



西安交通大学

XI'AN JIAOTONG UNIVERSITY

基于3D Unet头颈部CT图像器官分割的研究与应用

姓 名：刘婷

学 号：3117393099

指导老师：田丽华

组织结构

一、研究背景与意义

二、研究现状

三、基于3D Unet头颈部CT图像器官分割算法的研究

四、基于模型压缩的3D Unet头颈部CT图像器官分割算法的研究

五、结论与展望





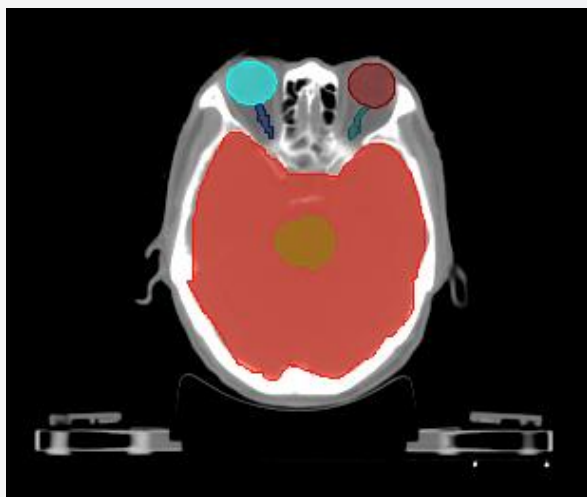
西安交通大学
XI'AN JIAOTONG UNIVERSITY

一：研究背景与意义

研究背景与意义

放疗是目前头颈部肿瘤治疗的主流方法。在放射治疗中，第一步、也是最关键的一步就是勾画出肿瘤和危及器官的轮廓，以便放疗师制定相应的放射剂量方案。

由于CT图像是由几十甚至上百张二维图像堆叠而成，手工勾画是非常耗时且繁琐的。



头颈部CT截面图



头颈部CT三维图



西安交通大学
XI'AN JIAOTONG UNIVERSITY

二：研究现状

国内外研究现状

- 2017年之前，主要是传统分割方法，**基于Atlas的分割方法、基于统计形状/外观模型的方法。**

处理流程复杂、分割精度低且不稳定、耗时

- 2017年至今，深度学习开始应用到CT图像器官分割，**CNN、Vnet, 3D Unet等。**

处理流程相对简单、分割精度相对较高



西安交通大学
XI'AN JIAOTONG UNIVERSITY

三：基于3D Unet头颈部CT图像器官 分割算法的研究

研究内容

分割头颈部CT图像的**八个**危及器官：脑干、左右眼睛、左右视神经、左右腮腺、下颌骨，分割最终结果应如下图：（三维）

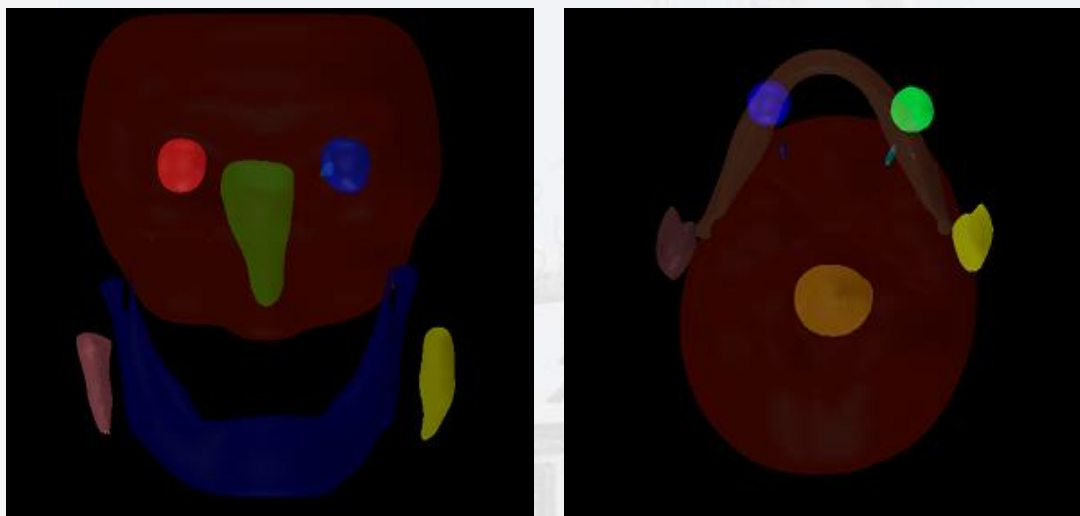


图3-1 分割最终效果图

研究目标

- 提高头颈部8个器官分割的整体精确度
- 提高小器官分割的精确度
- 减少分割花费的时间和占用的内存

研究挑战

- 器官相对于整张CT图像 占比非常小
- 不同器官的大小、体积、形状 差异性大
- CT图像中 软组织对比度差 部分器官与周围组织分界不明显

实验数据与环境

实验环境：python3.5, Tensorflow1.6框架,
2块Nvidia GeForce 1080Ti GPU,
Ubuntu16的操作系统。

实验数据：研二实习公司合作医院。
其中训练集30个，验证集10个。

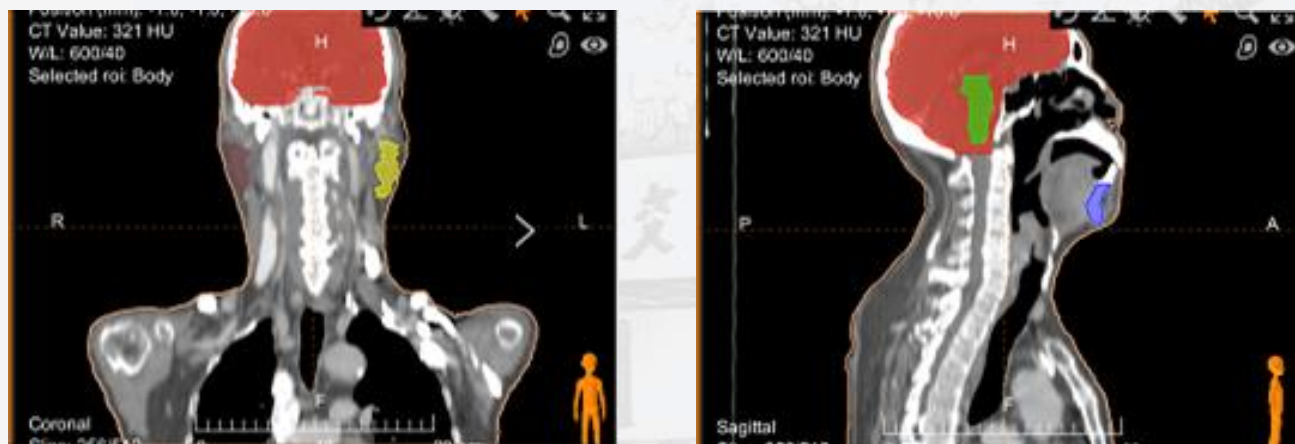


图3-2 训练数据 冠状面 矢状面

数据扩增

采用平移、旋转和缩放操作，对头颈部CT图像进行数据扩增。

平移: $[-8, +8]$

旋转角度: $[-5, +5]$

缩放比例: $[0.9, 1.1]$

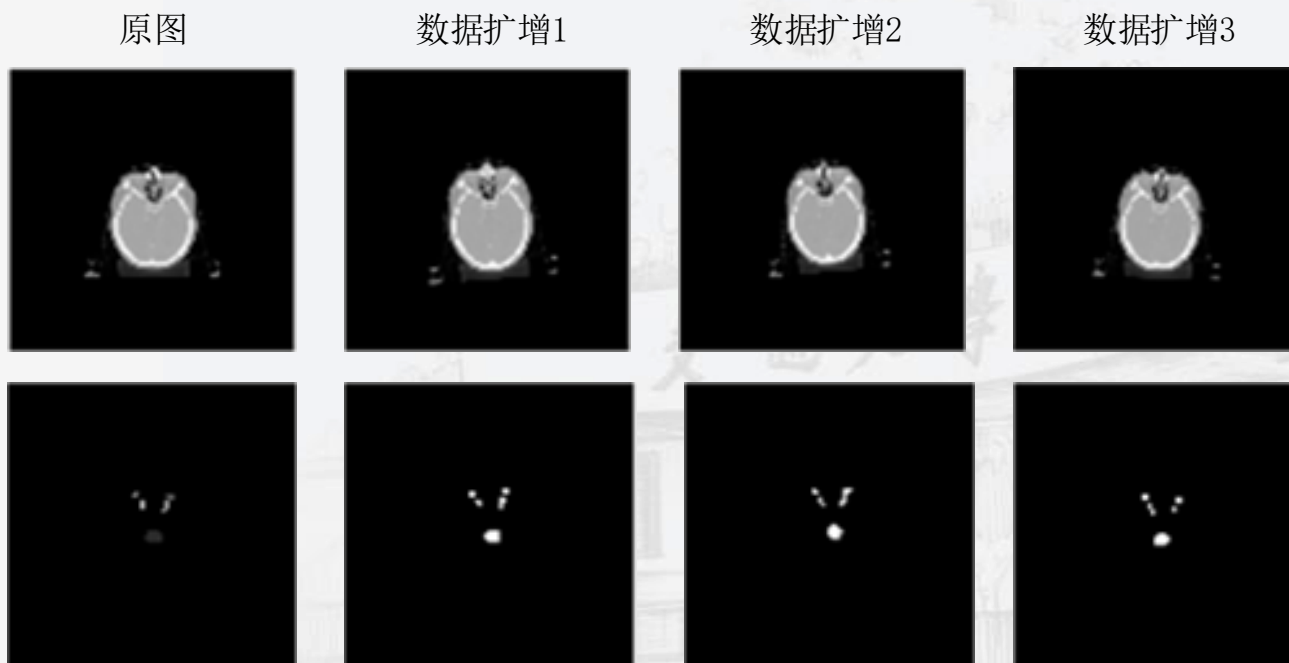


图3-3 数据扩增示意图

神经网络整体结构

本文从池化次数、卷积深度、损失函数三个方面对原始3D U-net做了改进，最终网络结构如下：

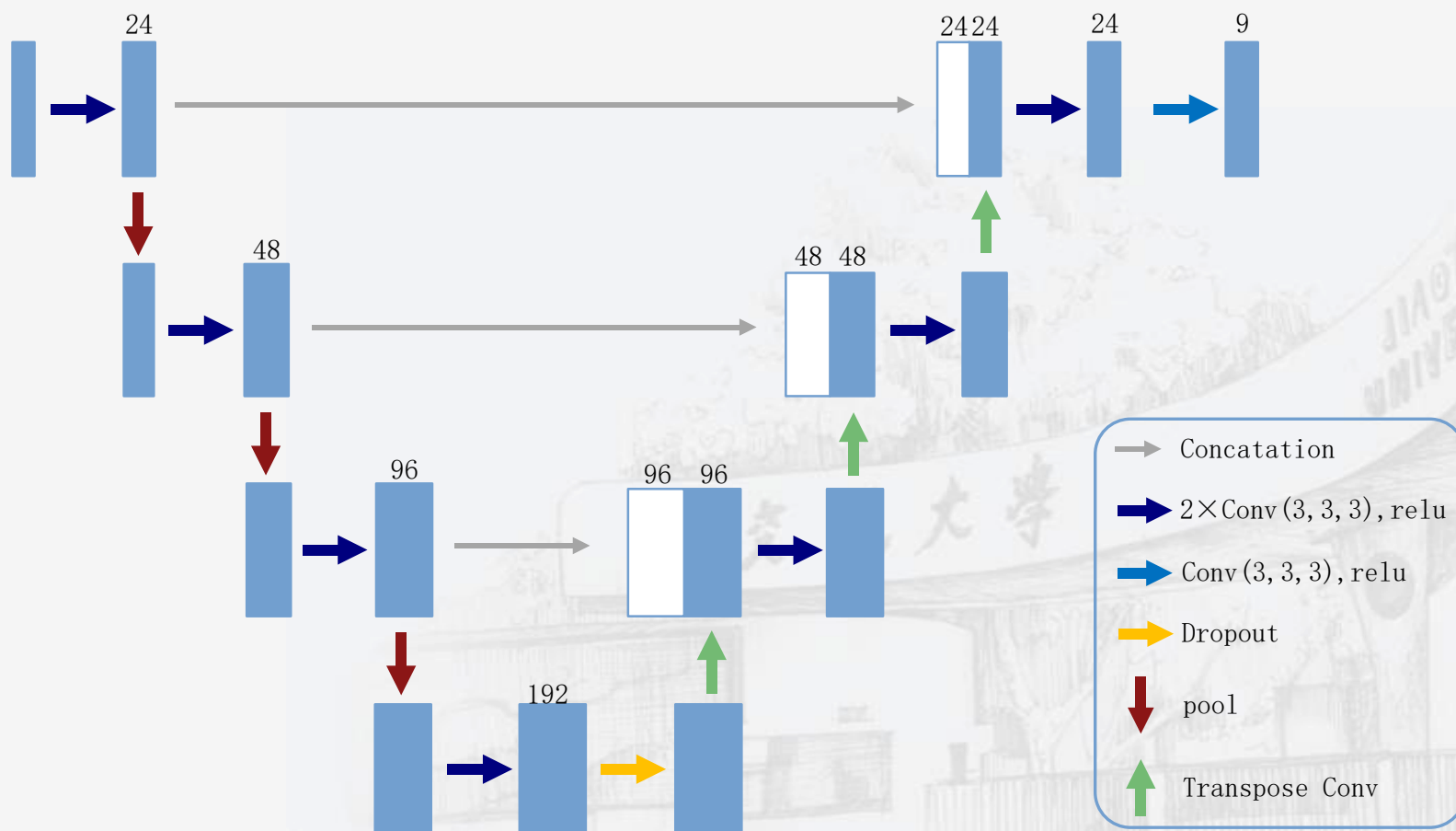


图3-4 神经网络结构图

池化次数

考虑到池化次数过多，会丢失信息；而池化次数过少又会影响到整个网络的“视野”，为了得到最佳的池化次数，本文将2次、3次与4次池化的模型进行实验对比：

	Pool2	Pool3	Pool4
脑干	0.524	0.645	0.647
左眼睛	0.780	0.843	0.849
右眼睛	0.768	0.832	0.836
下颌骨	0.862	0.878	0.882
左视神经	0	0.325	0.331
右视神经	0	0.317	0.320
左腮腺	0.633	0.731	0.736
右腮腺	0.608	0.706	0.712
Average	0.522	0.660	0.664

卷积深度

卷积深度过小，可能漏掉某些重要特征，导致模型欠拟合；而卷积深度过大，可能把噪音当做图像本身的特征，导致模型过拟合。为了得到最佳的卷积深度，本文对初始卷积核深度为12，24，36的网络模型进行实验对比：

	12	24
脑干	0.541	0.645
左眼睛	0	0.843
右眼睛	0	0.832
下颌骨	0.835	0.878
左视神经	0	0.325
右视神经	0	0.317
左腮腺	0	0.731
右腮腺	0	0.706

损失函数

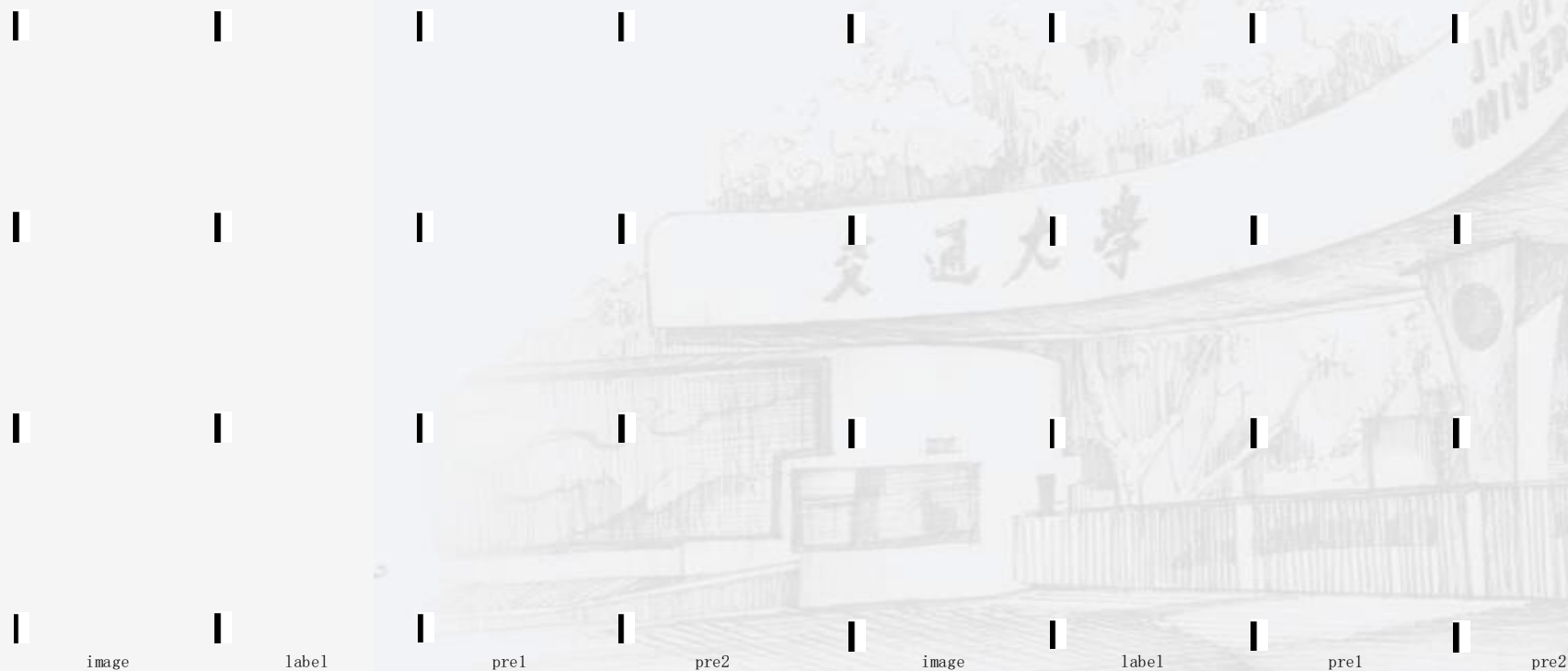
Dice系数和交叉熵都是分别对每个类别进行计算，取平均值作为结果，无法解决类别不平衡的问题；Focal loss在交叉熵的基础上，添加了平衡因子和调节因子，理论上可以解决类别不平衡的问题。

为了验证Focal loss可以解决类别不平衡问题、提高小器官（如：视神经）的分割精确度，本文设计了使用不同损失函数的对比实验：

	dice	cross entropy	focal loss
脑干	0.617	0.645	0.658
左眼睛	0.803	0.843	0.851
右眼睛	0.795	0.832	0.840
下颌骨	0.871	0.878	0.883
左视神经	0.287	0.325	0.338
右视神经	0.276	0.317	0.327
左腮腺	0.702	0.731	0.743
右腮腺	0.695	0.706	0.715
Average	0.631	0.660	0.670

Two Stage流程

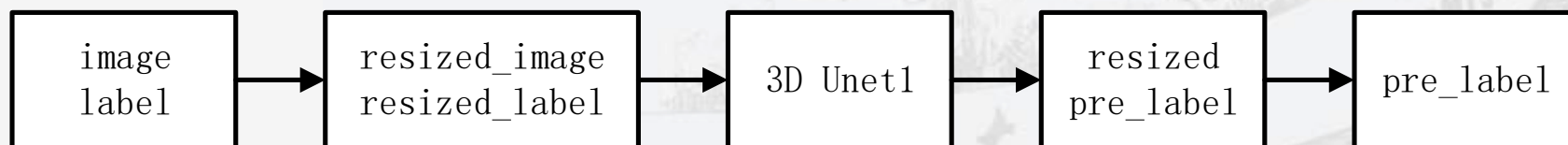
针对头颈部CT图像软组织对比度不明显，器官在整张CT图像上占比小的问题，引入Two Stage训练与分割，把第一阶段的分割结果用于定位器官，第二阶段分别对每个器官进行分割，旨在整体提高各个器官分割的精确度。



Two Stage训练流程

第一阶段：

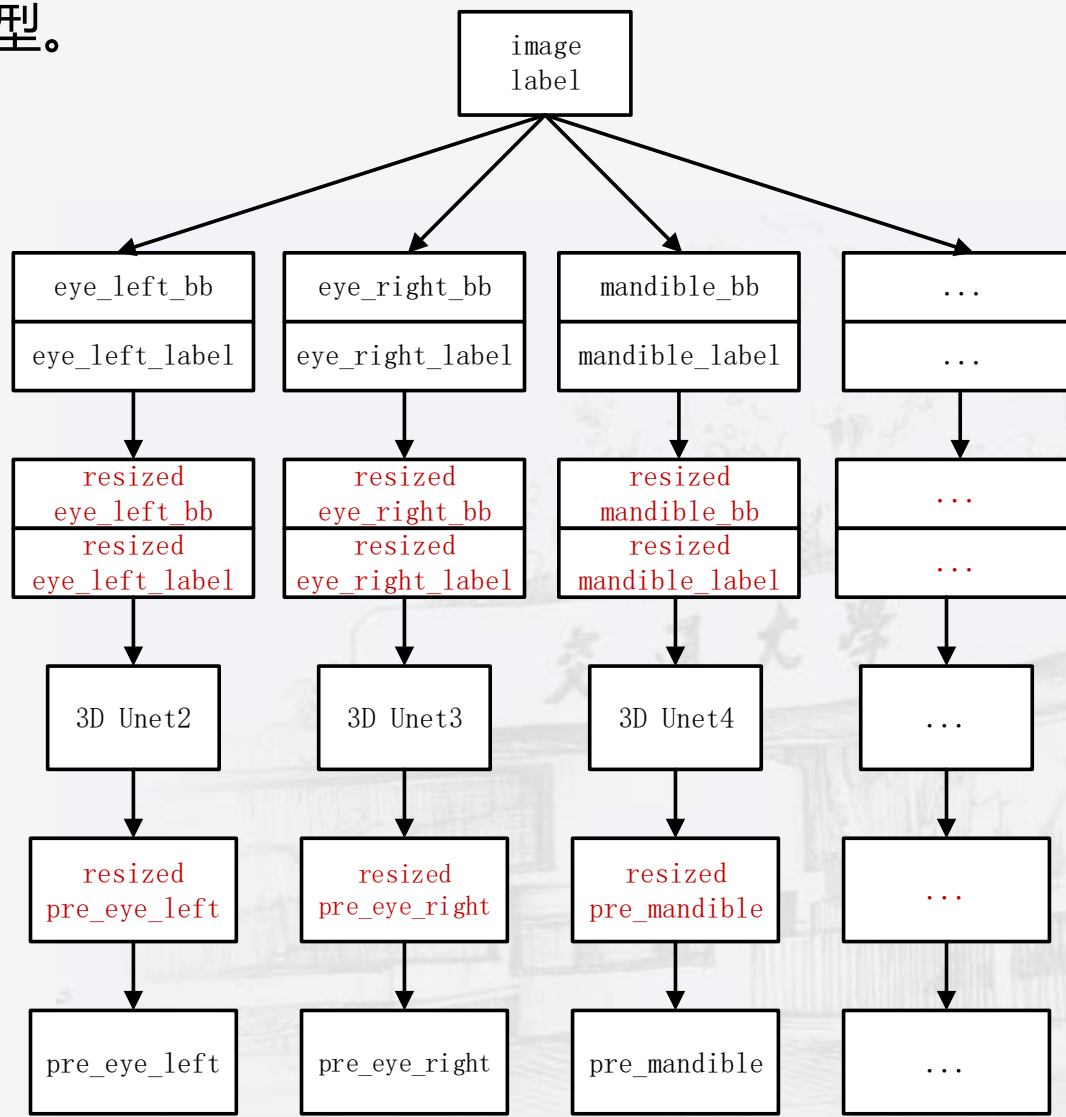
训练一个可以同时分割出八个器官的网络模型。



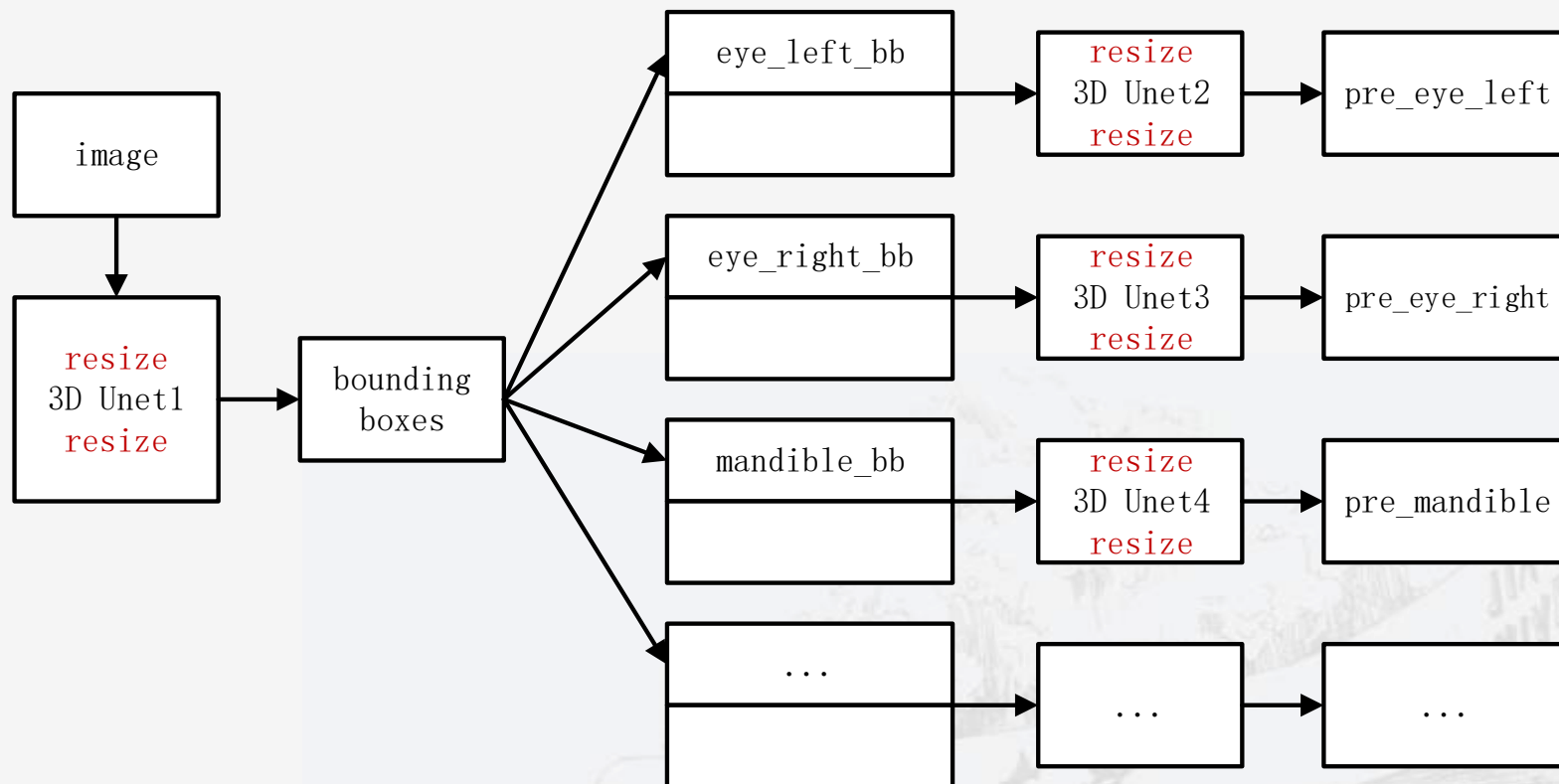
Two Stage训练流程

第二阶段:

从训练集中提取器官的bounding box, 分别为每个器官训练分割模型。



Two Stage分割流程



首先，运行第一阶段网络，一次性得到八个器官的分割结果；

然后，提取各个器官的bounding box，运行各自的网络模型进行更加精细的分割。

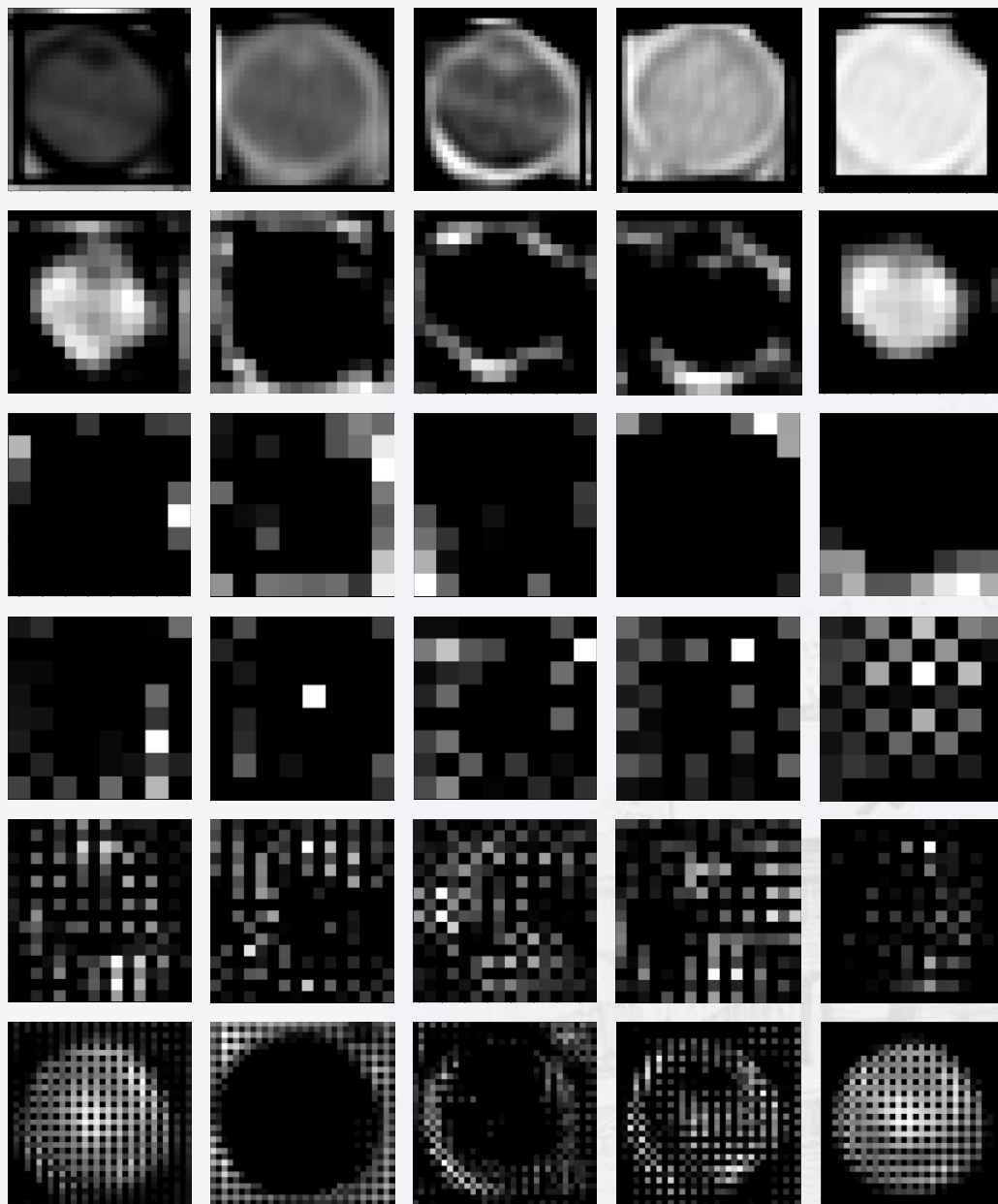
Two Stage 第一阶段特征图



器官占比非常小

噪声多

Two Stage 第二阶段特征图



器官占比大

噪声少

Two Stage流程效果

	第一阶段	第二阶段
脑干	0.658	0.747
左眼睛	0.851	0.897
右眼睛	0.840	0.881
下颌骨	0.883	0.906
左视神经	0.338	0.422
右视神经	0.327	0.414
左腮腺	0.743	0.794
右腮腺	0.715	0.776
Average	0.670	0.730

算法对比

	本文算法	AnatomyNet
脑干	0.747	0.751
左眼睛	0.897	0.884
右眼睛	0.881	0.875
下颌骨	0.906	0.910
左视神经	0.422	0.406
右视神经	0.414	0.398
左腮腺	0.794	0.786
右腮腺	0.776	0.752
平均值	0.730	0.721

算法对比

	本文算法	AnatomyNet	Fritscher et al	Ren et al
脑干	0.857	0.865	N	N
左眼睛	N	N	N	N
右眼睛	N	N	N	N
下颌骨	0.930	0.925	N	N
左视神经	0.741	0.721	N	0.720
右视神经	0.733	0.710	N	0.700
左腮腺	0.890	0.881	0.810	N
右腮腺	0.882	0.874	0.810	N

算法应用

头颈部		胸腔部		腹部		盆腔部	
器官	dice	器官	dice	器官	dice	器官	dice
脑干	0.747	脊髓	0.823	左肾	0.841	膀胱	0.705
左眼睛	0.897	左肺	0.964	右肾	0.905	左股骨头	0.745
右眼睛	0.881	右肺	0.970	肝	0.920	右股骨头	0.688
下颌骨	0.906	心脏	0.828	脊髓	0.804	小肠	0.514
左视神经	0.422	食管	0.656			直肠	0.550
右视神经	0.414					脊髓	0.671
左腮腺	0.794						
右腮腺	0.776						

成果

- Liu T, Tian L, et al. 3D U-Net Based Automatic Segmentation of Organs at Risk From CT[J]. Medical Physics, 2019, 46(6): E628-E628.
- 应用在70家医院，用于头颈部、胸腔部、腹部、盆腔部四个部位 危及器官分割



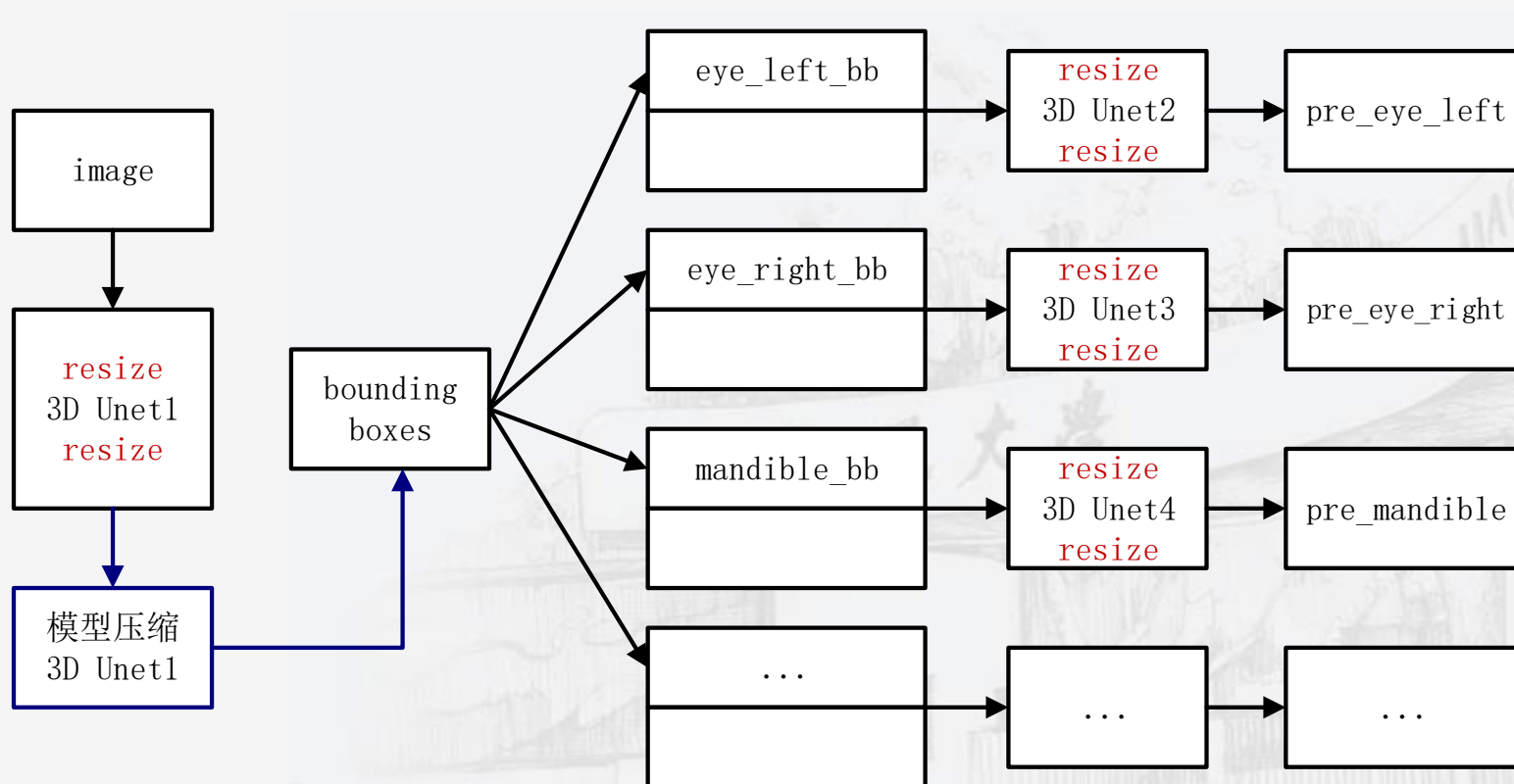
西安交通大学
XI'AN JIAOTONG UNIVERSITY

四：基于模型压缩的3D Unet头颈部 CT图像器官分割算法的研究

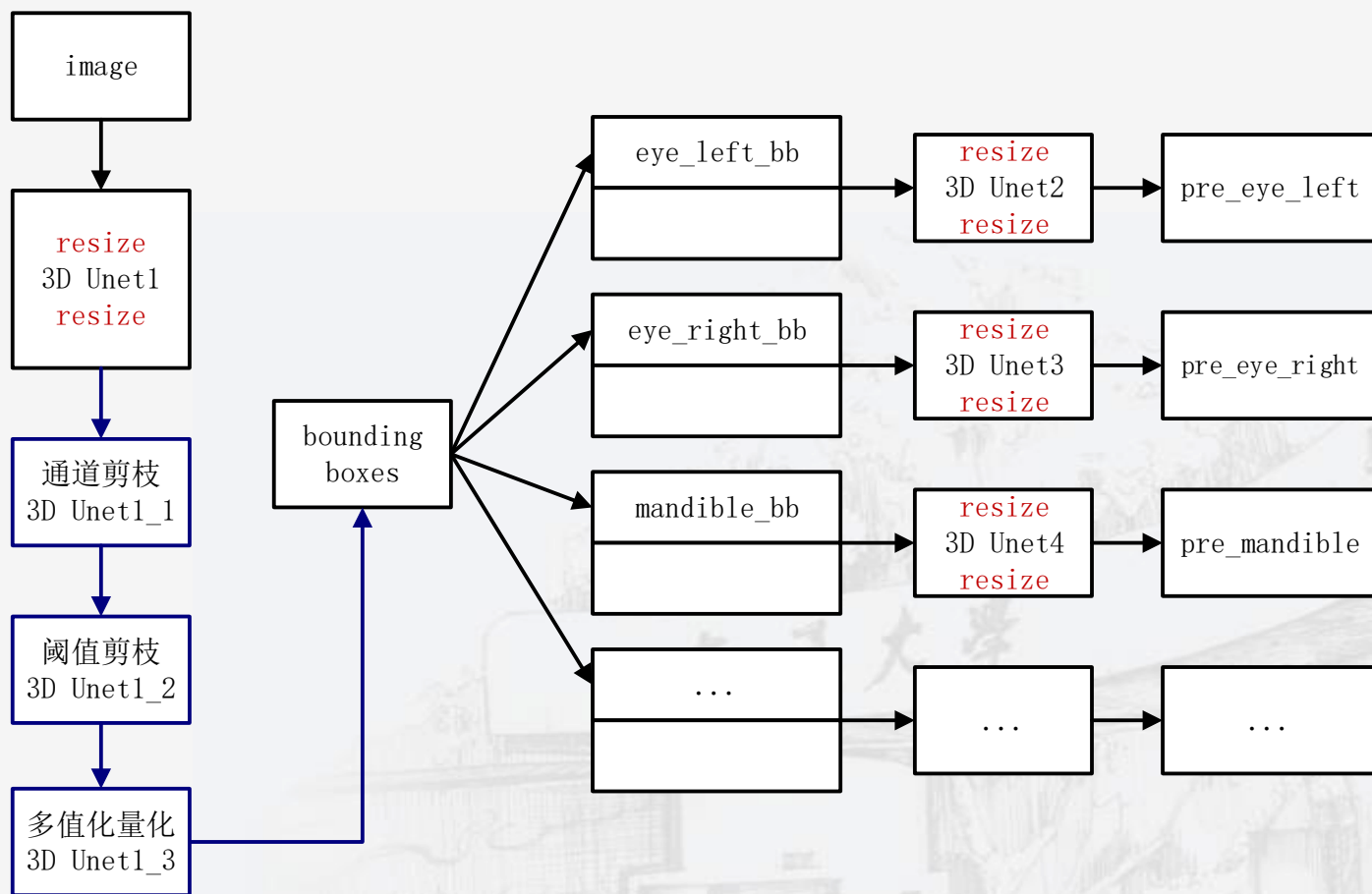
算法分析

第一阶段网络模型：3G, 30s

第二阶段网络模型：1G、200M、30M, 2s



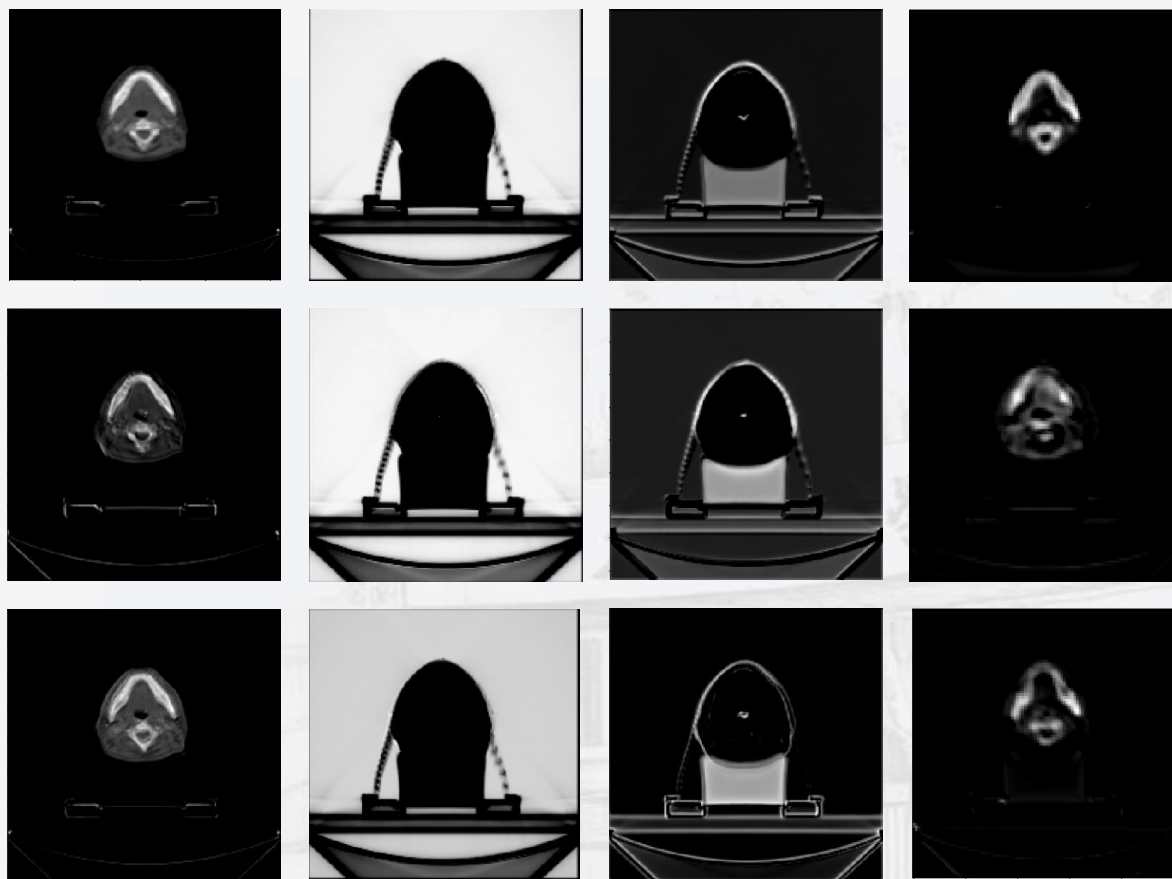
模型压缩



通道剪枝

每一列特征图都相似→ 特征图冗余

第2列 第3列提取的特征 与器官分割无关 → 特征图含有噪声



通道剪枝

输入： $\{w_k, 0 \leq k \leq C\}$ ：待剪枝的模型，共有C层； **start**：初始剪枝率，取值范围是 $[0,1]$ ，文中start=0.05； **end**：终止剪枝率，取值范围是 $[start, 1]$ ，文中end=0.25。

输出： $\{w_k, 0 \leq k \leq C\}$ ：剪枝后的模型

Repeat

for $k = 0, 1, \dots, C$ do

 计算 w_k 中卷积核的绝对值之和，并排序

 将 w_k 中卷积核绝对值最小的start部分置为零

end for

Repeat

 前向传播

 反向传播（为置为0的卷积核设置掩膜，这些卷积核不参与反向传播）

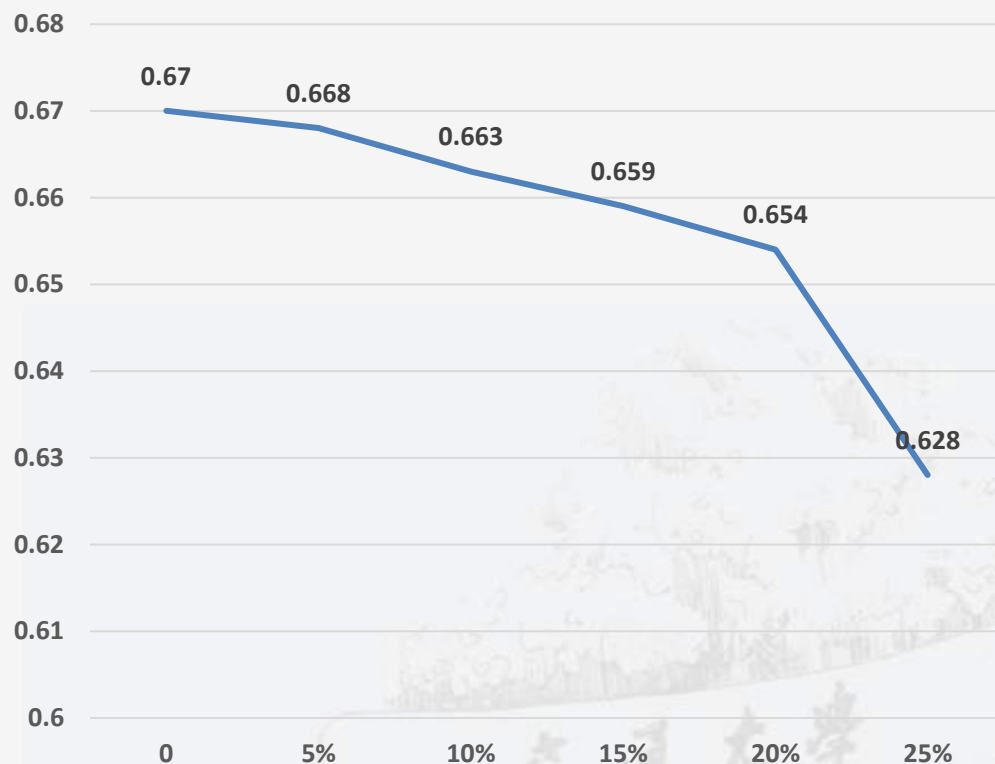
Until 达到最大迭代次数

$start \leftarrow start + 0.05$ 增大剪枝率，在上一次剪枝的基础上继续剪枝

Until $start \geq end$ 持续剪枝与再训练，直到剪枝率达到end

将权值均为零的卷积核从结构上剪掉

通道剪枝



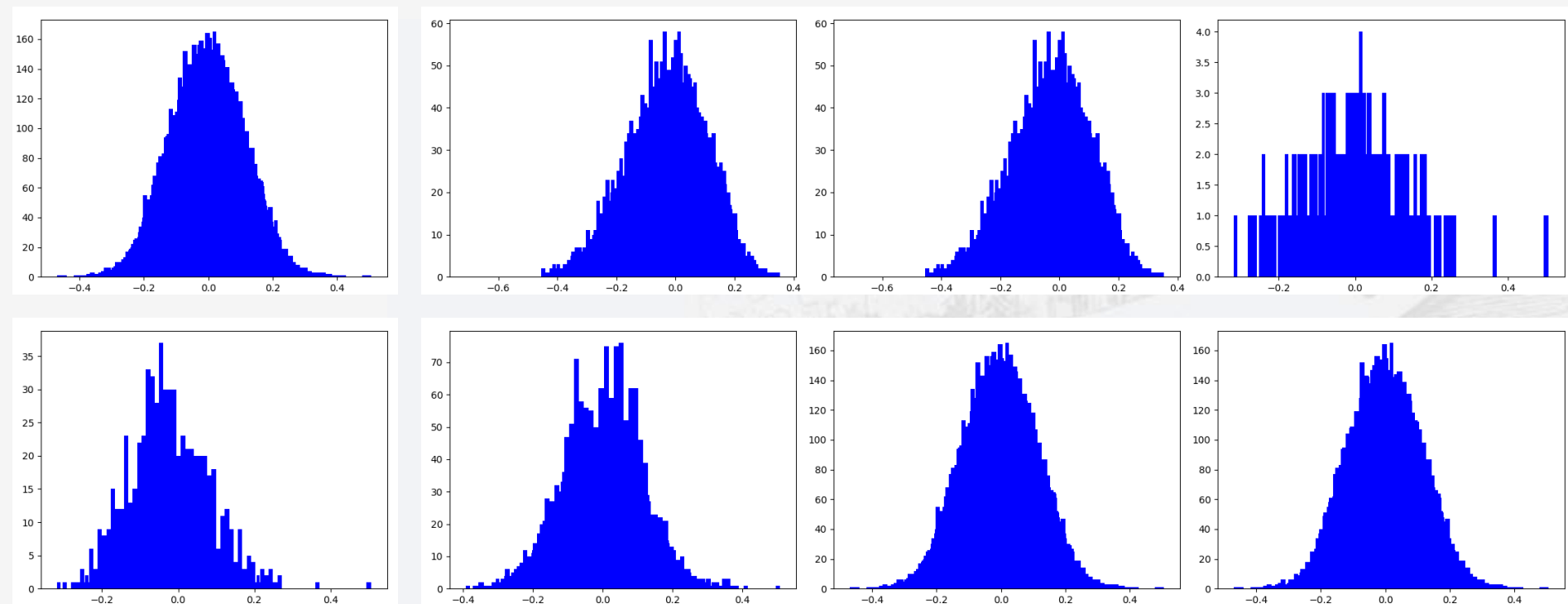
本文选择剪枝率为20%的模型作为最终通道剪枝模型。
分割精度降低约0.016;
模型压缩率为35.3%;
运行时占用内存由原来的3G降到了2G;
运行时间减少了约10s。

阈值剪枝

增加模型稀疏度、配合硬件加速

卷积核权值呈关于0对称的趋势

卷积核权值范围: $[-0.4, +0.4]$, 权值较小



阈值剪枝

输入： $\{w_k, 0 \leq k \leq C\}$ ：待剪枝的模型，共有C层； start：初始阈值，文中start=0.1； end：终止阈值，文中end=0.4；

输出： $\{w_k, 0 \leq k \leq C\}$ ：剪枝后的权重及其掩膜

Repeat

for $k = 0, 1, \dots, C$ do

 将 w_k 中绝对值小于start的值置为零

end for

Repeat

 前向传播

 反向传播（计算掩膜，置为0的元素不参与反向传播）

Until 达到最大迭代次数

 start \leftarrow start + 0.05 增大剪枝率，在上一次剪枝的基础上继续剪枝

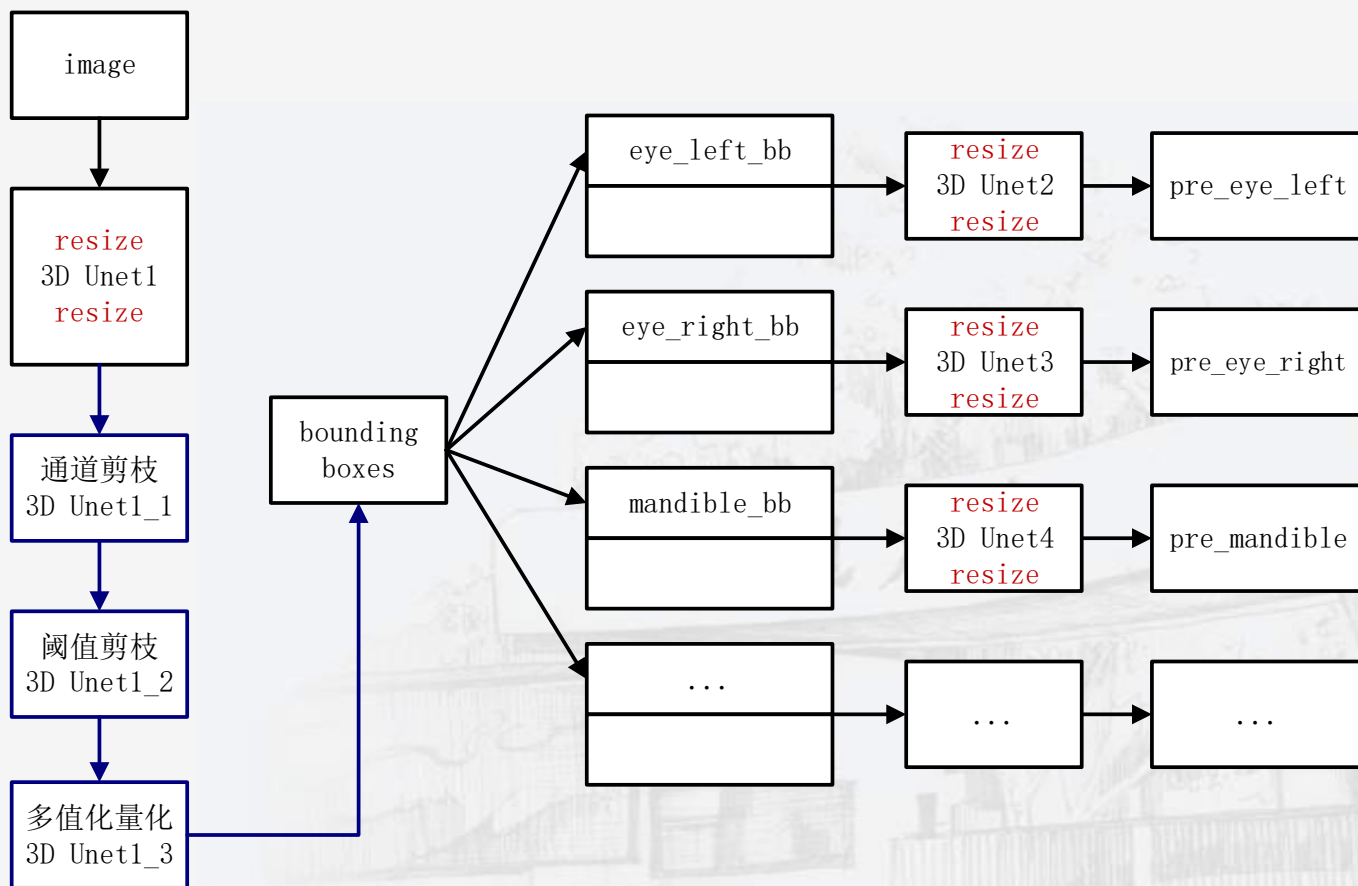
Until start \geq end 持续剪枝与再训练，直到剪枝率达到end

通道、阈值剪枝

	原模型	通道剪枝	阈值减枝
精确度	0.670	0.654	0.650
内存占用	12.03M	7.78M	
运行时内存占用	3G	2G	
运行时间	30s	20s	
稀疏度	0		0.150

量化

权值量化主要通过用有限的几个值表示神经网络中的权值来减小模型的存储空间，配合硬件加速。



量化

输入: $\{w_k, b_k, 0 \leq k \leq L\}$: 待量化的模型, 共有L层; C: 损失函数; μ : 学习率;

输出: $\{w_k, b_k, 0 \leq k \leq L\}$: 量化后的模型

1) 前向传播

for $k = 1$ to L do

$w_b \leftarrow \text{binarize}(w_k)$ 对卷积核权值进行二值化, 不对bias进行二值化

$z_k \leftarrow w_b \times a_{k-1} + b_k$ 中间值

$a_k \leftarrow \sigma(z_k)$ 激活函数

end for

2) 反向传播

$\frac{\partial C}{\partial a_L}$ 计算输出层激活值的偏导数

for $k = L$ to 2 do

$$\frac{\partial C}{\partial z_k} = \frac{\partial C}{\partial a_k} \cdot \sigma'(z_k)$$

$$\frac{\partial C}{\partial a_{k-1}} = (w_k)^T \cdot \frac{\partial C}{\partial z_k}$$

$$\frac{\partial C}{\partial w_b} = \frac{\partial C}{\partial z_k} \cdot (a_{k-1})^T \text{ 计算 } w_b \text{ 的梯度, 即: 对 } w_k \text{ 量化之后的梯度}$$

$$\frac{\partial C}{\partial b_k} = \frac{\partial C}{\partial a_k} \text{ 计算 } b_k \text{ 的梯度}$$

$$w_k \leftarrow w_k - \mu \frac{\partial C}{\partial w_b} \text{ 更新 } w_k, \text{ 相当于把量化之后的梯度赋值给了量化之前的梯度}$$

$$b_k \leftarrow b_k - \mu \frac{\partial C}{\partial b_k} \text{ 更新 } b_k$$

end for

量化

量化区间: $[-0.45, +0.45]$

量化策略: 均匀取点

$$w_b = \begin{cases} +1, & \text{if } w \geq 0 \\ -1, & \text{otherwise} \end{cases}$$

$$w_b = \begin{cases} +1, & \text{if } w > 0.15 \\ 0, & \text{if } -0.15 \leq w \leq 0.15 \\ -1, & \text{if } w < -0.15 \end{cases}$$

$$w_b = \begin{cases} 0.27, & \text{if } w > 0.27 \\ 0.09, & \text{if } 0.09 \leq w \leq 0.27 \\ 0, & \text{if } -0.09 \leq w \leq 0.09 \\ -0.09, & \text{if } -0.09 \leq w \leq -0.27 \\ -0.27, & \text{if } w < -0.27 \end{cases}$$

量化

	卷积核内存占用	精确度
原模型	7.78M	0.650
二值化	0.24M	0.231
三值化	0.49M	0.272
七值化	0.78M	0.382
十五值化	0.97M	0.539
三十一值化	1.9M	0.646

小结

	原模型	剪枝量化
精确度	0.670	0.646
内存占用	12.03M	1.9M
运行时间	30s	20s
稀疏度	0	0.150

由于第一阶段的网络仅用于器官的定位，第二阶段的网络会对器官进行更加精细的分割，本章的模型压缩仅将第一阶段网络模型的精确度降低了0.024，实验表明，这不会导致各个器官最终分割精度的下降。



西安交通大学
XI'AN JIAOTONG UNIVERSITY

五：结论与展望

结论

1) 提出一种基于3D Unet头颈部CT图像器官分割算法。首先，从池化层数、卷积深度和损失函数三个方面对原始3D Unet进行了改进，得到了一个适合于头颈部CT图像器官分割的网络模型；其次，引入Two Stage训练与分割，整体提高各个器官分割的精确度。最后，在本文的数据集上将本文使用的算法与其他的头颈部CT图像分割算法进行了实验对比，结果表明：本文算法相比于当下流行的两种算法有更好的分割效果。

2) 对基于3D Unet头颈部CT图像器官分割算法模型压缩的研究。本文仅对第一阶段用于定位的网络模型进行了压缩，压缩后的模型体积大大减小，只有原本的1/6；模型运行时占用的内存由3G降到了不足2G；模型的运行时间减少了10s；模型的稀疏度为0.15。由于第一阶段的网络仅用于器官的定位，第二阶段的网络会对器官进行更加精细的分割，实验结果表明：第一阶段（器官定位）网络模型由于模型压缩降低的0.024的精确度，不会影响第二阶段各个器官更加精细的分割，器官最终分割精确度不变。

展望

- 1) 本文采用Two Stage训练与分割流程，虽然极大的提升了分割的精确度，但由于第二阶段需要训练八个小网络模型，导致训练过程非常耗时，未来将对八个小网络模型特征复用进行研究。
- 2) 为了不损失模型的分割精度，本文仅对第一阶段的网络模型进行了压缩，未来将对第二阶段八个网络模型进行压缩。



西安交通大学
XI'AN JIAOTONG UNIVERSITY

谢谢大家

