

Review: CFS-FCN (Biomedical Image Segmentation)



SH Tsang

Follow

Nov 15, 2018 · 5 min read

This time, **CFS-FCN (Coarse-to-Fine Stacked Fully Convolutional Net)** is shortly reviewed, which is used for **segmenting lymph nodes in the ultrasound images**.

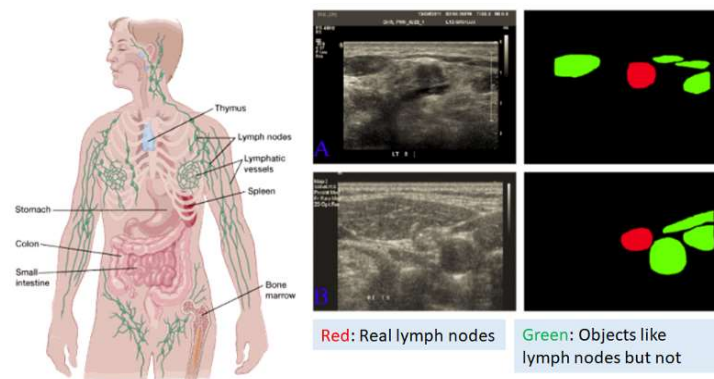
You may ask: "Is it too narrow to read about biomedical Image Segmentation?"

*However, we may learn the techniques of it, and apply it to different industries. Say for example, **quality control / automatic inspection / automatic robotics during construction / fabrication / manufacturing process**, or any other stuffs we may think of. **These activities involve quantitative diagnosis. If we can make it automatic, cost can be saved with even higher accuracy.***

This is a paper in **2016 BIBM**. It **outperforms two state-of-the-art approaches, CUMedVision1 and U-Net**. It also has the **concept of intermediate labels which assists the segmentation**. (SH Tsang @ Medium)

• • •

Lymph Nodes & Ultrasound Images



Lymph Nodes (green line) in Human Body (Left), Ultrasound images (Middle), Segmentation Results (Right)

Lymph nodes are important to our immune system. ultrasound images scanning is a kind of non-invasive scanning and it's commonly available in hospital. Based on the ultrasound image, we can have clinical diagnosis, cancer staging, patient prognosis, and treatment planning, etc.

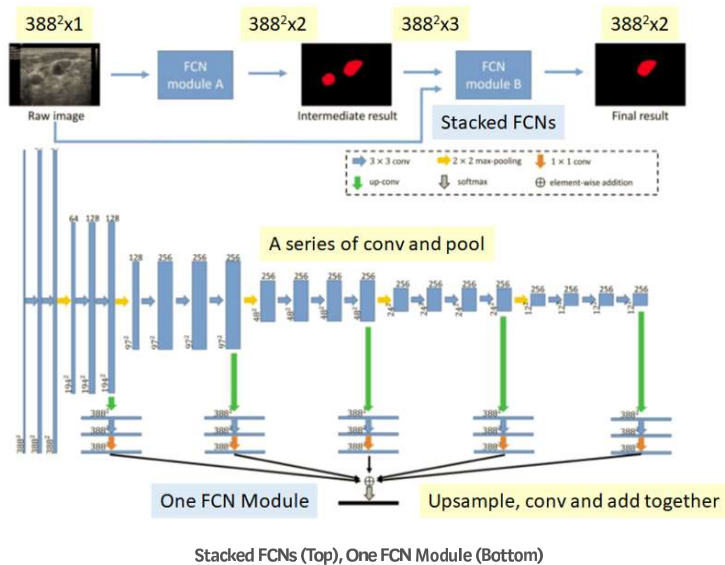
• • •

What Are Covered

1. **Coarse-to-Fine Stacked FCN**
2. **Intermediate Labels and Training Strategies**
3. **Results**

• • •

1. Coarse-to-Fine Stacked FCN



1.1. Stacked FCNs

1. First, a $388 \times 388 \times 1$ (width \times height \times color plane) gray-scaled ultrasound image is acted as input to FCN module A.
2. The FCN module A outputs the $388 \times 388 \times 2$ (width \times height \times output labels) intermediate results in which it segments out both real lymph nodes and objects look like lymph nodes but not lymph nodes from the background.
3. Then this output concatenates together with the input gray-scaled ultrasound image ($388 \times 388 \times 3$) and input to the FCN module B.
4. The FCN module B outputs the $388 \times 388 \times 2$ (width \times height \times output labels) final results in which it only segments the real lymph nodes.

1.2. A FCN module

The FCN module, as shown above, actually is similar to the one in [FCN](#) or [CUMedVision1](#).

1. A series of convolutions and max pooling to extract the features.
2. Each layer before max pooling are used for unsampling, convolution, then fused (element-wise added) together to get the results for each FCN module.

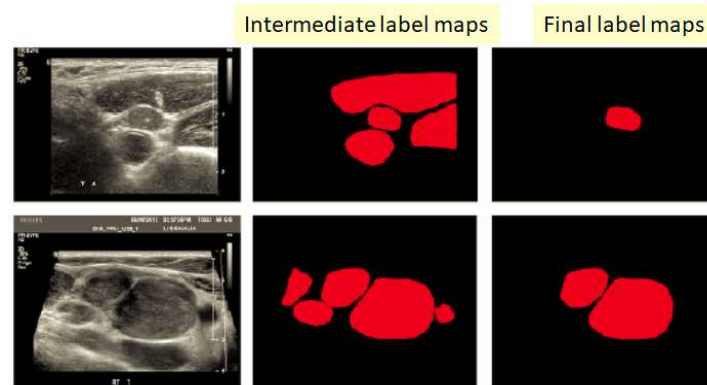
Except that the number of channels input to the FCNs are different. For FCN module A, it only got 1 channel, while for FCN module B, it got 3

channels.

. . .

2. Intermediate Labels and Training Strategies

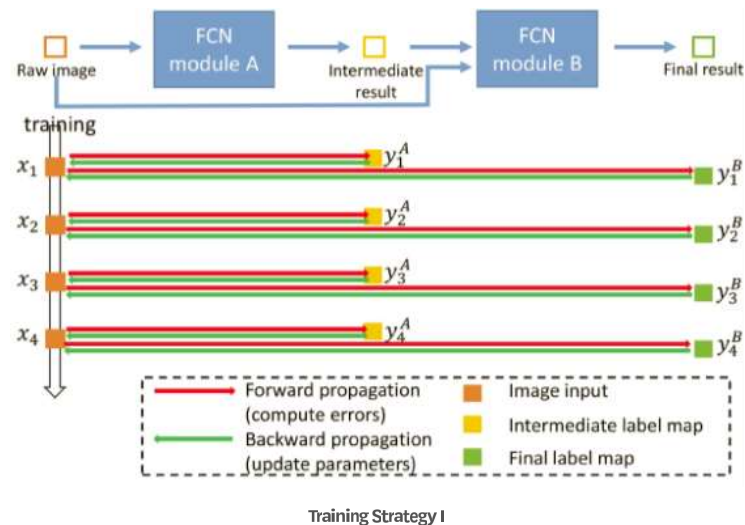
2.1 Intermediate Labels



Input Images (Left), Intermediate Label Maps (Middle) and Final Label Maps (Right)

Besides annotating the final label maps, **we also need experts to annotate the intermediate label maps**, in order to have the training of FCN module A.

2.2. Training Strategies



Different training strategies are tried.

Naive stacked FCN

- Train the whole network without using the intermediate label maps.

Training Strategy I

- Train FCN A and FCN B at the same time alternatively using the same image data as in the figure above.
- Both modules have influence to each other in the training.

Training Strategy II

- Train FA using the intermediate label maps.
- Then train FB using the final label maps.

Training Strategy III

- Train FA using the intermediate label maps.
- Then fixed FA and train FB using the final label maps.

As we can guess, Training Strategy I is the best one here.

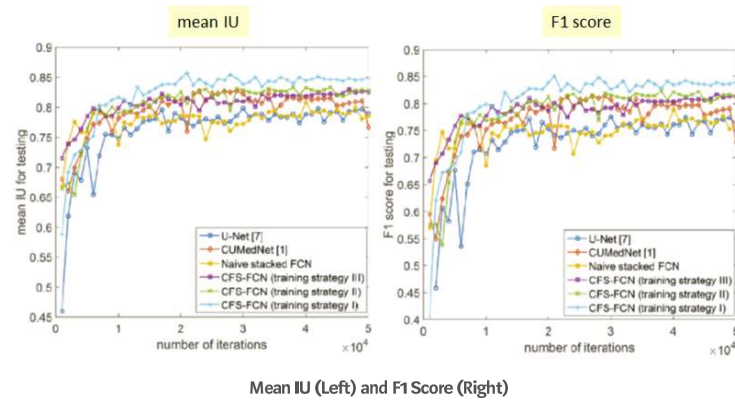
. . .

3. Results

3.1 Dataset

- 80 ultrasound images
- Two-fold cross validation is used, in which 2 sets are split.
- Train: set 1; Test: set 2
- Train: set 2; Test: set 1
- And the mean IU and F1 score are calculated

3.2 Mean IU and F1 Score



U-Net and Naive Stacked FCN have similar performance.

CUMedNet (i.e. CUMedVision1), CFS-FCN (Training Strategy II) and CFS-FCN (Training Strategy III) have similar performance.

CFS-FCN (Training Strategy I) has the best performance for both mean IU and F1 score, which is the curve at the top for each graph.

| | Mean IU | F1 score | Memory cost |
|------------------------------------|---------|----------|-------------|
| U-Net [7] | 0.798 | 0.775 | ~5200 MB |
| CUMedNet [1] | 0.816 | 0.798 | ~950 MB |
| Naive stacked FCN | 0.799 | 0.778 | ~1920 MB |
| CFS-FCN (training strategy III) | 0.830 | 0.818 | ~1920 MB |
| CFS-FCN (training strategy II) | 0.833 | 0.822 | ~1920 MB |
| CFS-FCN (training strategy I) | 0.851 | 0.843 | ~1920 MB |
| CFS-FCN (training strategy I) + BR | 0.860 | 0.858 | ~1920 MB |

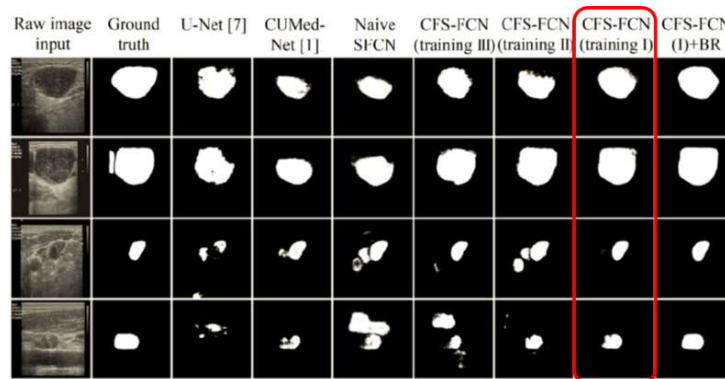
BR: Boundary refinement for concave places

Mean IU, F1 Score and Memory Cost

CFS-FCN (Training Strategy I) has 0.851 mean IU and 0.843 F1 score. With BR (i.e. Boundary Refinement to fill the concave places), which is a post-processing step making it not end-to-end learning, obtain a better mean IU of 0.860 and F1 score of 0.858.

CFS-FCN needs double memory of CUMedNet (CUMedVision1) since two FCNs are stacked.

3.3. Some Visual Segmentation Results



Some Visual Segmentation Results

. . .

Though the training losses from intermediate label map and final label map can be combined so that FCN A and FCN B can be trained together, they use a simple idea that they just need to stack the FCNs, and a better result is obtained. Thus, they don't need to think of a new FCN architecture to tackle the problem or improve the result. But the downside is to prepare the intermediate label maps. In terms of the number of labels, CFS-FCN needs double number of labels (200%).

. . .

References

1. [2016 BIBM] [CFS-FCN]
Coarse-to-Fine Stacked Fully Convolutional Nets for lymph node segmentation in ultrasound images

My Related Reviews

[U-Net] [CUMedVision1] [CUMedVision2] [FCN] [DeconvNet]

