

# Literature Study Report

冯浩哲

2017 年 10 月 14 日

# 提纲

本周，我主要阅读了 Professor Chen 教授所推荐的 2 篇 MICCAI 会议论文，它们分别为文献 [1] 与文献 [2]。这两篇论文都来自香港中文大学的计算机视觉实验室，它们的主题都是 3D 医学图像的语义分割。

在此报告中，我将以两篇论文为基础，详细地整理三类 3D 医学图像语义分割任务（肝脏分割，心血管分割以及肺结节分割）所需要解决的主要难题，以及这两篇论文描述的方法针对什么样的问题采用何种方式进行了解决，这些解决方式又是在什么样的假设下如何起作用的。

在报告最后，我将根据上面的论文，就我们针对的问题（肺结节分割问题）提出下一步的工作方案，包括我们可以对哪些问题进行解决，以及我们可以利用论文中的哪些方法。

# 应用与问题

3D 医学图像语义分割有 3 个应用：

- 胸部 CT 图像肝脏分割
- 磁共振 (MR) 图像心脏与大血管分割
- 肺部 CT 图像肺结节分割

完成这些分割任务面临的问题可以分为四类：

- 3D 图像区别于 2D 图像的问题
- 医学图像区别于自然图像的问题
- 语义分割网络 (3D FCN) 训练问题
- 具体应用自身的问题

## 3D 区别于 2D 的问题

3D 图像区别于 2D 图像主要产生 2 个问题:

- 特征数目

3D 图像对比 2D 图像所产生的大量额外的形状特征需要用拟合能力更强的模型来拟合, 直接导致了基于人为寻找特征的方法不适用于 3D 图像分割。

- 空间信息

模型需要考虑并应用 3D 图像的空间信息, 因此在每个 Slice 上进行分割, 然后堆叠成 3D 的做法并不适用。这意味着很多在 2D 图像上表现优异的方法无法直接迁移到 3D 上。

# 医学图像区别于自然图像的问题

与自然图像相比，医学图像语义分割具有 2 个额外难题：

- 边界模糊

在自然图像中，如果我们需要进行语义分割（例如人与背景分割），那么目标区域与周围区域往往具有明确的分界。但是医疗图像中，目标区域与周围区域具有在物理特性，颜色以及结构上连续渐变的特点，这导致目标区域的边界本就是模糊的。但是监督训练中使用的 Ground Truth 的边界往往来源于经验丰富的影像师的主观判断，这会导致模型学到的边界特征只是对于某个医生的主观判断的拟合。

- 数据量少

3D 医疗图像的精确标注需要医生对数十张 Slice 进行准确标注，因此数据获取成本大，相应地可用的 Ground Truth 数目较少，一般一项分割任务的 Ground Truth 为 10 个左右。

# 语义分割网络 (3D FCN) 训练问题

FCN 是当下处理语义分割问题表现最好的模型，以上两篇论文也以 3D FCN 为主流架构。FCN 训练过程中的主要问题有 4 点，其中，参数数目庞大，优化收敛困难以及容易过拟合这三个问题是神经网络训练的普遍问题，但是由于 3D 图像本身的特殊性，这些问题在 3D FCN 训练过程中更为突出。

但是就模型本身而言，还有一个值得解决的模型设计问题：

- Patch-Based 方法缺陷

由于医学 3D 图像体积过大，因此一般采用 Patch-Based 方法从原图像中切割出小的 patch 进行分割，然后再进行拼凑。然而，Patch-Based 方法缺陷在于有可能选的待分割 patch 只截断了目标区域的一小部分，这就要求模型对目标区域位置不敏感，而对于目标区域的部分子域区别能力也较强。同时原图像有些部分被多次覆盖，重复计算问题严重。

## 具体应用问题

具体到应用方面，不同应用所侧重的特性也不同。因此，针对性地选取方法以解决当前应用所侧重的特性很重要。针对上面提到的三个应用，我们通过作应用-特性对比表格来进行比较如下：

	应用			
		胸部 CT 图像 肝脏分割	MR 图像心脏 与大血管分割	肺部 CT 图像 肺结节分割
特性	目标区域所占比例	😊	😊	😞
	目标区域位置不变性	😊	😊	😞
	边界与周围组织粘连程度	😊	😊	😞
	分割精确性	😊	😊	😞
	患者间差异程度	😞	😞	😊
	分类任务类型	单目标二分类	单目标多分类	多目标二分类
符号注解：		😊：程度大	😞：程度中等	😞：程度小

# 文献 [1] 方法分析

文献 [1] 针对胸部 CT 图像肝脏分割与磁共振 (MR) 图像心脏与大血管分割两项任务，主要解决以下 4 个问题：

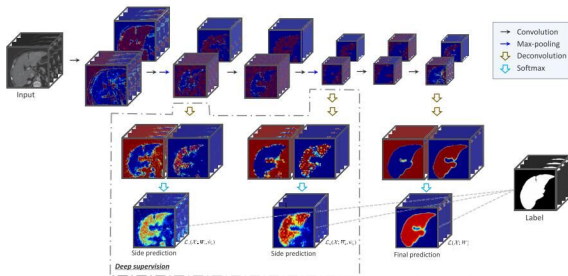
- 数据量不足

该文章采用 FCN 的结构来解决数据量不足的问题，文章提出的理由是，FCN 网络输出与输入尺寸相同的 Per Voxel 预测，这相当于在每一个体素上都得到了预测误差。使用基于每个体素误差的误差传播相当于增加了数据量，从而能减少过拟合的危险。



# 文献 [1] 方法分析

- 梯度消失与梯度爆炸  
文章提出了一种 Deep supervision 结构来保证梯度能有效传递到每一层。如图所示:



在每次使用池化后对池化输出使用反卷积与 Softmax 操作来生成与输入相同尺寸的辅助输出，如图中进行了 2 次池化操作则得到两个辅助输出。最后我们将网络输出与这些辅助输出的加权和作为最终的预测，并在此预测的基础上进行参数迭代。同时作者给出的加权策略为先赋予辅助输出较大的权值以加速网络开始的优化，并在每一轮迭代后让辅助输出权重指数衰减，使得网络输出在预测中占主导地位。文章同时通过使用 DS(Deep Supervision) 前后的 Loss 曲线证明这种结构在训练初期能让误差快速下降，从而加速收敛。

## 文献 [1] 方法分析

- 使用 CRF(Conditional Random Field) 进行边界优化

针对边界模糊的问题，作者采用 CRF 方法进行了轮廓优化。CRF 方法多被用于多特征的分类任务中，其核心思想是预先给出很多决定标记的特征 (例如词性标记任务中我们给出规则为动词后面的词性为名词)，然后使用带标记数据去训练这些给定特征在分类任务中的权重，从而获得合适的综合分类标准。

关于 CRF 的介绍可以参考 “Introduction to Conditional Random Fields” (<http://blog.echen.me/2012/01/03/introduction-to-conditional-random-fields/>)。

但是文章也指出，该方法仅在胸部 CT 图像肝脏分割中取得了显著有效的成果，而在磁共振 (MR) 图像心脏与大血管分割中没有作用。同时我们也注意到，CRF 方法仍然需要人为寻找大量初始特征，同时它仍然是基于医生所给出的主观认定的边界进行学习的算法，并未真正解决模糊边界的问题。

## 文献 [1] 方法分析

- 解决 Patch-Based 方法在重合区域的问题

文献使用了 Patch-Based 方法进行预测，针对重合区域部分被多次预测的问题，文献提出的解决方法是使用投票法决定重合区域的最终预测结果。

# 文献 [1] 方法分析

此外，文献 [1] 还详细描述了判断分割结果好坏的 7 种度量方式和各自的侧重点，并将三种传统方法，没有使用 DS 的 3D-FCN 与使用了 DS 的 3D-FCN 在不同度量方式下的表现进行了对比，从而详细叙述了加入 DS 的 FCN 结果优越性的侧重点。同时，文献对于 3D 医学图像语义分割问题的背景与相关工作进行了综述式详细叙述，这些都是该文献值得阅读的地方。

## 文献 [2] 方法分析

文献 [2] 主要针对磁共振 (MR) 图像心脏与大血管分割这一问题，继承了文献 [1] 中的 Deep Supervision 结构并引入 Densely Connection 结构，主要解决了以下两个问题：

- 深层 3D-FCN 的参数优化问题

Densely Connection 结构是 2017 年 CVPR 会议的 Best Paper 所提出来的结构，它的基本逻辑在于强调参数复用，在每个 DenseBlock 中下一层使用前面所有在 Block 中的层数的特征，从而增强输入信息流，使得每个 Block 中的参数能够迅速收敛。同时，数个 DenseBlock 进行连接从而扩宽网络深度。采用 Dense Connection 的网络收敛迅速，训练时间大大缩短。

## 文献 [2] 方法分析

- 梯度消失与梯度爆炸问题

本文档继承了文献 [1] 中的 Deep Supervision 结构，同样在每一个池化层后增加相应的反卷积运算来得到辅助输出，利用网络输出与辅助输出的加权作为最终输出，同时与文献 [1] 采用相同的策略，即辅助输出的权重随迭代次数增加指数衰减。

## 文献 [2] 方法分析

文献 [2] 是文献 [1] 利用了 Dense Connection 结构的一个改进，它展示了如何将 Dense Connection 成功运用于 3D-FCN 以解决训练时间的问题，同时在 Dense Block 之间保留了 Deep Supervision 的结构并通过对比证明了这一结构不仅能增加收敛速度，还能提高模型的分辨能力。文献展示了一个值得借鉴的方法融合的例子，具有很大启示。

# 启示

以上两篇文献为我们提供了丰富且具有意义的 3D-FCN 实践经验，揭示了 3D-FCN 的优越之处与缺点，同时文献中用以解决收敛性与梯度问题的方法应纳入我们的模型结构。

但是，我们同时要注意到，两篇文献的实例，胸部 CT 图像肝脏分割与磁共振 (MR) 图像心脏与大血管分割与我们的肺结节分割在分类任务类型，目标区域位置不变性与目标区域所占比例等方面具有很大的差异，我们在应用过程中应将差异纳入考量。

特 性	应用			
		胸部 CT 图像 肝脏分割	MR 图像心脏 与大血管分割	肺部 CT 图像 肺结节分割
	目标区域所占比例	☺	☺	☹
	目标区域位置不变性	☺	☺	☹
	边界与周围组织粘连程度	☺	☺	☹
	分割精确性	☺	☺	☹
	患者间差异程度	☺	☺	☺
	分类任务类型	单目标二分类	单目标多分类	多目标二分类
符号注解：		☺：程度大	☹：程度中等	☹：程度小



# 启示

同时，两篇论文仍然留下了 3 个亟待解决的典型问题，我认为这可以是我们接下来的重点方向。

- 模糊边界问题

两篇论文对于医疗图像语义分割的模糊边界的问题并没有提出有效的方法，仍然侧重于认为医生的主观标注就是金标准，模型仍然停留在对主观分割边界的学习。我们可以结合 Danny Chen 教授暑假所提供的关于神经元细胞的 Fuzzy 分割方法来尝试在 3D 层面解决问题，基本思路还是由网络来给出大致的拓扑结构，由图分割等基于 cost 函数的方法并结合相应的医学知识来计算出分割边界。同时我们注意到医生对于边界的划分有激进风格也有保守风格，因此所设计的 cost 函数应该可以通过超参数的修改来实现边界分割处的微调。

# 启示

- Patch-Based 方法问题

Patch-Based 方法弊端在于有可能选的待分割 patch 只截断了目标区域的一小部分，这就要求模型对目标区域位置不敏感，而对于目标区域的部分子域区别能力也较强。同时原图像有些部分被多次覆盖，重复计算问题严重。对于重复覆盖问题，两篇文献只是简单采用重复覆盖部分的标签通过投票决定的方法进行处理，但是对于其它问题并没有深入讨论，因此如何改进这种方法或是使用一种新方法是一个值得研究的问题。

对于这个问题我仍然没有一种好的思路，我的初步想法是参考 Faster-RCNN，将区域选择与神经网络相融合，或者是采用某些区域筛选方法进行预处理，但是仍然没有形成体系。

# 启示

- 现有数据标签的 Ground Truth 问题

与文献中少而精的数据不同，我们现在所拥有的数据量很大，但是标签并不精，往往会出现一个结节中有部分 Slice 没有标记的情况，同时边缘部分的划分也并不精细，而这种数据是数据获取的常态。如何充分利用此种类型的数据进行网络训练，对于数据中空缺的标签应该如何处理（例如空缺标记部分权重置 0），我们想在这个方面利用现有的数据进行一些方法尝试，而具体思路仍然要进行文献阅读。

# 目标

总结上文内容，我认为我们下一步应该利用这两篇文献所提出的结构与他们的网络训练经验，结合相关文献改进我们的网络结构，尽快得到一些初步结果，并尝试使用 Suggest Annotation 的方案。

同时，对于上文提到的模糊边界问题，现有数据标签的 Ground Truth 问题以及 Patch-Based 方法问题，我认为我们应当以模糊边界问题为开始尝试解决。考虑到肺结节分割对于边界要求并不精细，而给出的数据边界标准也不统一，该数据集很适合作为解决模糊边界问题的背景。



Q. Dou, L. Yu, H. Chen, Y. Jin, X. Yang, J. Qin, and P. A. Heng.  
3d deeply supervised network for automated segmentation of volumetric medical images.

*Medical Image Analysis*, 2017.



Lequan Yu, Jie-Zhi Cheng, Qi Dou, Xin Yang, Hao Chen, Jing Qin, and Pheng-Ann Heng.

*Automatic 3D Cardiovascular MR Segmentation with Densely-Connected Volumetric ConvNets*, pages 287–295.

Springer International Publishing, Cham, 2017.