Weekly Report

Haozhe Feng

2018.4.8

What I have done this week

Last week, I complete the semantic segmentation task for lung nodule, and Professor Danny Chen suggests that I should **formulate some specific research problem and research targets** based on this staged progress. I remember that Professor Danny Chen discussed with me on 3 topics on 16th March, and the 3 topics are :

- Using the adversarial strategy to criticize the raw model (in this instance the model is a semantic segmentation model). For example, we can use adversarial strategy to use the unannotated images to improve the performance of the raw model and learn from the unannotated images.
- Using the adversarial strategy to do suggest annotation. For example, in 3D cubes,there maybe a part of prediction which is not predicted well and need some criticize. If we can build a adversarial model which can criticize the raw model and predict some specific parts which need some more focus and stronger label, it will be very meaningful. This ideal is an extension of the original suggest annotation ideal.
- Using some cluster method to classify the unannotated dataset and do suggest annotation in one step.

This week I try to figure out some research targets based on this 3 topics. Professor Chen reminds me that I could propose some problems based on the ShangHai Hospital's dataset, so I want to propose some interesting problems from this perspective. I do these attempts this week :

Since the dataset from ShangHai Hospital is a fresh unannotated dataset with lung nodule, I think it is a good dataset to do adversarial training just like paper Deep Adversarial Networks for Biomedical Image Segmentation Utilizing Unannotated Images, so I **try the same adversarial strategy** in this paper to see if it can improve the performance of the semantic segmentation network. However, this strategy do not work well on the 3D dataset, and the test accuracy isn't improved with this adversarial strategy. I

recheck the dataset and find some problems which may be the reason for this problem.

In this dataset, I find the nodules are all like this type :
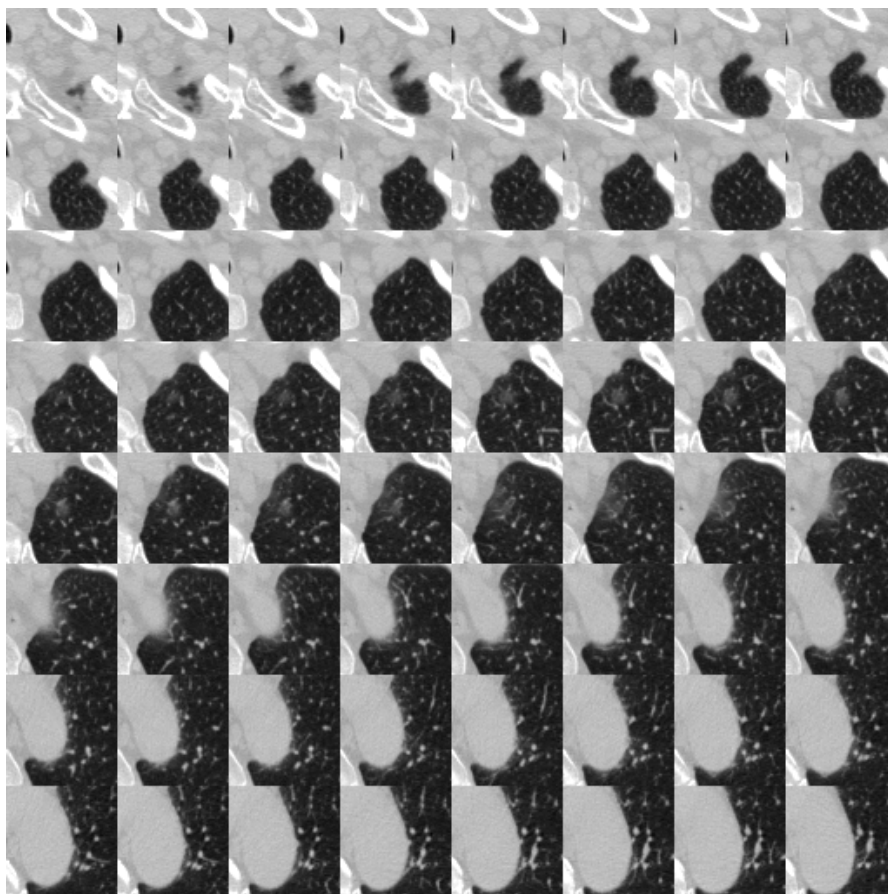
And the nodule label is :
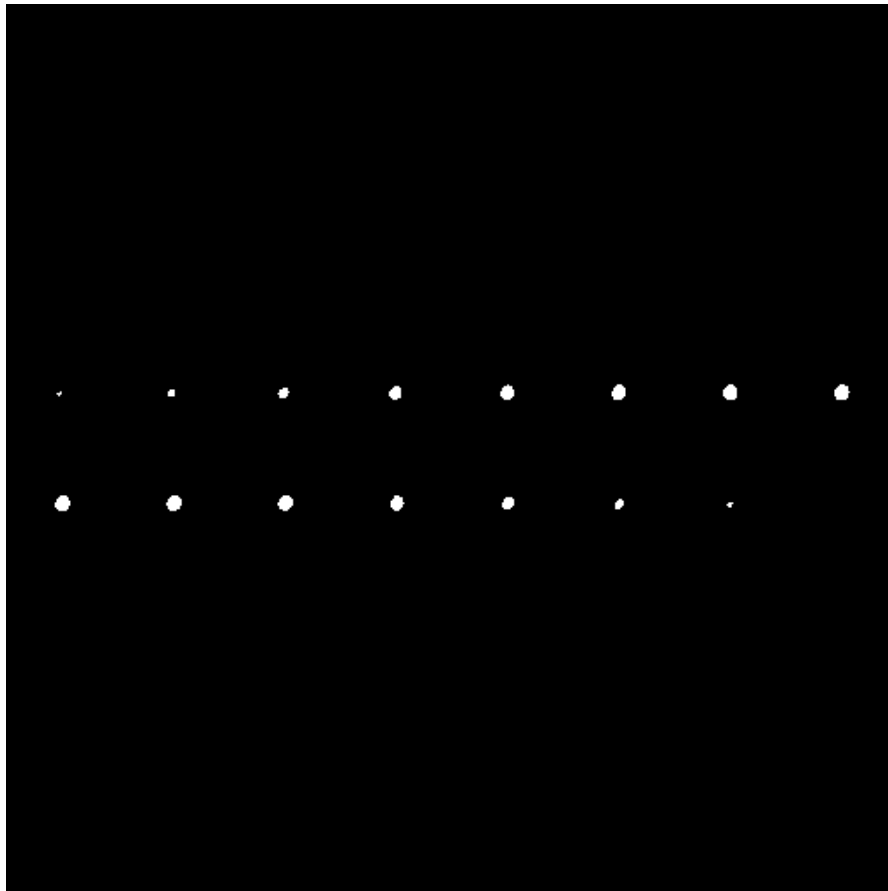
It represents that the morphological characteristics of the dataset most belong to the ground glass nodule, which means that if the test dataset don't have much ground glass nodule, it may not improve the performance in test dataset. Actually I also find that in LIDC dataset, there are only a few ground glass nodule, which means the adversarial strategy may not reflect on the test dataset well.

So it is a problem **when we meet a branch of dataset consist of many different types**, some of which may also exists in the train dataset, but others may be new to our dataset.If we can **know which data has the feature that never appears in the raw dataset**, it may be useful that we can precisely know which data can be predict better, and which data need the adversarial strategy to let the system learn its features.

I'm not sure this problem is of some value, but I think **the cluster method for dataset** is very useful not only for adversarial strategy but also for the suggest annotation in topic 3, so I do some work on this method this week.My basic ideal is to train a deep auto encoder-decoder to encode the input image into an 100-dimension vector and use this encode feature vector to do some cluster method. I have build a deep encoder-decoder network using the dense block in the semantic segmentation network, and the result is like this:
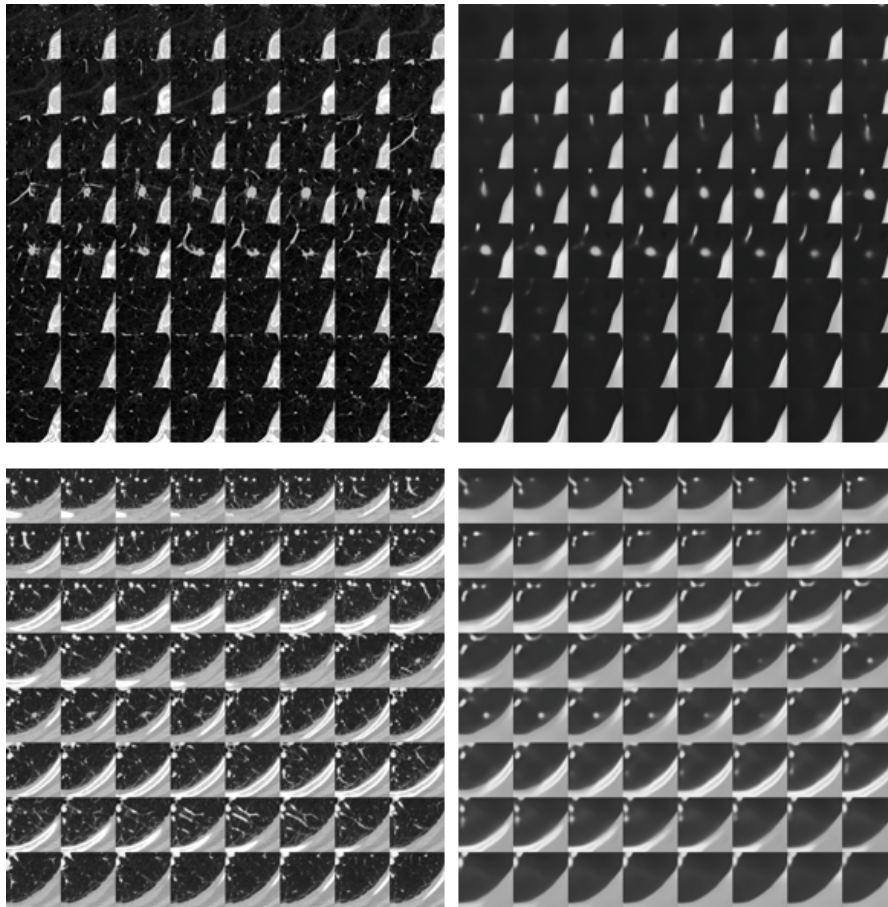
A good news is that my network filter the blood vessel structure and retain the nodule information, which may means the feature vector can be used in the cluster method.

What I plan to do next week

Next week, I may do some research on cluster methods and implement a cluster method using these encode feature vectors. I will test the cluster result in suggest annotation strategy and the adversarial strategy mentioned above.

However,I'm not sure this topic I proposes is a good research problem. If I can get some suggestions, It will be very helpful.