 ****

**实验报告**

**实验课程： 自然语言处理作业**

**实验题目：**

**姓名： 金渤蛟**

**学号：2271326 同组人： 无**

**实验日期： 2022-12-5 指导教师： 肖桐**

**实验成绩（教师签字）：**

# **实验一：bert与tiny\_bert的对比实验**

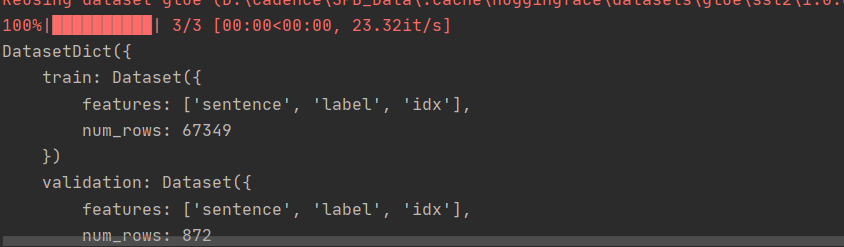
实验一：

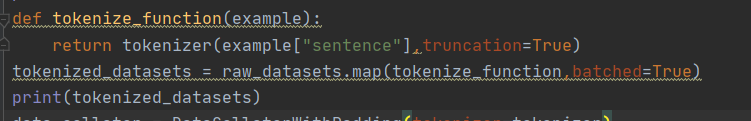
实验内容：使用finetune方法搭建/训练一个情感分类模型

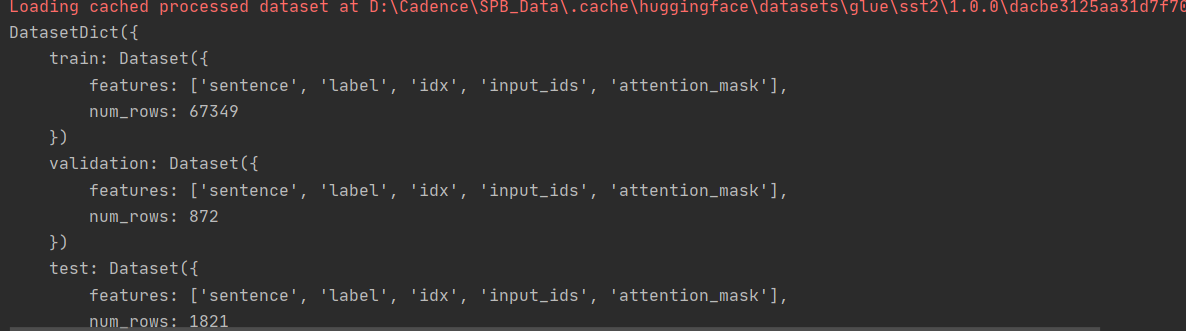
数据集：SST-2

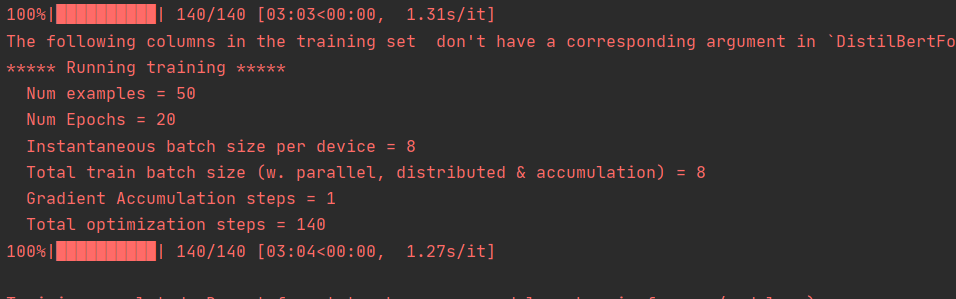
预训练模型：基础需要为BERT-base

基础实验：根据课上提供的方法，我首先使用了bert-base-uncased作为基础实验，SST-2是huggingface中的原有数据集，我通过datasets中的data\_loader从网站下载，接着我对数据进行了观察数据是由句子标签和编号构成并且分为train，validation，test，使用tokenize进行文字编码将文字转换为向量，使用hugging-face使用TrainingArguments更改训练参数（输出模型文件路径，训练次数，学习率），接着使用AutoModelForSequenceClassification这个数据时进行二分类问题所以要分成两种标签，因为我没有gpu所以我从数据集中进行随机切片，取得小批量数据进行训练。









对比实验：因为设备原因，我使用了更小的distilbert模型进行实验比较,通过模型对比，更小的distilbert提升了很高的速度，但是相比于Bert的准确性要减少了一些，对于学习率的影响我使用了速度更快的Distilbert进行了对比,可以很明显的看出由于学习率的下降导致损失率下降变慢，使得模型训练效率降低，而学习率过大会导致损失率下降过快易导致模型不收敛。

Epoch = 20 lr = 5e-5

|  |  |  |
| --- | --- | --- |
| Model | Bert-base | Distilbert |
| Time | 342.4 | 188.2 |
| Loss | 0.255 | 0.334 |

Epoch = 30 lr = 5e-5

|  |  |  |
| --- | --- | --- |
| Model | Bert-base | Distilbert |
| Time | 464.63 | 260.96 |
| Loss | 0.184 | 0.237 |

Epoch = 40 lr = 5e-5

|  |  |  |
| --- | --- | --- |
| Model | Bert-base | Distilbert |
| Time | 630.01 | 358.45 |
| Loss | 0.103 | 0.142 |

Epoch = 100 lr = 5e-5

|  |  |  |
| --- | --- | --- |
| Model | Bert-base | Distilbert |
| Time | 1630.75 | 871.83 |
| Loss | 0.043 | 0.052 |

Epoch = 20 lr = 1e-1

|  |  |
| --- | --- |
| Model | Distilbert |
| Time | 187.92 |
| Loss | 0.694 |

Epoch = 40 lr = 1e-1

|  |  |
| --- | --- |
| Model | Distilbert |
| Time | 354.16 |
| Loss | 0.694 |

Epoch =100 lr=1e-1

|  |  |
| --- | --- |
| Model | Distilbert |
| Time | 354.16 |
| Loss | 0.691 |

Epoch =20 lr=10e-10

|  |  |
| --- | --- |
| Model | Distilbert |
| Time | 187.8 |
| Loss | 0.115 |

Epoch =40 lr=10e-10

|  |  |
| --- | --- |
| Model | Distilbert |
| Time | 187.8 |
| Loss | 0.056 |

# 实验二 肝脏及癌变部位的图像分割

### 实验思路

根据肝肿瘤分割算法的不同原理分为传统方法和深度学习两大类，传统分割算法多归类为阀值分割、区域生成、水平集等。而深度学习根据卷积维度将分割卷积模型分为2D、2.5D以及3D。传统分测算法形依赖于先验知识，不利于临床应用中的自动化扩展。近年来，深度卷积神经网络(Deep Convolutional Neural Network, DCNN)被广泛应用在肝肿瘤分割任务上，通过局部感知和参数共享的特性，让模型能够从大量样本中有效学习到目标特征。

肝脏和肿瘤区域属于三维结构，分割网络设计时常用3D卷积代替2D卷积作为特征提取模块，该设计方式将急剧地增加模型复杂度，导致参数冗余问题。针对此问题，我想到将三维数据拆解成二维切片数据，对每一张切片进行分割。Unet的优点是在很小的数据样本下也能训练出很好的模型。SegNet的优点是不用保存整个encoder部分的特征图，只需保存池化索引，节省内存空间。因此本次报告中用到的深度学习分割方法有Unet、SegNet和SwinUnet。

### 数据处理

#### 数据观察

医学领域将dicom格式文件存为nii格式，一个nii格式由三部分组成：hdr，ext，img。

Hdr（header）所包含的内容：1.维度（x,y,z）和时间t；2.voxel size（体素大小）：毫米单位的xyz大小；3.数据类型4.Form和转换矩阵

Ext（extension）：是自己可以随意定义数据的部分，可以自己用。但是通用的软件公司都无法使用这部分。

Image：储存3D或者4D的图像数据。

软件方法观测

肝脏训练数据集包含130次CT扫描，测试数据集包含70次CT扫描，对于处理CT图像的软件有很多，我选择ITK-SNAP软件进行数据观察。

ITK-Snap的主要功能是对[医学图像](https://so.csdn.net/so/search?q=%E5%8C%BB%E5%AD%A6%E5%9B%BE%E5%83%8F&spm=1001.2101.3001.7020" \t "_blank)进行分割，包括2D与3D的分割，包括人工分割与半自动分割。

ITK-Snap可以CT扫描的视角观察，主要包括三个切片窗口，分别为：轴向面，矢状面和冠状面，有利于我观察和处理图像。

我首先使用Lager Inspect查看图像的信息，我首先将窗宽窗位设置为30和150方便观察到肝脏位置，之后通过Info窗口我发现图片大小为512\*512

图形用户界面

描述已自动生成

图形用户界面

描述已自动生成

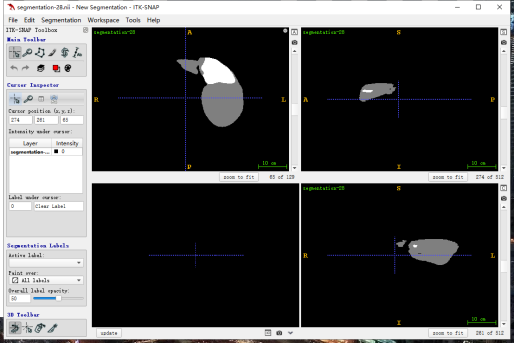


图3.1

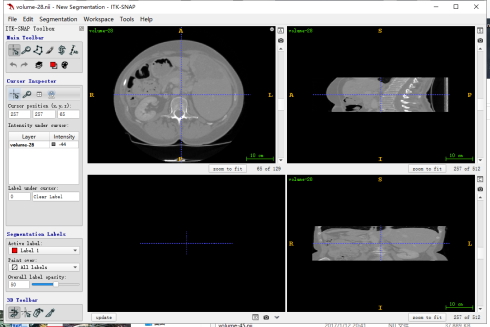


图3.2

python代码观测

通过python，我使用nibabel包进行nii文件数据的读取，并将存储在nii文件中的3D image通过matplotlib包展示出来（如图3.3和3.4），我根据数据维度将图片分段显示。

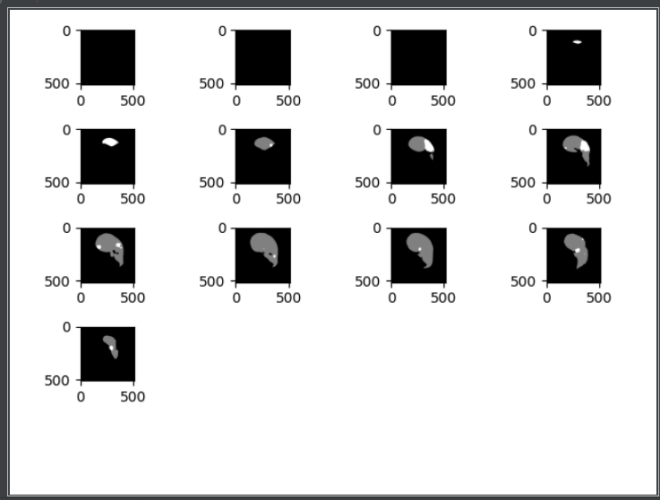


图3.3

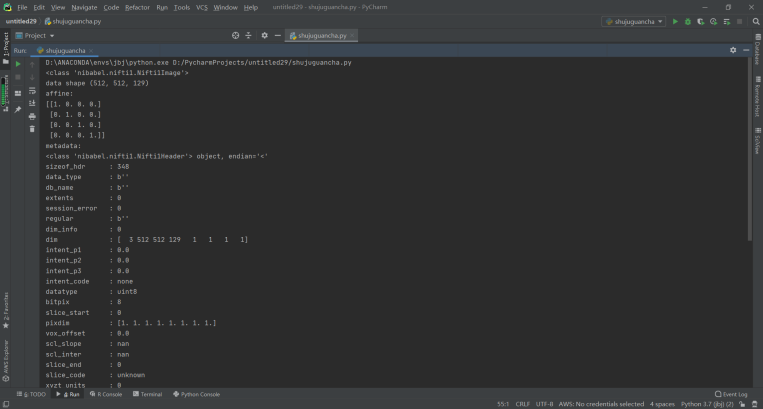
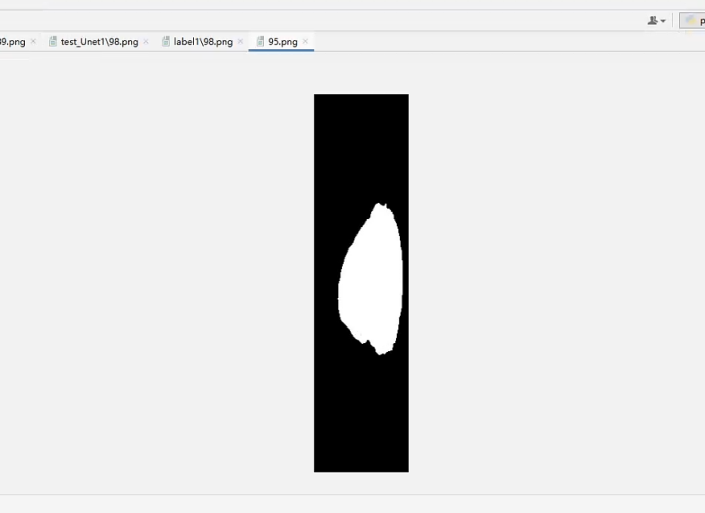


图3.4

#### 数据处理

（1）参考方法：在获取的标签里肿瘤与肝脏是一块得到的，我将肿瘤和肝脏分离开作为两种标签同时将原图片数据集分成两个标签对应的两个数据集进行对应，与原方法不同的是我么将标签色素变为255而不是122，以此来对应我的网络参数。

（2）自我构思：在对数据进行观测后我可以发现标签是由两部分组成的也就是肝和肿瘤，我的第二种方法没有按照老师提供的方法进行两个标签分割因为我发现标签中没有肿瘤的肝脏是白色的，而有肿瘤部分的肝脏是灰色的而肿瘤部位是白色的（标签如下图所示），因此我选择将nii图片转化为png形式并且在没有肿瘤的情况下分割出肝脏，在有肿瘤的情况下分割出肝脏的肿瘤部位，在将png图片放入模型之前我需要将图片和标签大小变成相同的512\*512。



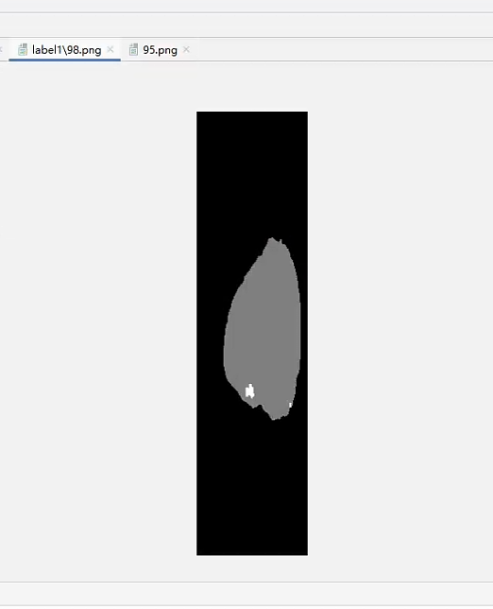


图3.5

### 深度学习模型

#### U-Net

我选择在医学图像处理中较为常见的UNet，UNet属于FCN的一种变体，它可以从小训练集中训练，非常适合LITS Challenge，U-Net网络结构如图所示，前半部分作用是特征提取，后半部分是上采样。U-Net网络结构是对称的，形似英文字母U所以被称为U-Net。

图表, 箱线图

描述已自动生成

图4.1 UNet网络结构

UNet由一个编码器（下采样器）和解码器（上采样器）组成。左半部分，由两个3x3的卷积层（RELU）再加上一个2x2的maxpooling层组成一个下采样的模块，右半部分，由一个上采样的卷积层（去卷积层）、特征拼接concat和两个3x3的卷积层（ReLU）构成。

如图所示输入572\*572的单通道图像通过第一个3\*3的卷积后，高和宽从572变为570，再通过一个3\*3卷积层又变为568。接下来，通过一个max pooling下采样，池化核大小为2x2，步距为2，由568变成了284，channel没有发生变化的还是64。再通过两个3\*3的卷积核，通过3\*3卷积后channel变为128。接下来，通过下采样max-pooling，将特征层的高和宽由280变为140，然后再通过两个3x3的卷积层，将channel由128调整为256。之后再通过下采样max-pooling，将特征层的高和宽由136变为68，然后再通过两个3\*3的卷积层，将channel由256调整为512。最后，再通过下采样max-pooling，将特征层的高和宽由64变为32，然后再通过两个3\*3的卷积层，将channel由512调整为1024。

之后进行上采样，这里上采样采用的是转置卷积，通过转置卷积后它将特征层的宽和高变为原来的两倍，由28变成56，然后channel减半，由1024变为512，对应图中的蓝色部分。灰色的箭头表示copy and crop，左侧为64\*64大小的特征层，但箭头右侧的特征图为56\*56，没法办法直接将他们进行拼接。所以对左侧64\*64大小的特征层进行一个裁剪，裁剪成56\*56大小，之后进行channel 方向的concat拼接，拼接后channel就变为了1024。然后再通过2个3\*3的卷积核将channel调整成512。重复以上步骤，以此类推，最后得到388x388，channel为64的特征层。

我所使用的UNet是将3\*3的卷积加上一个padding，就是说我每次通过3\*3的卷积层不会改变特征层的高和宽。并且我在卷积核Relu之间会加上一个Bn层。所以卷积后不改变特征层大小，两个特征图可以直接拼接，就不需要中心裁剪了，并且最终得到的卷积层高和宽与输入的卷积层高和宽是保持一致的。

#### SegNet

SegNet是通过对图像中每一个像素点进行分类，识别每一个像素点的类别来实现图像的分割。其思路与FCN的思路十分相似，只是两者的Encoder与Decoder技术不同。SegNet的创新点在于：使用记录maxpool层最大响应特征位置的tensor（pooling indices）来进行上采样，避免了FCN中学习上采样带来的消耗，而后再使用可训练的卷积层使稀疏的feature map密集，这将避面上面描述保存feature map产生的额外空间消耗。如下图是SegNet的网络结构图。

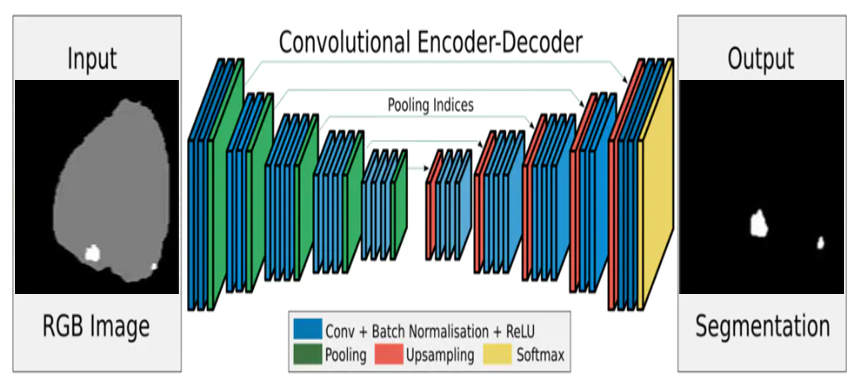


图4.2 Segnet网络结构

网络结构包括Encoder和Decoder两个部分，Encoder是由4层上采样组成，Decoder由4层下采样组成，Encoder与Decoder之间呈一种对称关系。左边使用了VGG-16的前13层卷积网络，并做了少许的改动，用于提取特征并保存池化索引；右边是一个反卷积与upsampling的过程，通过反卷积使得图像分类后特征得以重现，upsampling还原到图像原始尺寸，该过程称为Decoder；最后使用Softmax输出不同分类的最大值，得到最终分割图。

编码部分：通过卷积提取特征，SegNet使用的卷积为same卷积，即卷积后保持图像原始尺寸；在Decoder过程中，同样使用same卷积，不过卷积的作用是为upsampling变大的图像丰富信息，使得在Pooling过程丢失的信息可以通过学习在Decoder得到。其Encoder部分使用了VGG-16的前13层，与VGG-16不同的是，多了一步保存max-pooling索引的操作，目的是Decoder可以利用其来做非线性上采样。

解码部分：pooling在CNN中能够将图像缩小，包括两种方式：max和mean。文章中用到的是max-pooling的方式，在一个22的filter中，取出其中权重最大的值，同时保存最大权值在22 filter中的位置。从网络框架图可以看到绿色的pooling与红色的upsampling通过pool indices相连，实际上是pooling后的indices输出到对应的upsampling（因为网络是对称的，所以第1次的pooling对应最后1次的upsamping，如此类推）。

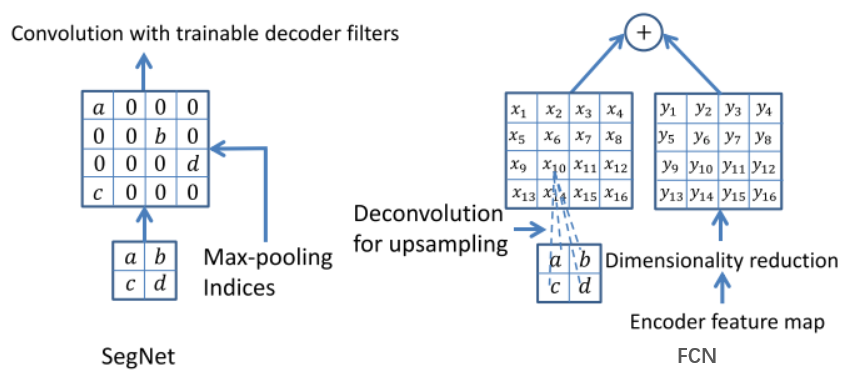


图4.3 SegNet与FCN上采样方式对比

SegNet：图像经过max-pooling后在每一个filter中会损失3个权值，但是因为在pooling过程中同时保存了最大权值的索引位置，根据pooling-indices可以在upsampling中可以pooling相对于pooling filter的相对位置，将数据放回对应位置后进做卷积训练。在与encoder中第一层对应的decoder层中（即decoder的最后一层卷积)，与原图像为RGB的3通道不同，该层产生的是一个通道为K(类别数）的multi-channel feature maps，然后将其送入softmax分类器，做逐像素的分类处理。

#### Swin-UNet

SwinUnet是基于Swin Transformer为基础，结合了UNet网络的特点组合而成的新的分割网络。SwinUnet的整体架构如图所示，SwinUnet由编码器、瓶颈、解码器和跳过连接组成。SwinUnet的基本单元是Swin transformer块。编码器、瓶颈和解码器都是基于Swin transformer块构造的。对于编码器，为了将输入转换为序列嵌入，将医学图像分割成大小为4 × 4的非重叠patch。通过这种划分方法，每个patch的特征维数变为4 × 4 × 3 = 48。此外，将投影的特征维度应用线性嵌入层，转换后的补丁标记经过几个Swin Transformer块和补丁合并层生成分层特征表示。其中，补丁合并层负责下采样和增加维度，Swin Transformer块负责特征表示学习，受UNet的启发，设计了一种基于对称transformer的解码器。该解码器由Swin transformer块和补丁扩展层组成。提取的上下文特征通过跳跃连接与编码器的多尺度特征融合，以弥补降采样造成的空间信息丢失。与补丁合并层不同，补丁扩展层被专门设计用于执行上采样。补丁扩展层通过2×上采样分辨率将相邻维度的特征地图重塑为一个大的特征地图。最后，利用最后一个补丁扩展层进行4×上采样，将特征映射的分辨率恢复到输入分辨率(W ×H)，然后在这些上采样特征上应用线性投影层输出像素级分割预测。

图示

描述已自动生成

图4.4 SwinUnet网络

Swin transformer块如图所示，每个Swin transformer块由LayerNorm (LN)层、多头自注意模块、剩余连接和2个具有GELU非线性的MLP组成。在两个连续的transformer模块中分别采用了基于窗口的多头自注意(W-MSA)模块和位移的基于窗口的多头自注意(SW-MSA)模块。

图示

描述已自动生成

图4.5 Swin transformer块

编码器：

  在编码器中，将分辨率为H/4×W/4的c维标记化输入输入到连续的两个Swin Transformer块中进行表示学习，特征维度和分辨率保持不变。同时，patch 合并层会减少token的数量(2× down sampling)，将特征维数增加到2×原始维数。此过程在编码器中重复了三次。

图示

描述已自动生成

图4.6 编码器

patch合并层：

  输入patch被分为四部分，通过Patch merging layer连接在一起。这样的处理会使特征分辨率下降2倍。并且，由于拼接操作的结果是特征维数增加了4倍，因此在拼接的特征上加一个线性层，将特征维数统一为原始维数的2倍。

Bottleneck：

  由于Transformer太深，无法收敛，因此只使用连续两个Swin Transformer块构造瓶颈来学习深度特征表示。在瓶颈处，特征维度和分辨率保持不变。

解码器:

  与编码器相对应的是基于Swin transformer模块的对称解码器。为此，与编码器中使用的patch 合并层不同，在解码器中使用patch扩展层对提取的深度特征进行上采样。patch 扩展层将相邻维度的特征图重塑为更高分辨率的特征图(2×上采样)，并相应地将特征维数减半。

图示

描述已自动生成

图4.6 解码器

patch扩展层:

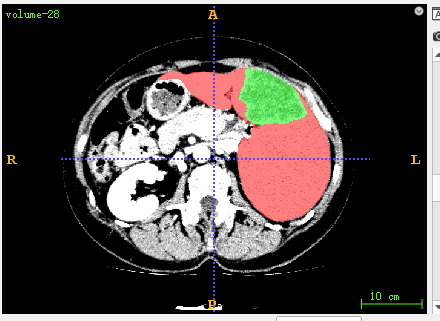
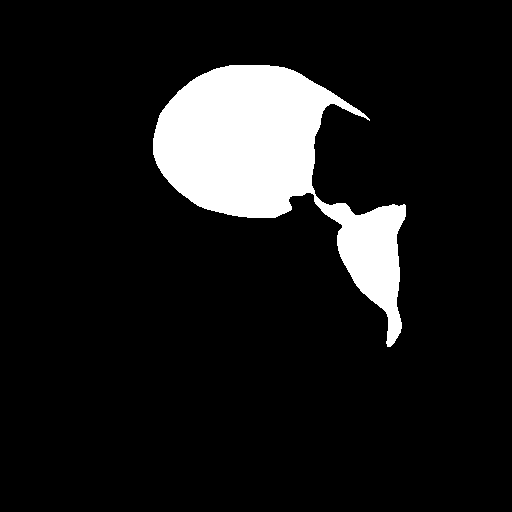
  以第一个patch扩展层为例，在上采样之前，在输入特征上加一个线性层(W/32×H/32×8C)，将特征维数增加到原始维数的2倍(W/32×H/32×16C)。然后，利用重排操作将输入特征的分辨率扩展为输入分辨率的2倍，将特征维数降低为输入维数的1/4 (W/32×H/32×16C→W/16×H/16×4C)。

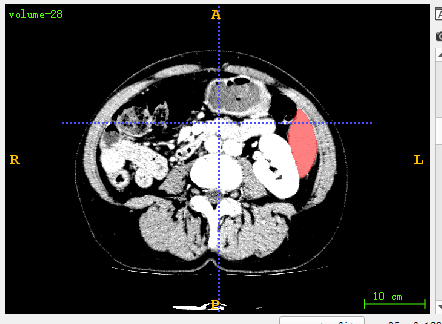
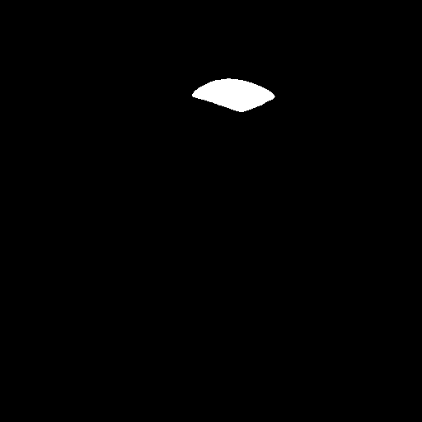
跳跃连接：

与U-Net类似，跳跃连接用于融合来自编码器的多尺度特征与上采样特征。将浅层特征和深层特征连接在一起，以减少降采样带来的空间信息损失。然后是一个线性层，连接特征的尺寸保持与上采样特征的尺寸相同。

### 实验结果

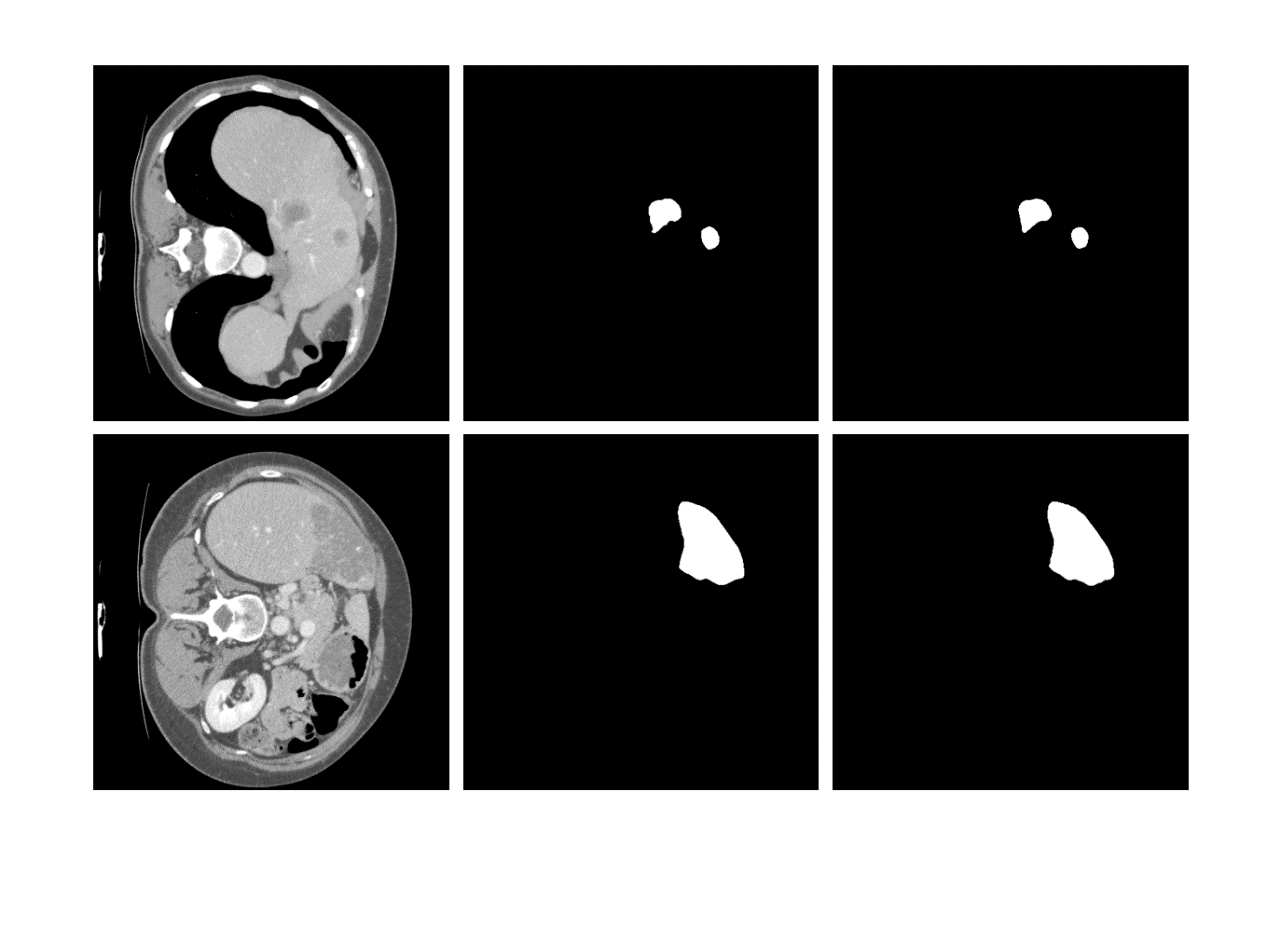
#### UNet预测结果

原始图像与标签 UNet预测肝分割图像

图5.1



原始图像 标签 UNet预测肿瘤分割图像

图5.2

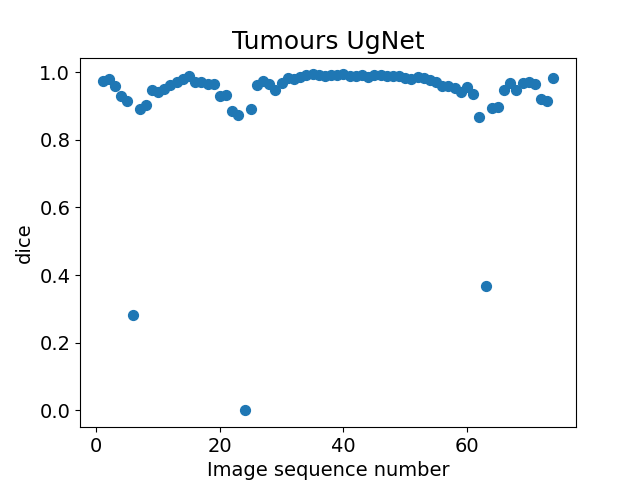
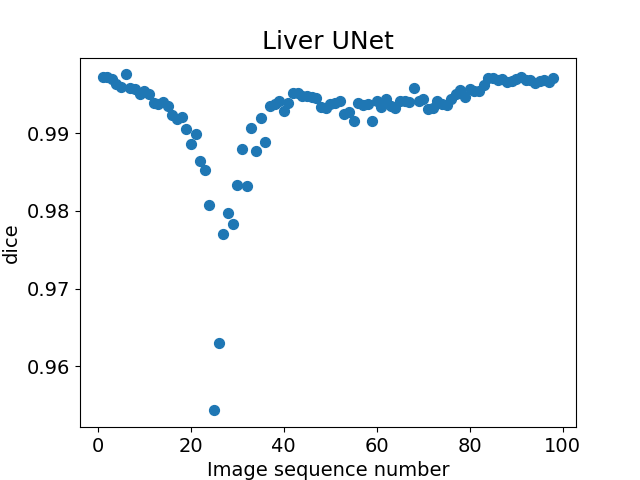
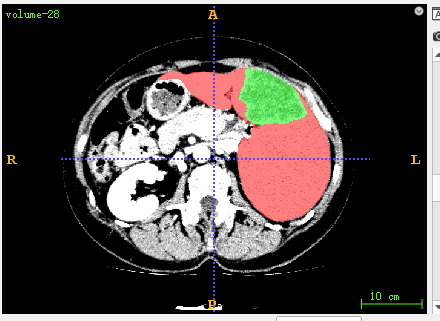
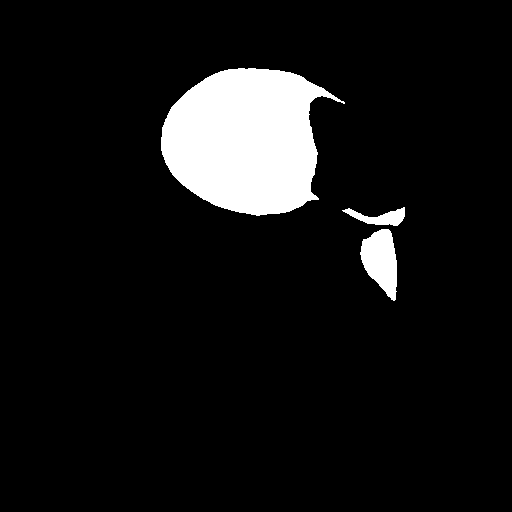
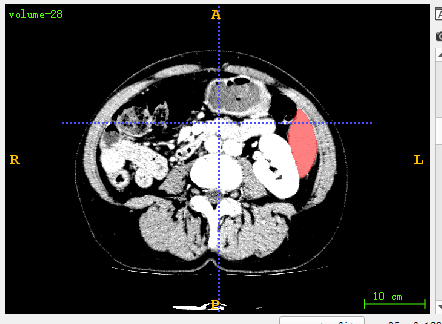
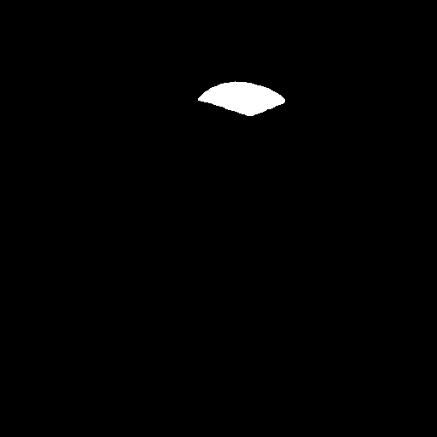


图5.3 UNet预测肝与肿瘤的Dice系数

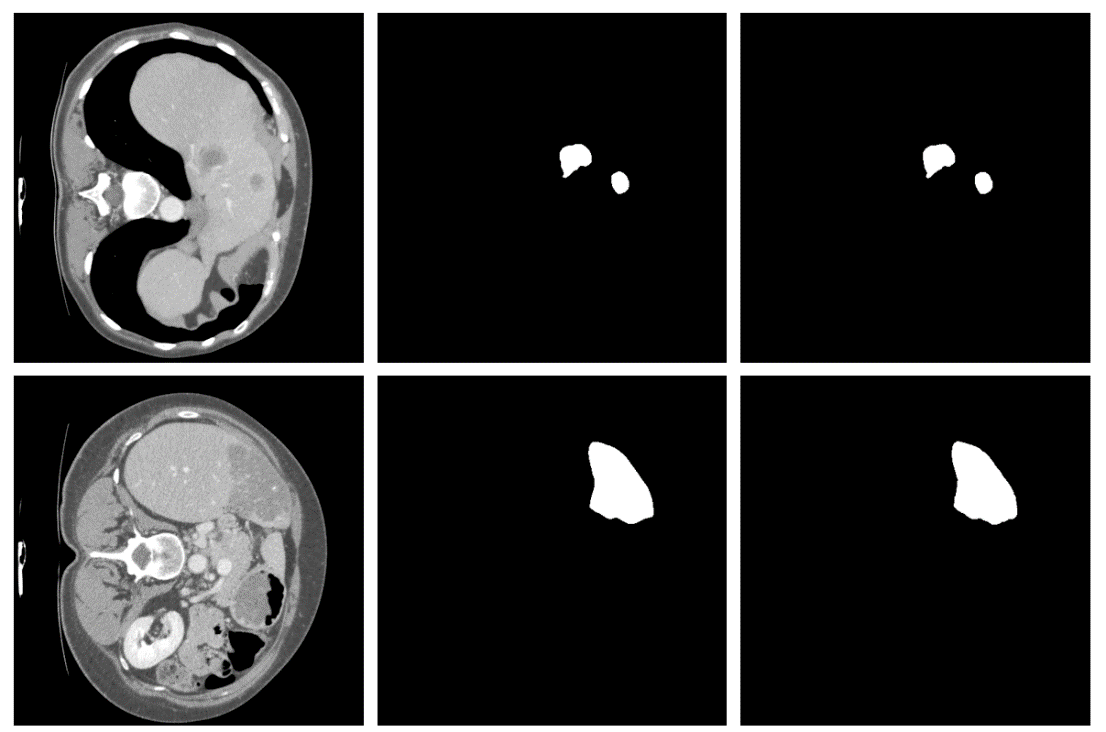
#### SegNet预测结果

原始图像与标签 SegNet预测肝分割图像

图5.4



原始图像 标签 SegNet预测肿瘤分割图像

图5.5

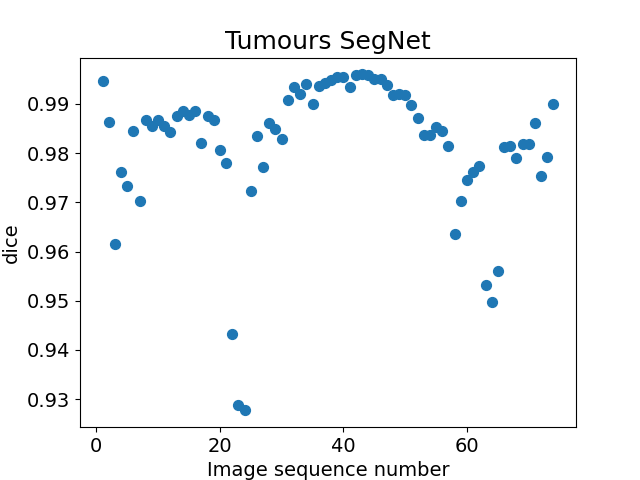
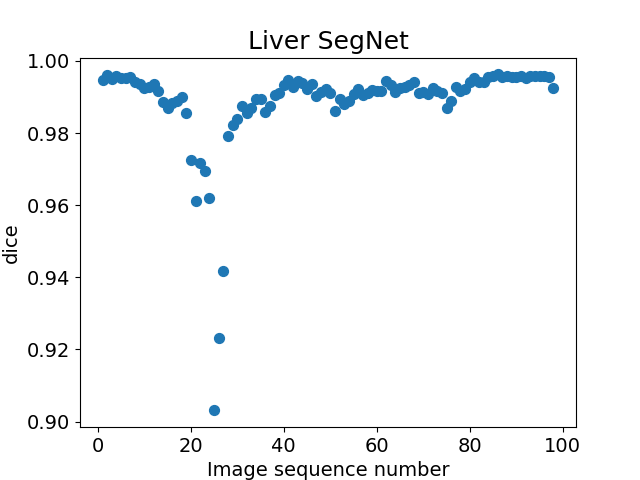


图5.6 SegNet预测肝与肿瘤的Dice系数

#### Swin-UNet预测结果

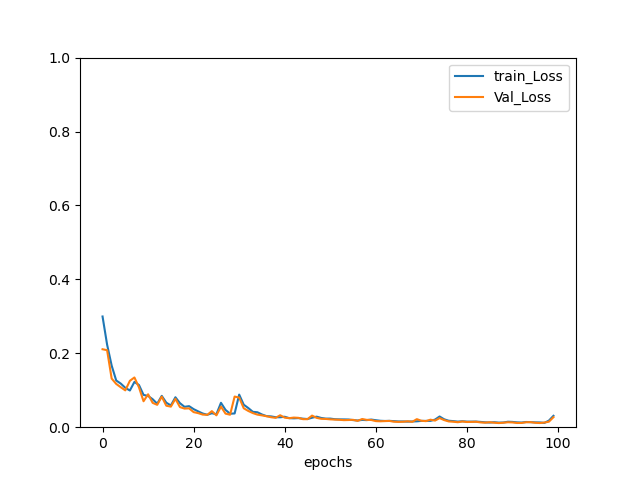
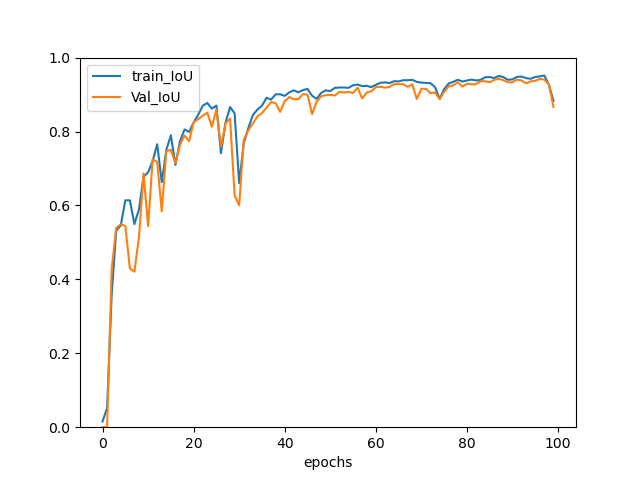


图5.6 Swin-UNet训练肝分割的IoU与loss系数

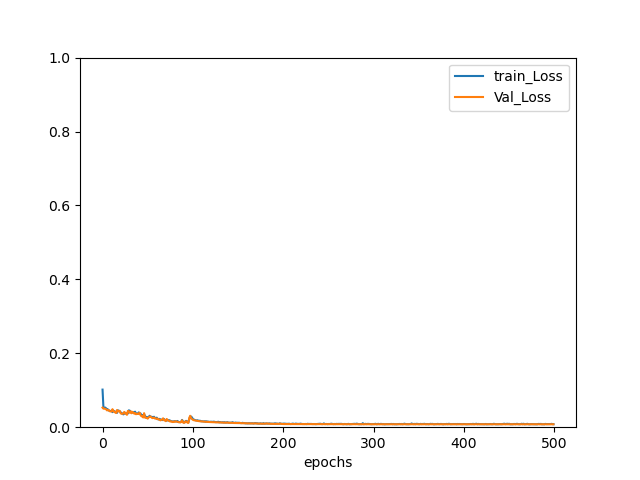
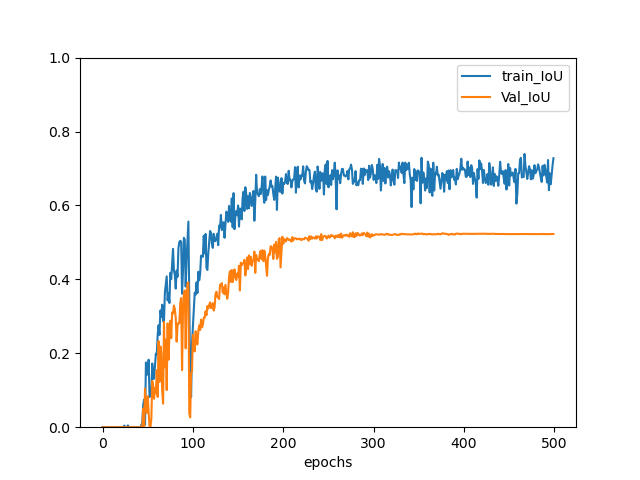
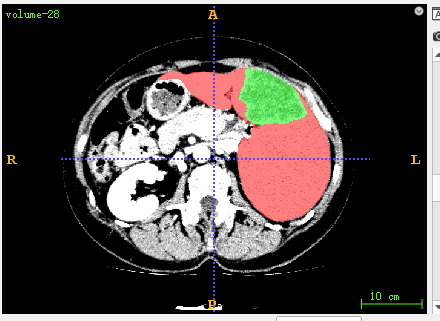
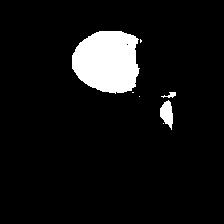
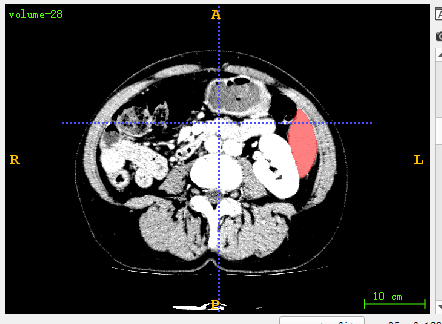
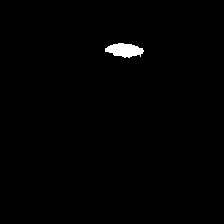


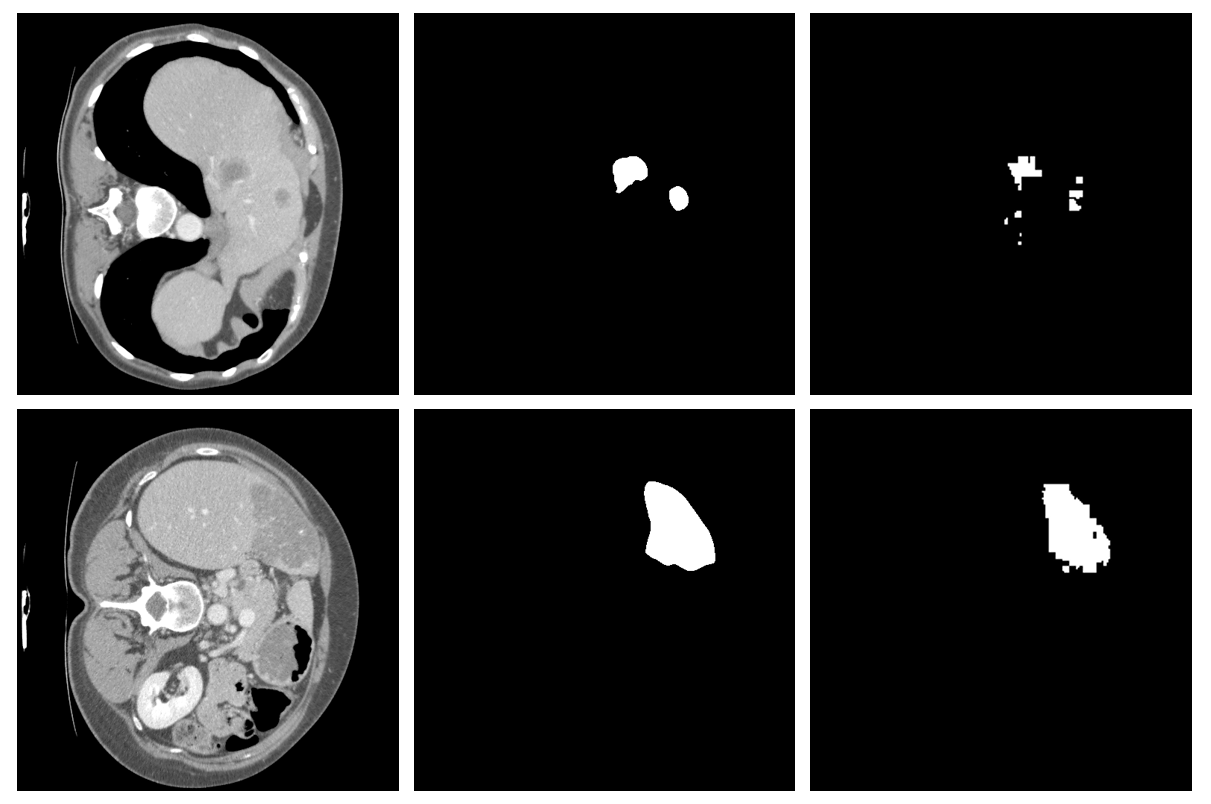
图5.7 Swin-UNet训练肿瘤分割的IoU与loss系数

原始图像与标签 SegNet预测肝分割图像

图5.8



原始图像 标签 Swin-UNet预测肿瘤分割图像

图5.9

表5.1 三种模型预测图片的平均Dice系数

|  |  |  |  |
| --- | --- | --- | --- |
|  | 平均dice | | |
|  | UNet | SegNet | Swin-UNet |
| 肝 | 0.9926055239595509 | 0.9885138801459568 | 0.5994900699009889 |
| 肿瘤 | 0.9292816462685554 | 0.9818620937126523 | 0.11053745865965918 |