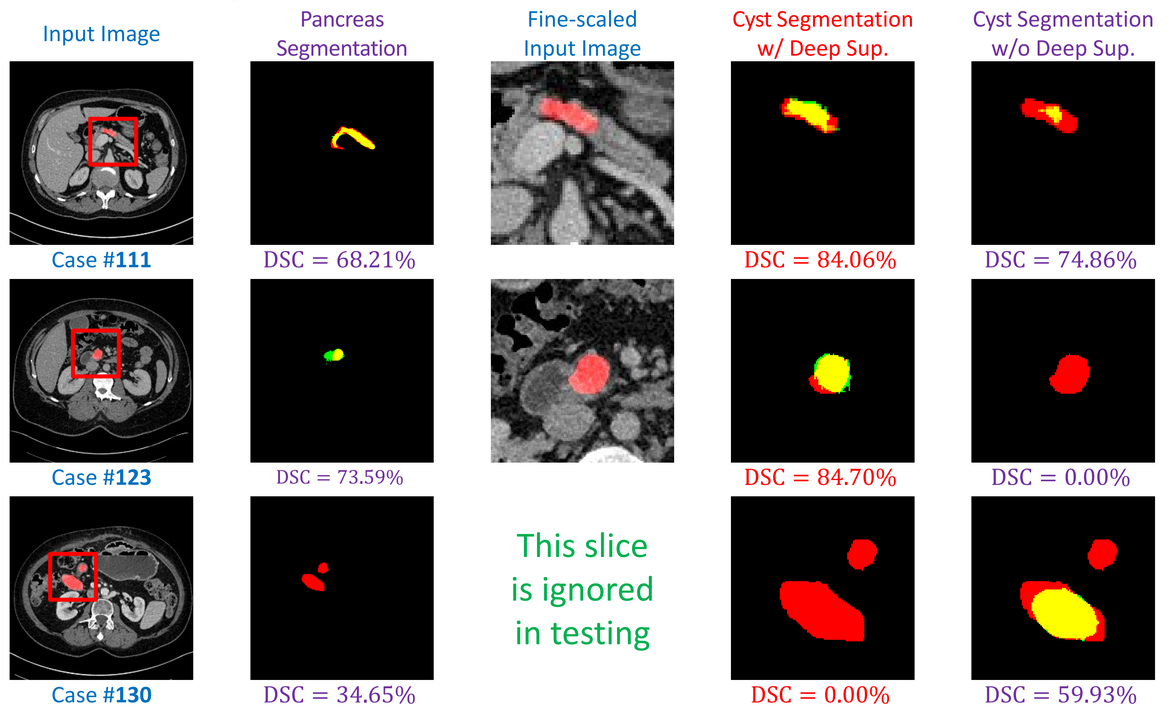
Conclusions

In this work, we present a two-stage framework to detect the pancreatic cyst from abdominal CT scan. Our approach is motivated by a simple idea, that using the rough location of the pancreas to shrink the input image for cyst segmentation. We design a reasonable transformation function, and thus make it possible to train and evaluate the two stages individually (this reduces computation overheads).

Despite the promising results, we note that this problem remains challenging, as our approach still fails on 8 cases. Meanwhile, the two-stage framework is not always working well, due to inaccurate pancreas segmentation (see the right figure).

Visualization

We compare the segmentation results without and with deep supervision in some typical cases.



In the first two cases, the pancreas segmentation mask is relatively accurate, so we can provide a reasonable input region which improves cyst segmentation accuracy. In the last case, however, the low quality in pancreas segmentation causes complete failure in cyst segmentation.