近期工作汇报

Danny Chen教授所提出的值得探索的科研课题

1. 3D语义分割问题与对抗策略

   考虑用对抗网络来对所产生的结果进行批评，从而通过批评策略增加优化分割结果，或者通过批评策略精确定位是哪个局部区域产生了不好的结果，并对这个局部区域进行推荐标注优化（局部推荐标注策略）

   这个方法在未来面对非常大的数据集的时候(如(1000,1000,1000)这种规模的3D影像数据集的时候)将会起到非常重要的作用。因为这个时候医生已经很难在整体上进行金标准评估了，而局部推荐标注策略可以解决这个问题。

2. 特征聚类算法与推荐标注策略

   我们可以采用一套方法尝试对未标注数据进行聚类，利用聚类结果来尝试直接用一步来做推荐标注系统，或者采用不同的模型来处理基于聚类结果的不同特征，即可以从聚类结果开发出局部自适应的模型出来。即如果我们将原数据集分为A,B,C,D4类，并对每一类采用适当的模型进行训练，此后我们进行预测的时候可以对输入的数据进行类别匹配，然后用相应类的模型来进行预测。

3. 对于标注的鲁棒性问题

   另外一个值得考虑的问题是医学影像标注的鲁棒性问题。以心理疾病为例，有可能10个医生的诊断都不一样，如何选取ground truth,如何定义ground truth都是非常值得思考的问题。我们可以设计一个模型，这个模

型综合不同的诊断，最终能够超越单个医生的表现，比如说十个医生给了诊断，我们可以考虑采用一种方法将这个诊断投射到几何空间上去，找到这些诊断的中心来代表更好的综合一件。

以上是Danny Chen教授所提出的非常有价值的课题。

2018.3 - 2018.4.15 期间所做的科研工作汇报

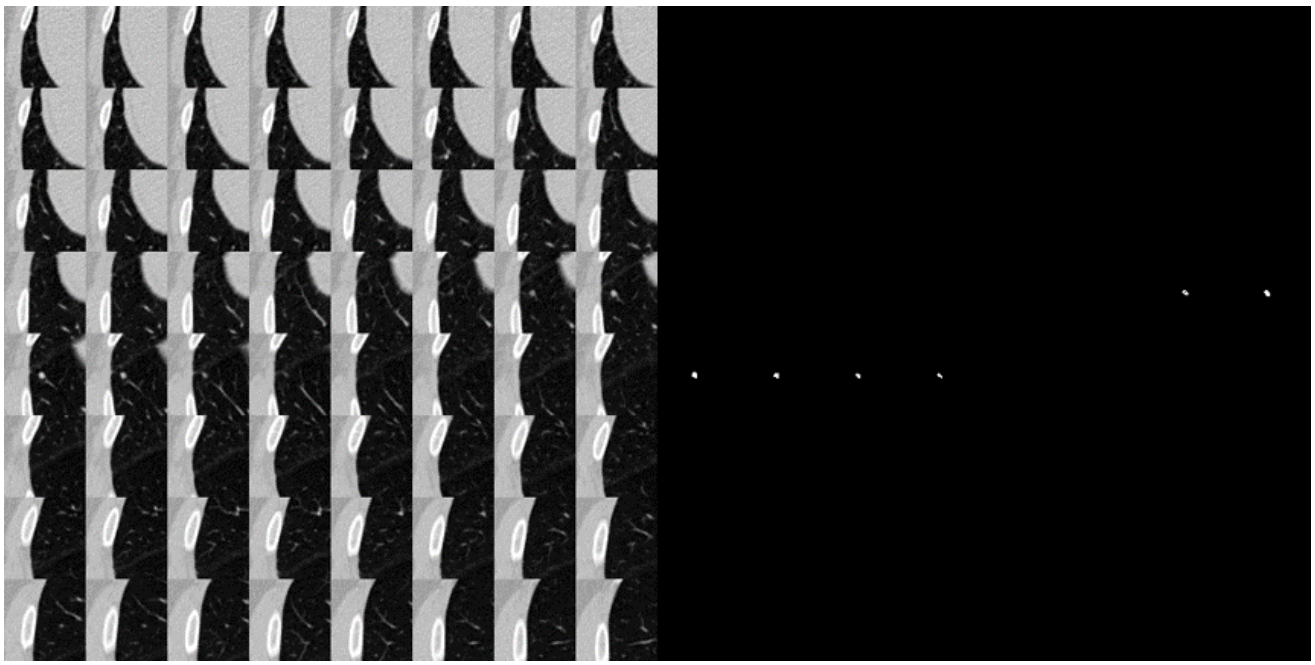
因为这些课题都是很有价值的课题，而我并不知道哪一种课题能够做出来，因此在这段时间内我主要是进行一些前期的工作，并对课题都进行了探索（因此显得没有什么方向感），在探索中我在特征聚类课题上做出了一些东西，之后就一直抓住特征聚类课题了。

大四专业琐事

花了一部分时间在毕业设计(开题报告，答辩）以及大四下的必修学分上(讲座，形势与政策等)。

数据库构建

3月份完成了整个数据清洗与数据库构建工作，数据集如图所示：

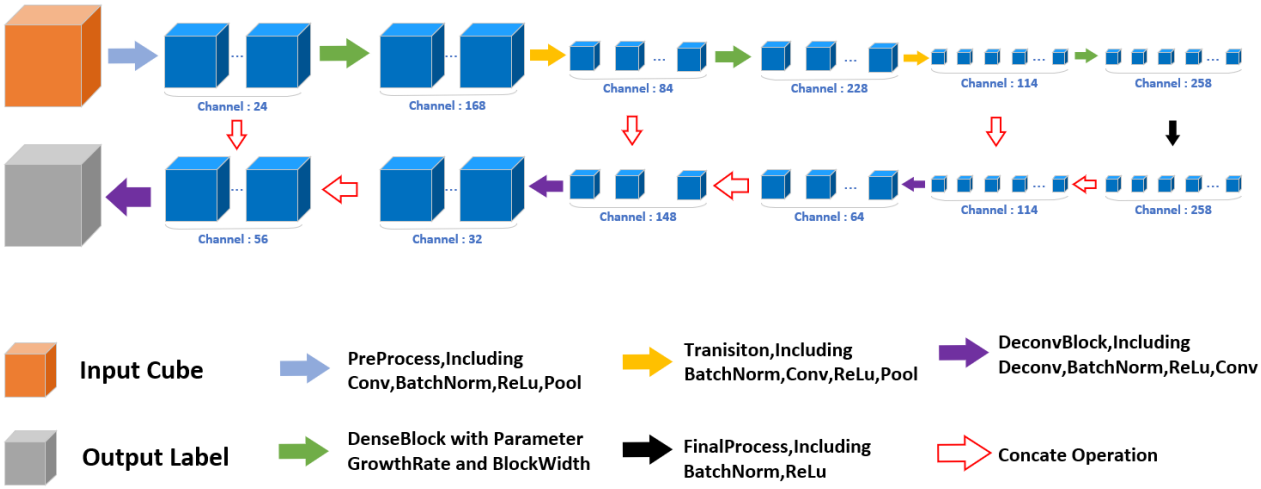


我主要用这个肺结节分割数据库作为研究实例，这个数据库是做3D语义分割或3D Object Detection相关实验的很方便的数据集。

分割模型构建

先用3D U-Net进行了尝试。U-Net的分割效果很好，但是收敛很慢(训练一次大概需要21个小时才能收敛)，而且非常耗内存（需要6.6G显存，一块GPU显存10G,因此都无法在一个GPU上跑2个网络）,因此U-Net无法进行对想法的快速验证。

我利用Densenet的结构设计了一个新的网络Dense Segnet，如图所示：



这个网络只需要花原来一半的内存(3.1G)以及5个小时的训练时间就能收敛。3D U-Net的分割结果能够达到平均 $Dice:0.763, VoE:0.37$，而Dense Segnet能够达到平均 $Dice:0.757$, $VoE:0.373$，两个网络结果基本等同。这个网络一块GPU可以跑3个，这就给我多次试验并对想法快速验证提供了相应的帮助。

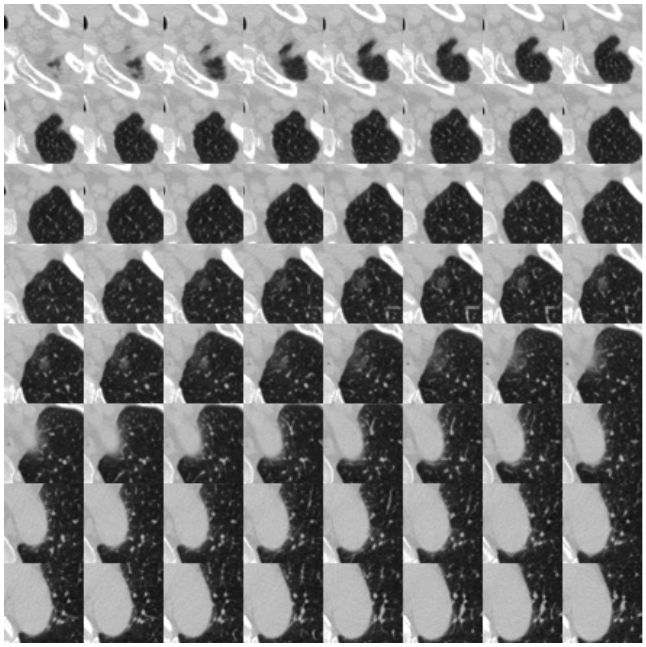
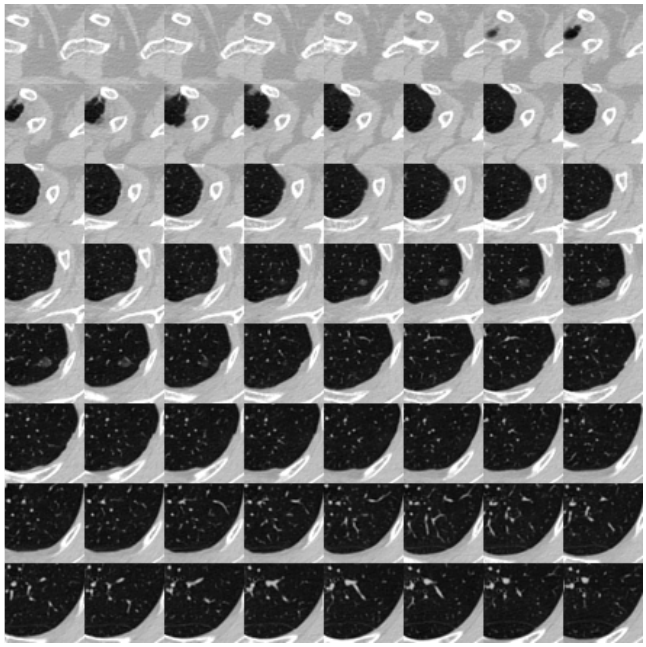
尝试了论文Deep Adversarial Networks for Biomedical Image Segmentation Utilizing Unannotated Images的训练策略

基于Danny Chen教授所提出的用对抗策略来辅助训练或者用对抗策略来提升分割效果或者做局部Suggest这个课题，同时我们有一批来自上海肺科医院的未标注数据集，因此我考虑用这批数据集来对原数据集做批评，因为没有什么思路于是我就想先尝试把论文的策略在我的数据集上复现一下。

我是用与上面的Densenet相似的结构构建的Adversarial Net的结构，未标注的数据集我选用的是上海肺科医院的数据集，因为这个数据集只给出了结节的中心坐标，没有给出Bounding Box或者是精细分割标注，而且独立于LIDC数据集。但是它有额外的标注是关于每个结节的分类(原位腺癌，浸润腺癌)，我想可能这个信息也可以想办法用在对抗策略中。因此我用LIDC的数据集作为有标注的数据集，上海肺科医院的数据集作为无标注数据集，采用论文中的策略对我的数据集进行了训练

训练的结果是整个训练过程中所用的trick很多，比如如果一开始分割网络训练的步数太多，那么就会导致批评网络没办法得出很好的批评结果，这里的超参数，以及交替训练的策略步数需要很多次实验然后确定。同时这里的对抗策略并没有如预期一般对网络最后的结果与泛化性有提升。

我对比了LIDC与上海肺科医院数据集的结果，发现一个可能的原因。LIDC的数据集里面有各种形态的肺结节，但是它里面如下图所示的毛玻璃形态的结节数目非常少：

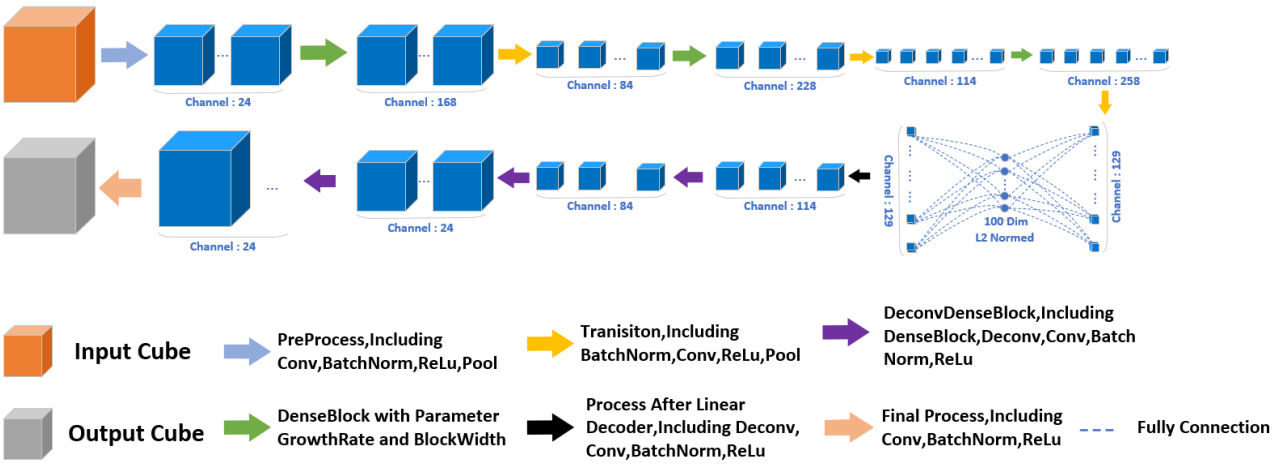
我提出的一个**可能的猜想**是可能肺科医院的肺结节的形态特征的概率分布与LIDC是不一样的，也就是说我用这个数据集进行对抗所额外学到的特征并不是在validation集中的特征，因此对最后的结果没有显著的提升。

以上是关于对抗策略的尝试，我大概花了8个工作日左右来构建这个对抗策略，但是因为结果并不是很好，而且基于我对这个不好的结果的一个可能的猜想，我就试图采用分类策略来解决问题，但是之后在分类策略上做出了一些成果，因此我暂时把工作重心与课题全部放在了分类任务上，做完这个课题之后也许可以重新回头审视对抗策略。

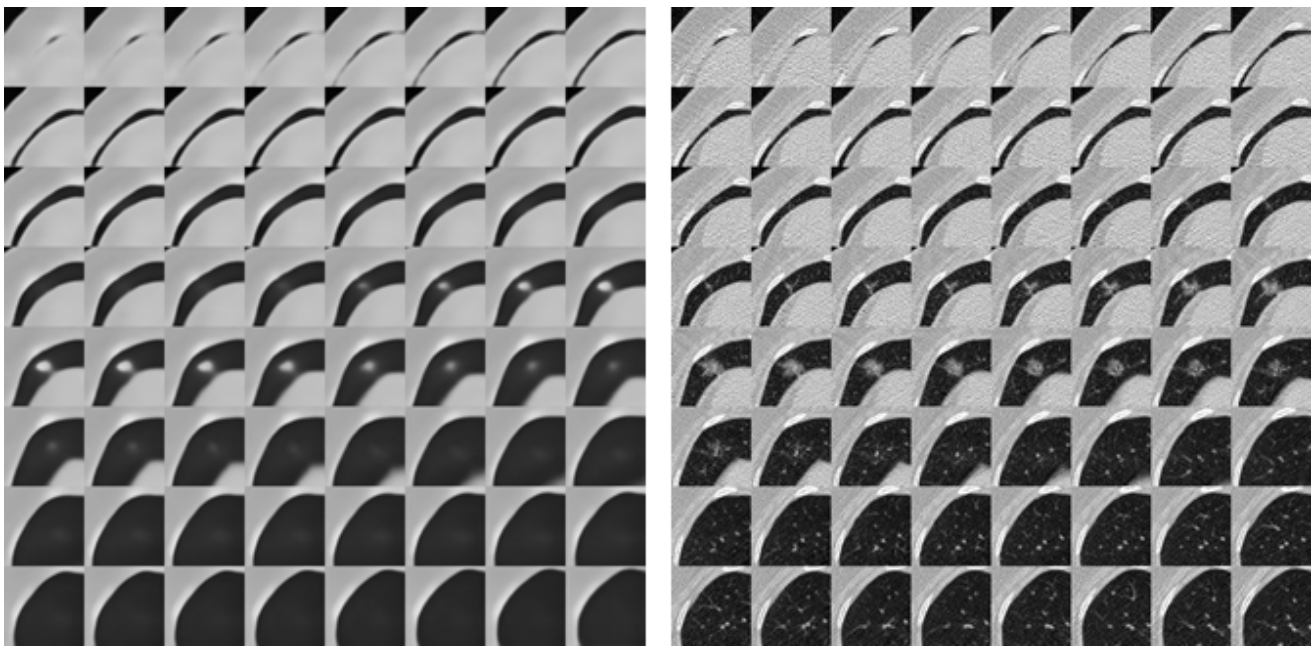
分类策略的初步尝试

我尝试采用分类策略来进行Suggest方法，即直接将输入数据采用某种方法进行分类，然后对分类结果进行进行数据筛选。

然后我设计的分类器是基于自动编码器来进行的。基本构造如下所示：



这个自动编码器与Dense Segnet结构非常相似，因为我希望自动编码器所能学习到的特征与语义分割网络所能学习到的特征是相似的。我用这个网络将原数据集编码到一个100维特征向量，最后的训练结果是这样的：



这个编码器学到了结节特征，然后起到了滤波的结果。

我采用传统聚类方法(K-Means,DBSCAN)对这些向量进行聚类，同时试验了基于深度学习的特征子空间聚类法，最后发现DBSCAN会有一些提升。我按照10：1：1的比率对原数据集进行训练，测试与验证的随机划分划分，得到的初步结果如下：

在取原训练集50%的数据时，在肺结节这个实例上密度聚类能够达到用原数据一样的结果，但是有一些随机选的策略也能达到。

在取原训练集30%的数据时，密度聚类能够达到$Dice$：0.75的平均值

在取原训练集20%的数据时，密度聚类比随机选策略有显著优越性。

随机选策略在test与val的平均dice为$0.715, 0.697$，而密度聚类策略为$0.741, 0.725$，密度聚类确实起到了一些效果。

2018.4.17 - 2018.5.6 期间所做的科研工作汇报

基于前一段时间的工作，Danny教授对我的工作提供了如下建议：

1. 我的聚类方法是否允许不同聚类之间的Overlapping

2. 对于那些只有几个数据的小类，比如孤立点问题，我的聚类方法如何实现类别均衡的。
3. 对于从聚类结果中的数据选取，比如从600个数据集中选取60个数据集，这个选取的策略可以用最小支配集的方案来选取
4. 尝试在选取原数据集60%,50%,40%的基础上画图看看最后的结果怎么样

我在接下来一段时间的工作主要是基于这些建议做的。在4.17日-4.30日我主要对以上的建议进行了探索，之后主要把时间花在写作本科毕业论文上。我主要进行了以下3个方面的改善：

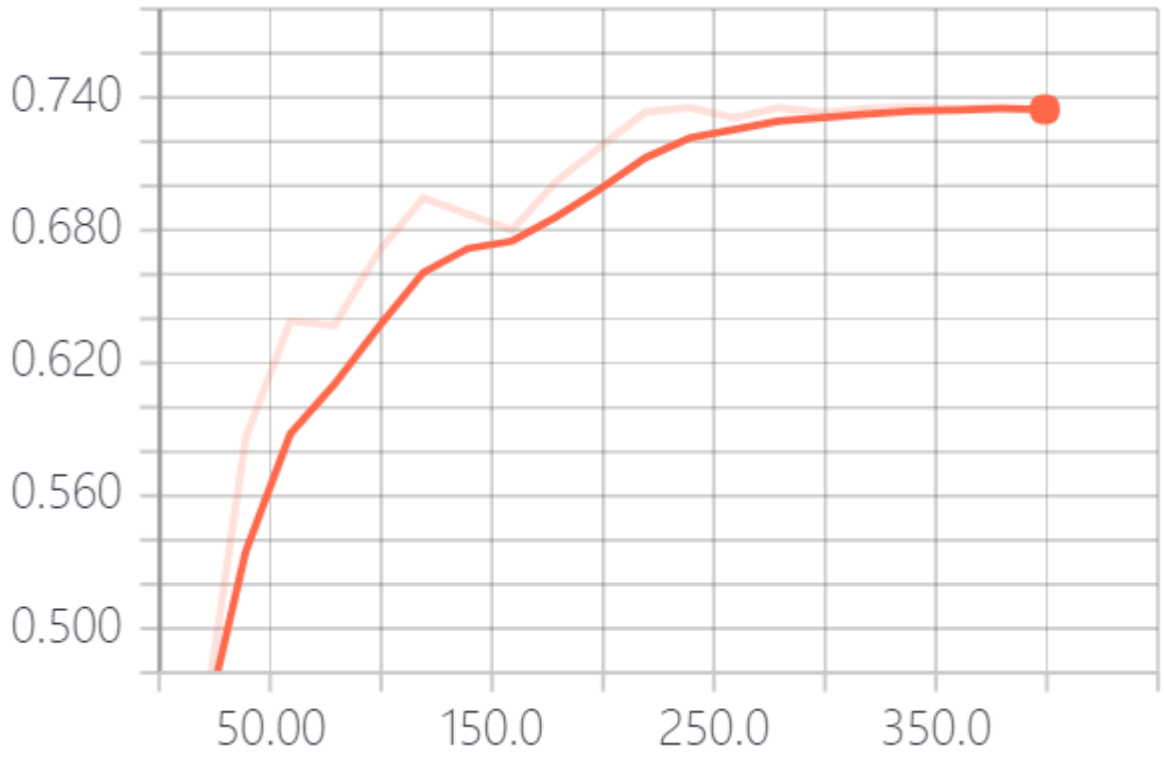
对衡量一个聚类方法的效果的改善

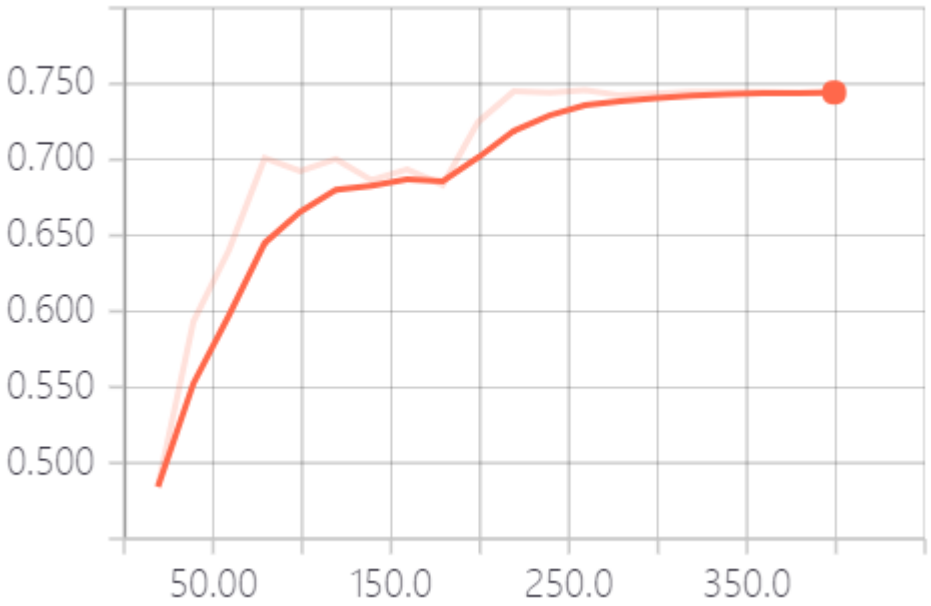
LIDC-IDRI数据集并没有给出训练集，验证集，测试集的划分，因此之前每次验证一个方法的有效性我需要用5折交叉检验，非常费时间。我想了一个新的逻辑来测试聚类方法。因为我们一共有1148个有标注的数据集，而我们聚类算法的目的是为了证明我们所挑选出的数据集能包含原数据集的大部分特征，因此我采用从1148个数据集中选取X个数据集作为我们的推荐标注数据集，用剩下的所有数据集作为测试集，如果我们的策略在测试集上表现得显著优越于随机挑选方法，那么就能证明聚类算法达到了目的。

改善Danny Chen教授提出的建议3

我选取了密度聚类与递归谱聚类结合的方法来进行挑选，即先用密度聚类将数据集划分为若干大类并找出孤立点，然后用谱聚类对密度聚类所得到的若干大类进行递归式的聚为更小的子类，直到每一个子类的数据量小于一个阈值。然后在每一个最小子类中采用Danny的Suggest Annotation论文中所采用的max k cover问题的贪心算法进行选取。

我采用从1148个数据集中选取400个数据集的方案，用以上策略进行实验，这个策略的实验结果如图所示：





随机选取策略能够达到Dice=0.736，基于我们的策略的结果能够达到Dice=0.748,比随机选取显著高出0.012。

短期工作目标

完成毕业论文，预计5.15号之前完成。

中期工作目标

研究Danny Chen教授提出的第二个建议，即实现类别均衡问题。

对原算法进行改进，我希望达到我的算法比随机选要高出0.02以上，且测试结果高于0.755。

长期工作目标

将聚类算法用于Danny 提出的其它重要课题上，重新审视对抗策略

投稿目标

暂时目标定为11月的CVPR会议，但是我不太确定自己这些基于医学影像的结果能不能投CVPR。