

## Bridging the Gap Between 2D and 3D

### Organ Segmentation with Volumetric Fusion Net

#### 一、摘要

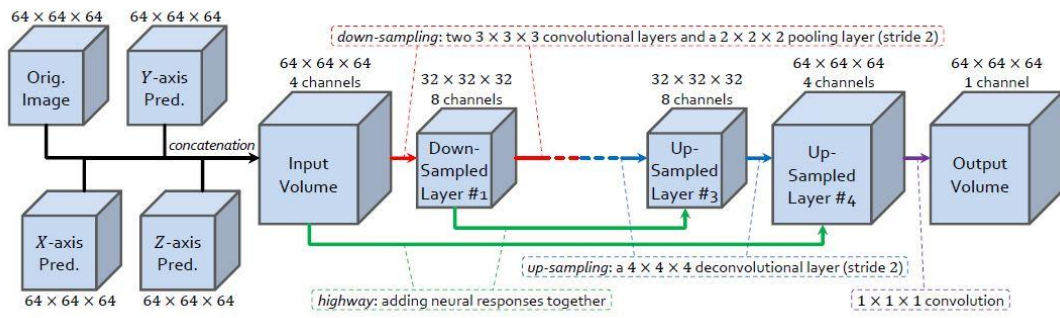
There has been a debate on whether to use 2D or 3D deep neural networks for volumetric organ segmentation. Both 2D and 3D models have their advantages and disadvantages. In this paper, we present an alternative framework, which trains 2D networks on different view-points for segmentation, and builds a 3D Volumetric Fusion Net (VFN) to fuse the 2D segmentation results. VFN is relatively shallow and contains much fewer parameters than most 3D networks, making our framework more efficient at integrating 3D information for segmentation. We train and test the segmentation and fusion modules individually, and propose a novel strategy, named cross-cross-augmentation, to make full use of the limited training data. We evaluate our framework on several challenging abdominal organs, and verify its superiority in segmentation accuracy and stability over existing 2D and 3D approaches.

关于是否使用 2D 或 3D 深度神经网络进行体积器官分割存在争议。2D 和 3D 模型都有其优点和缺点。在本文中，我们提出了一个替代框架，它在不同的观点上训练 2D 网络进行分割，并构建一个 3D 体积融合网络（VFN）来融合 2D 分割结果。VFN 相对较浅，包含的参数比大多数 3D 网络少得多，使我们的框架更有效地整合 3D 信息进行分割。我们单独训练和测试分割、融合模块，并提出一种称为交叉增强的新策略，以充分利用有限的训练数据。我们在几个具有挑战性的腹部器官上评估我们的框架，并验证其在分割准确性和稳定性方面优于现有的 2D 和 3D 方法。

#### 二、精读论文

1. 2D 和 3D 模型在进行体积器官分割都有各自的优缺点。因此作者提出一种新的训练模型：Volumetric Fusion Net (VFN)，它可以在不同的平面训练 2D 网络进行分割，同时它包含的参数比大多数 3D 网络少得多，从而可以更有效的整合 3D 信息进行分割。

2. VFN 结构模式图：



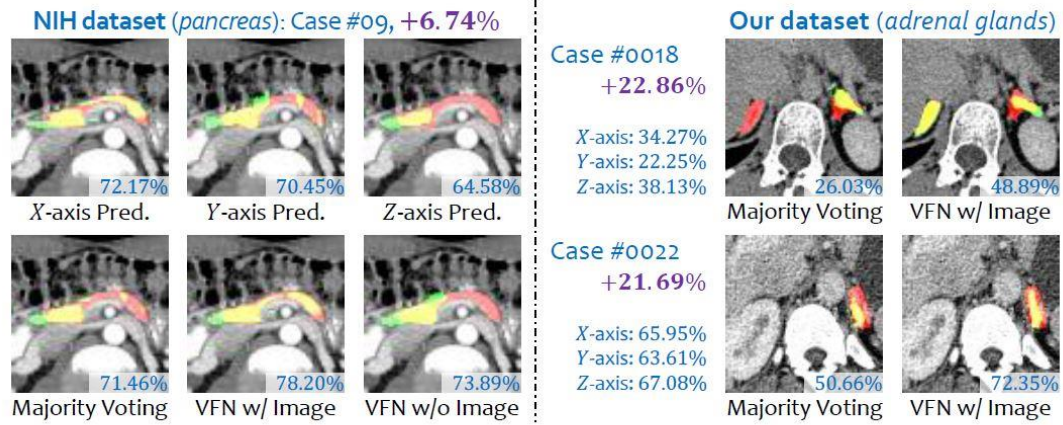
该模型只显示一个下采样和一个上采样阶段，但是每个阶段包括 3 个层面。每个下采样阶段将空间分辨率缩小 1/2，但是使通道数增加一倍并且建立了 3 个高速连接通道。之后，在每个卷积和反卷积层之后执行批量标准化和 ReLU 激活。

3. 作者为了解决训练样本不够的问题，提出了一种 cross-cross-augmentation(CCA) 的方法。

4. 在公开数据库和自己中心数据库对 VFN 模型进行测评，发现使用该模型进行腹部器官分割的准确性都优于其他的 2D 或者 3D 分割方法。

Approach	Average	Min	1/4-Q	Med	3/4-Q	Max	Time (m)
Roth <i>et al.</i> [9]	71.42 ± 10.11	23.99	—	—	—	86.29	6–8
Roth <i>et al.</i> [10]	78.01 ± 8.20	34.11	—	—	—	88.65	2–3
Roth <i>et al.</i> [11]	81.27 ± 6.27	50.69	—	—	—	88.96	2–3
Cai <i>et al.</i> [3]	82.4 ± 6.7	60.0	—	—	—	90.1	N/A
Zhu <i>et al.</i> [19]	84.59 ± 4.86	69.62	—	—	—	91.45	4.1
Zhou <i>et al.</i> [18]	82.50 ± 6.14	56.33	81.63	84.11	86.28	89.98	0.9
[18] + NLS	82.25 ± 6.57	56.86	81.54	83.96	86.14	89.94	1.1
[18] + VFN	<b>84.06 ± 5.63</b>	62.93	81.98	85.69	87.62	91.28	1.0
Yu <i>et al.</i> [17]	84.48 ± 5.03	62.23	82.50	85.66	87.82	91.17	1.3
[17] + NLS	84.47 ± 5.03	62.22	82.42	85.59	87.78	91.17	1.5
[17] + VFN	<b>84.63 ± 5.07</b>	61.58	82.42	85.84	88.37	91.57	1.4

**Table 1.** Comparison of segmentation accuracy (DSC, %) and testing time (in minutes) between our approach and the state-of-the-arts on the NIH dataset [9]. Both [18] and [17] are reimplemented by ourselves, and the default fusion is majority voting.



**Fig. 2.** Two typical examples, each with the original image, segmentation results from three viewpoints, and different fusion results. In each label map, red, green and yellow indicate ground-truth, prediction and overlap, respectively (best viewed in color).

Approach	<i>adrenal g.</i>	<i>duodenum</i>	<i>gallbladder</i>	<i>pancreas</i>
Zhu <i>et al.</i> [19]	36.74 ± 25.14	68.80 ± 14.38	42.01 ± 29.47	85.25 ± 6.04
Zhou <i>et al.</i> [18]	66.09 ± 18.19	71.65 ± 13.15	90.39 ± 5.30	84.52 ± 6.23
[18] + VFN	<b>69.24 ± 17.42</b>	<b>72.77 ± 12.80</b>	<b>91.40 ± 5.19</b>	<b>86.39 ± 6.20</b>
Yu <i>et al.</i> [17]	71.40 ± 12.87	77.48 ± 8.70	91.81 ± 4.90	87.22 ± 5.90
[17] + VFN	<b>72.09 ± 13.61</b>	<b>77.77 ± 8.46</b>	<b>92.15 ± 5.05</b>	<b>88.06 ± 5.33</b>

**Table 2.** Comparison of segmentation accuracy (DSC, %) on our multi-organ dataset. The baseline for [18] and [17] is majority voting. The numbers of [17] are different from those in their original paper, because we are using a different dataset.

### 三、论文复现

本文的代码复现详见[/code/](#)文件夹，复现的过程及运行我复现代码所需要的操作详见[/code/README.md](#)。

### 四、实验及结果

1. 在腹部 CT 不同的 slice 分割腹部器官：

图 1

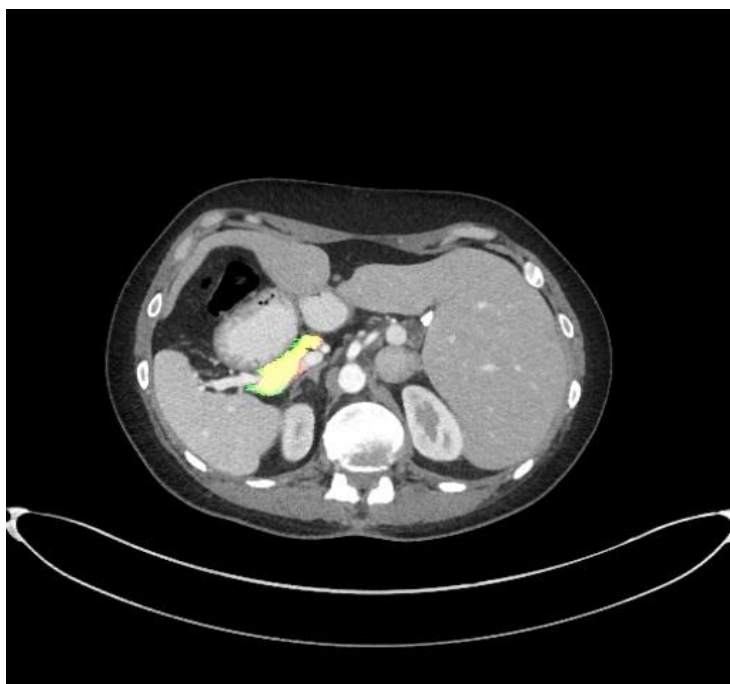


图 2

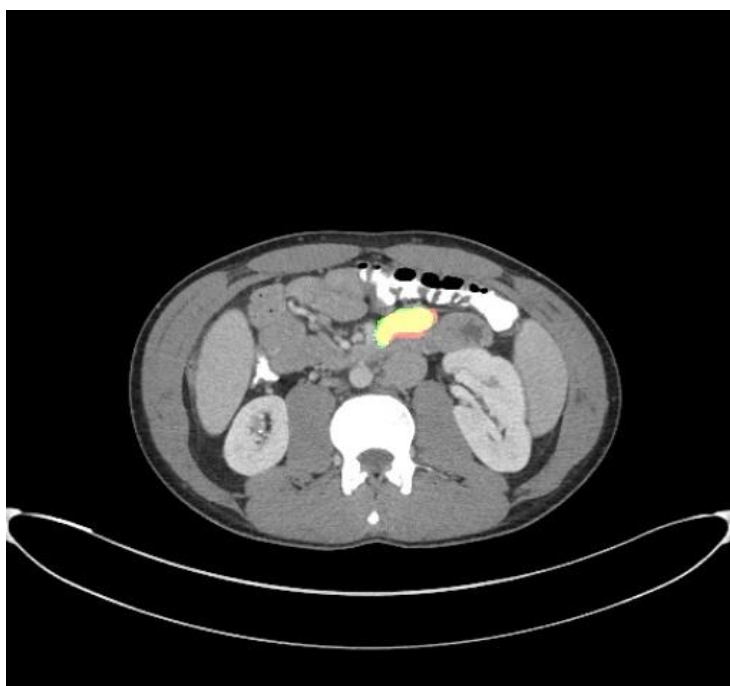
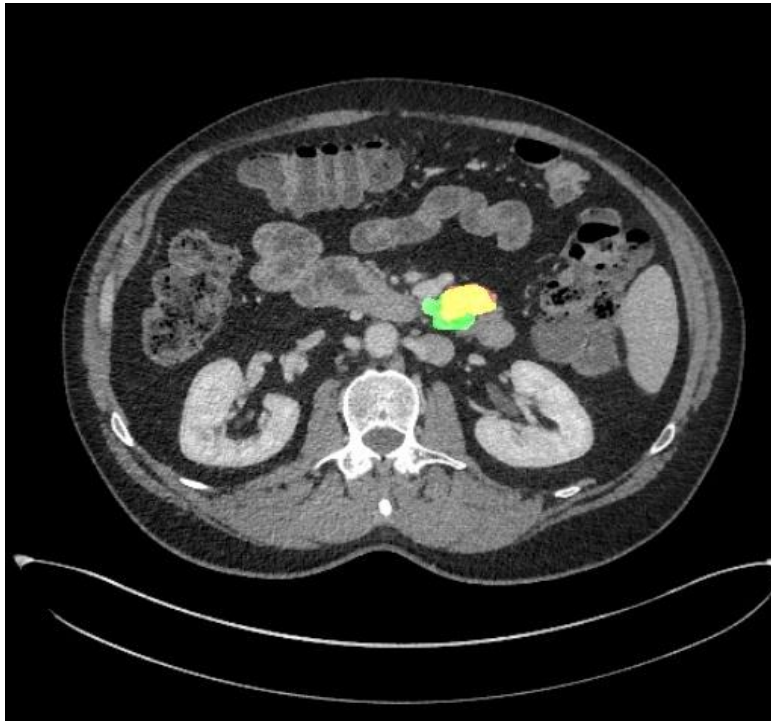
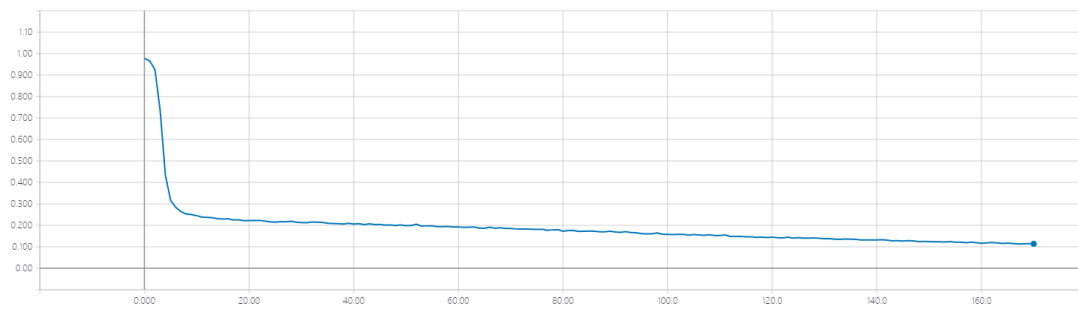


图 3



2. 模型损失曲线图:



## 五、实验结论

在不同的视点训练 2D 网络进行分割，并构建一个 3D 体积融合网络（VFN）来融合 2D 分割结果，可以提高分割的准确性和稳定性。如何更加有效地训练 VFN 模型，例如：使用硬件示例挖掘，这需要我们进一步的实验探究。