基于卷积神经网络的直肠癌肿瘤淋巴转移智能识别技术

摘要

近年来，随着图像处理与人工智能技术的发展，应用基于医学影像大数据的分析方法来辅助医生决策或者解决临床实践中的棘手问题成为研究热点[1]。本文采用卷积神经网络，根据直肠癌患者的CT影像资料和患者的个人基本资料，来解决以下的问题，建立直肠癌肿瘤淋巴结智能识别技术。

问题一对患者CT影像的识别，首先采用卷积神经网络进行深度监督学习，对所有的CT影像进行特征提取。然后采用Softmax回归系数[2,3]判断对所提取的特征进行分类（含肿瘤区域、不含肿瘤区域），建立一个二分类模型。最后根据给出的107名直肠癌患者的CT数据，随机选取80%进行模型训练，20%来进行模型测试，来进行模型的训练与测试，并采用F-Score系数进行模型的评价。

问题二对肿瘤的精准切割，本文根据患者所有含有肿瘤的CT影像，采用基于全卷积神经网络的U-Net框架，直接对肿瘤区域进行精准切割。为了增强网络对肿瘤区域的特征学习增强模型的拟合度，将所有数据分为80%训练、20%测试，并对训练数据集进行数据集增强。将经过模型切割出来的肿瘤掩模图和给定的肿瘤掩模图进行计算，求出评价模型图像切割的Dice系数。

问题三对肿瘤区域的特征提取和淋巴结转移预测，这里采用CNN和Pyradiomics分别提取了肿瘤区域的卷积特征和放射学特征，并结合患者的个人基本信息年龄和性别（0/1：男/女），将这三种信息结合在一起使用随机森林（Random Forest）[4]进行分类模型的建立。选取80%的数据进行模型训练，20%的数据进行模型测试，建立合理的模型并求出评价系数F-Score进行模型评价。

最后，经过上述三问题的解决所提取的基于卷积神经网络的直肠癌肿瘤淋巴转移智能识别技术对医学影像的识别、切割和对直肠癌肿瘤淋巴转移识别有较大的精度，基本解决了所有问题。

关键字：直肠癌、卷积神经网络、U-Net、图像切割，Random Forest

目录

1 绪论 1
1.1 背景、目的及意义 1
1.2 相关工作 2
1.3 本文工作 2
2.数据预处理 3
2.1 数据变换 3
2.2 数据规约 4
3 基于 CNN 的肿瘤识别 5
3.1 训练样本分类 5
3.2 CNN 神经网络结构 5
3.2.1 卷积层 6
3.2.2 池化层 8
3.2.3 全连接层 8
3.3 CNN 模型训练 8
4 基于 U-Net 的肿瘤图像切割 11
4.1 FCN 与 CNN 对比 12
4.2 U-Net 神经网络结构 13
4.2.1 下卷积层 13
4.2.2 上采样层 14
4.2.3 输出层 14
4.3 U-Net 模型训练 15

5 肿瘤特征提取 16
5.1 卷积特征提取..... 17
5.2 放射学特征提取..... 18
6 预测是否发生淋巴结转移算法设计..... 18
6.1 算法设计 19
6.2Random-Forest 简介..... 19
6.3 模型训练 20
7 实验设计与分析 21
7.1 算法评价指标..... 21
7.2 实验结果与分析..... 22
8 总结展望 23
参考文献..... 24

1 绪论

1.1 背景、目的及意义

近几年在中国，直肠癌的发病率越来越高，特别在一些大城市，它已经跃居至恶性肿瘤发病率排行榜前三位。大多数（占 65%）发病在 40 岁以后，男女之比为 2～3：1。以 40～50 岁年龄组发病率最高。近年我国结直肠癌发病率呈明显上升趋势，而治愈率并未明显改善。各地资料显示，随着人民生活水平的提高，饮食结构的改变，其发病率呈逐年上升趋势。直肠癌易向肠外浸润并发生淋巴结及远处转移，一旦发生转移病人常常需先进行辅助放化疗才能获得手术机会，患者预后较早期直肠癌患者的预后差。直肠癌患者是否有淋巴结转移对治疗方案的决策以及病人预后有重要的影响，因此对是否有淋巴结转移的准确判断是直肠癌治疗的重要步骤。

几十年来，我国的医疗影像学发展迅速，影像检测已经成为了常见手段，但是相对应的医生增长速度尚不能满足市场需求，这对众多医院造成了比较大的困扰。影像检查的发展积累了海量的医学影像，医生平均需要花费 10 到 15 分钟来进行有效的诊断和报告，但是当医生读完十个影像片子时，会出现视觉疲劳，容易造成漏诊，人工智能通过模拟人类的思考方式来对图像进行识别，因此被给予了巨大的期望。针对直肠癌的 CT 影像，我们可以利用机器学习来自动判别患者的 CT 影像有误肿瘤区域，并且自动识别出肿瘤位置，根据 CT 肿瘤区域特征，智能判断直肠癌有无发生淋巴结转移。

图像识别技术在医学上主要应用在对患者病变部位的诊断[5]，帮助医生快速对患者病变部位做 出准确的诊断。随着 CT、核磁共振、X 射线等大型医疗影像设备的出现，图像识别技术成功应用在医学诊断上，但是，由于医疗行业的局限性，科研工作者们无法使用大量的医学图像作为原数据[6]，因此，建立大规模的医学图像数据集、优化算法模型是医学图像识别的主要研究方向[11]。

1.2 相关工作

在对医学影像进行识别时，学者们提出了许多特征提取与识别的方法，如：周平提出基于小波变换的共生特征提取方法，但忽略了对小波变换细节子带的利用[8]；韩彦芳提出在小波细节子带中使用 GLCM 提取纹理共生特征，但仅用小波细节子带不足以描述不同的纹理细节[9]；付增良等人提出 7 个共生特征的 CT 图像分割方法[10]。

也有许多学者提出针对深度学习的医学影像的识别与切割。吕鸿蒙等人[12]提出了一种基于增强 AlexNet 的阿尔兹海默病的早期诊断，增加神经网络结构的层数以及优化各层参数，使之达到良好的诊断 效果。实验结果证明，增强的 AlexNet 网络在诊断 AD 上更有优势，它的灵敏度达到 100%，但它的特异度小于原始模型，说明增强的 AlexNet 在误诊率方面要大于原有模型。Ghesu 等人[13]提出了一种基于边缘空间深度学习（MSDL）的图像目标检测与分割方法[14]，利用层次边缘空间中有效对象参数化的优点，结合深度学习，网络自动设计新的特征框架。

1.3 本文工作

本文通过对基于卷积神经网络的直肠癌肿瘤淋巴转移智能识别技术的实现，主要做了以下工作：根据提供的直肠癌患者 CT 影像和医生给定的肿瘤掩膜图像以及对应患者是否发生淋巴结转移情况的数据，利用数据挖掘解决如下两方面问题。

（1）直肠癌肿瘤的精准识别。利用现有的直肠癌患者数据，进行数据挖掘，建立合理的图像识别模型，可是判别出患者 CT 影像有无肿瘤，如果有，则识别出肿瘤位置，进行图像切割，给出对应肿瘤掩膜图。

（2）智能判断直肠肿瘤是否发生淋巴结转移。设计有效的算法通过对直肠癌 CT 影像特征的判断来对淋巴结转移情况进行评估，提高影像学对淋巴结转移判断的准确性。

并且以此安排了文章的章节，如图 1-1 所示。

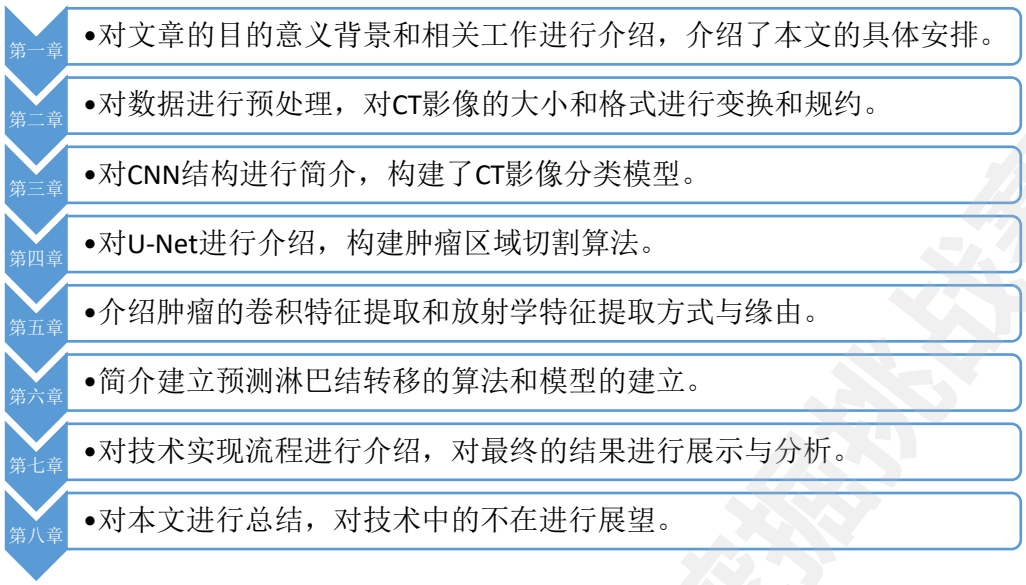

图 1-1 章节安排

2. 数据预处理

数据预处理是指在主要的处理以前对数据进行的一些处理。对于题目中给定的数据集，包括肿瘤患者的 CT 影像的 DCM 文件和 PNG 图像，以及患者的基本信息以及直肠癌淋巴结转移情况。对于这些原始数据，我们要做一些处理，将原始数据转换为预测模型易于处理的数据类型。

2.1 数据变换

题目中给定的样例数据，共包含 5 个病人的示例,每个病例包含动脉期和门脉期两套影像，分别保存在子文件夹“arterial phase”和“venous phase”中,每套影像中含有多幅 DICOM 格式的 CT 图像（后缀为.dcm），与每幅 CT 图像对应的肿瘤区域掩模图像（文件名中有后缀“_mask”，格式为 png）。如果某幅 CT 图像中不存在直肠肿瘤，则对应的掩模图像为全黑。所以针对给定的数据格式，我们本次研究一律使用动脉期的 CT 影像资料（arterial phase 文件夹内）进行研究。对于给的 DICOM 格式 CT 图像,这里我们引入 SimpleITK 医学图像处理工具，SimpleITK 是基于 ITK（一个开源的、跨平台的影像分析扩展软件工具）专门为 python 封装的函数包。我们利用 SimpleITK 和 Numpy 将 DICOM 文件读入，同时转换为一个二维数组。这里需要注意的是 SimpleITK. GetArrayFromImage()读

入的数据是一个三维坐标，坐标顺序是 z，y，x，只不过第一维 z 轴为 1。所以要用 Numpy 进行矩阵转置。同样转换为 Numpy.array 的三维数组[z,y,x],同样 z 轴数值为 1。

2.2 数据规约

在大数据集上进行复杂的数据分析和数据挖掘需要很长的时间，效率低下。数据规约产生更小的但保持原数据完整性的新数据集。在规约后的数据集上进行分析和挖掘将更有效率。

在题目中给出的样例数据，我们通过观察 png 格式的肿瘤图像（如图 2-1），图像大小均为 512*512 像素大小，但是图像周围大部分面积均为纯黑色填充，转换为数组矩阵后对应的 RGB 数值均为 0，所以这些区域对于图片特诊提取没有实际意义，而且所占区域高达图像的 2/3 区域，在进行建模分析时，大大降低了数据分析和挖掘效率。

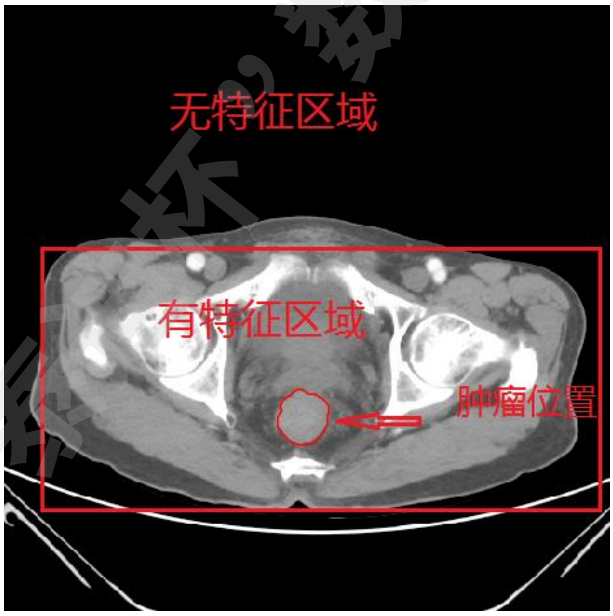

图 2-1 肿瘤示意图

在多次选择 DICOM 不同区域进行特征提取之后发现，在数组 y 轴长度为 230-430 区间内和 x 轴长度为 180-320 区间内的像素点特征提取最为明显。进行 Numpy 数组转置前，我们利用 Python 切片技术，截取三维数组[0,230:430,180:320]区间内的像素点，作为训练样本数据传入建立的训练模型中去。

3 基于 CNN 的肿瘤识别

卷积神经网络（Convolutional Neural Networks, CNN）是一类包含卷积计算且具有深度结构的前馈神经网络（Feedforward Neural Networks），是深度学习（deep learning）的代表算法之一。卷积神经网络仿造生物的视知觉机制构建，可以进行监督学习和非监督学习，其隐含层内的卷积核参数共享和层间连接的稀疏性使得卷积神经网络能够以较小的计算量对格点化特征，例如像素和音频进行学习、有稳定的效果且对数据没有额外的特征工程要求 。

3.1 训练样本分类

采用卷积神经网络进行建模分析时，我们采用监督学习对训练样本进行训练学习。针对题目给出的数据集中，每个病人给出的不同方位的 CT 影像，有的 CT 影像对应的掩膜图中没有肿瘤区域，所以我们要根据给定的 CT 影像掩膜图，对训练样本集进行分类，对 png 格式的掩膜图转换为数据矩阵，然后矩阵进行 sum 求和，如果结果为 0，即掩模图像为全黑，则说明该 CT 影像中没有发现肿瘤区域，然后根据图片名称编号找到对应 DICOM 文件，在进行读取该文件数据时，添加一个数据标签 label=0，标记没有癌症，相反，如果对掩膜图 sum 求和不为 0，则说明存在肿瘤区域，即添加一个数据标签 label=1。

3.2 CNN 神经网络结构.

卷积神经网络的低层是由卷积层和子采样层交替组成，在保持特征不变的情况下减少维度空间和计算时间，更高层次是全连接层，其输入是由卷积层和子采样层提取得到的特征，最后一层是输出层，可以是一个分类器，采用逻辑回归、Softmax 回归，支持向量机等进行分类，也可以直接输出一个结果。一个完整的卷积神经网络应包括三个阶段：第一阶段为卷积层，第二阶段为探索层（即激活层），第三阶段为池化层。模型结构如下图 3-1 所示。

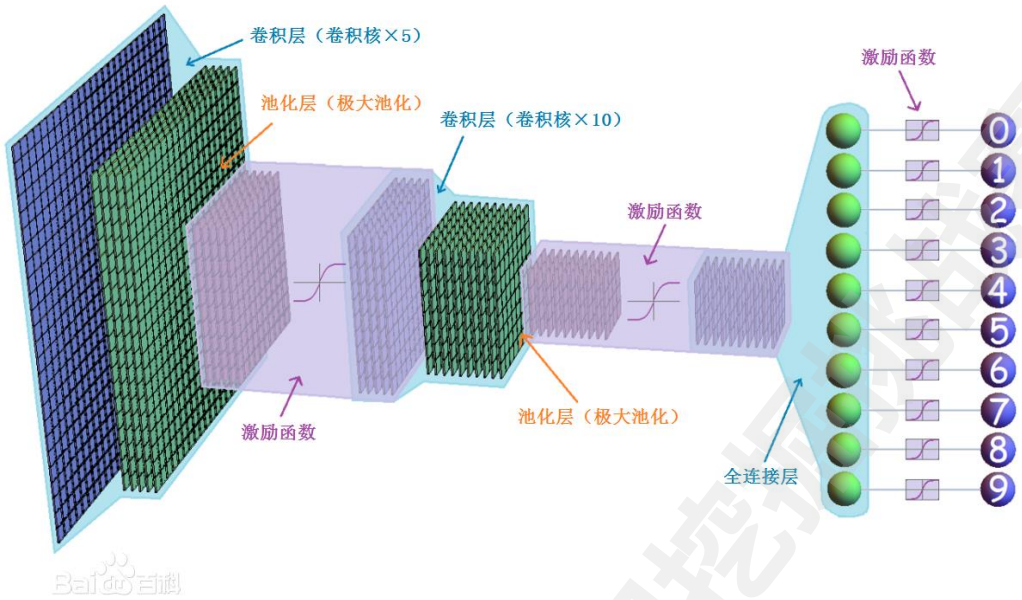

图 3-1 CNN 结构

3.2.1 卷积层

卷积层和子采样层是特征提取的核心模块，与其他前馈神经网络类似，卷积神经网络采用梯度下降的方法，应用最小损失函数对网络中的节点的权重参数逐层递减，通过反向递推，不断调整参数是的损失函数结果逐渐变小，从而提升整个网络的特征描绘能力，即对有肿瘤和图像识别和无肿瘤的图像识别的精准度和准确度不断提高。

通过卷积层的运算，可以将输入信号在某一特征上增强，从而实现特征的提取，也可以排除干扰因素，从而降低特征的噪声。卷积层同时对其输入使用多个过滤器，使之能够检测到输入的多个特征。具体来说，一个位于给定卷积层l的特征图k上的输入i行j列的神经元，与上一层$l-1$的神经元相连接，其位置为$i \times s_w$到$i \times s_w + f_w - 1$行，$j \times s_h$到$j \times s_h + f_h - 1$列，并且穿过$l-1$层中所有特征。注意所有位于不同特征图的i行j列的所有神经元都连接到前一层输出中完全相同的神经元。具体如公式 1：

$$z_{i,j,k} = b_k + \sum_{u=1}^{f_h} \cdot \sum_{v=1}^{f_w} \cdot \sum_{k'=1}^{f_{n'}} x_{x',j',k'} \cdot w_{u,v,k',k} \quad \text{其中} \begin{cases} i' = u \cdot s_h + f_h - 1 \\ j' = v \cdot s_w + f_w - 1 \end{cases} \qquad (1)$$

$z_{i,j,k}$是卷积层（l层）的特征图k的神经元在i行j列的输出。

s_h和s_w分别是水平方向和垂直方向的步幅，f_h和f_w分别是接受野的高和宽，$f_{n'}$是上一层（l − 1层）的特征数量。

$x_{x',j',k'}$是l − 1层的位于i'行和j'列的特征图k'（如果前一层是输入层，那么就是通道k'注：灰度图通道通常只有一个通道）上的神经元输出。

b_k是特征图k（l层）的偏执参数。即为矫正特征图k微调旋钮。

$w_{u,v,k',k}$是l层中任意特征图k和它位于特征图k'的u行v列的输入之间的连接权重。

卷积结束后，给与相应的激活函数。引入激活函数的主要目的是解决线性函数能力表达不够的问题。线性整流层作为神经网络的激活函数可以在不改变卷积层的情况下增强整个神经网络的非线性特征，不改变模型的泛化能力的同时提升训练速度。常见的卷积神经网络的激活函数有 Sigmoid、tanh、ReLU 函数。这里我们选择 ReLU 函数作为我们模型的激活函数，公式(2)如下：

$$f(x) = \max(0, x) \qquad (2)$$

ReLU 函数如图 3-2 所示。

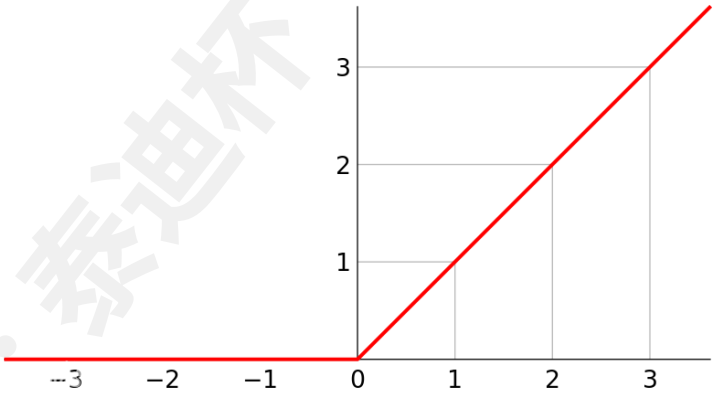

图 3-2 ReLU 函数

选择 ReLU 函数的原因是 ReLU 函数是非线性饱和函数，输出范围无限的，梯度下降速度快，训练时间更短。而 Sigmoid 和 tanh 函数是饱和线性函数，结果达到一定范围后不再变化。ReLU 则只需要一个阈值就可以得到激活函数，不需要对输入归一化来防止达到饱和。

3.2.2 池化层

池化层是一种向下采样的形式，在神经网络中也称之为子采样层。池化的结果是特征减少，参数减少，但其目的并不仅在于此。常用的池化方法有平均池化、最大池化、随机池化三种。在这里我们采用最大池化将特征区域的最大值作为新的抽取区域的值，减少数据空间大小。参数量和运算量也会减少，同时也减少全连接的数量和复杂度，一定程度上可以避免过拟合。

3.2.3 全连接层

经过几个卷积核池化之后，图像变得越来越小，但是由于卷积层，它变得越来越深（具有更多的特征）。在网络的结构的最顶层是由几个全连接层组成的常规前反馈神级网络，最后一层输出预测结果。

在 CNN 结构中，经多个卷积层和池化层后，连接着 1 个或 1 个以上的全连接层。全连接层中的每个神经元与其前一层的所有神经元进行全连接. 全连接层可以整合卷积层或者池化层中具有类别区分性的局部信息. 为了提升 CNN 网络性能，全连接层每个神经元的激励函数一般采用 ReLU 函数。最后一层全连接层的输出值被传递给一个输出，可以采用 Softmax 逻辑回归进行分类，或者直接输出一个结果，所以该层也可称为输出层。

对于一个具体的分类任务，选择一个合适的损失函数是十分重要的，损失函数评价模型对样本拟合度，预测结果与实际值越接近，说明模型的拟合能力越强，对应损失函数越小；反之，损失函数结果越大。损失函数比较大时，对应梯度下降比较快，则不利于特征提取。在我们的神经网络模型中，我们采用交叉熵损失函数与 softmax 损失函数相结合，TensorFlow 中将交叉熵损失函数和 softmax 统一封装实现了 softmax 后的交叉熵损失函数，TensorFlow 中可以直接使用如下代码设置：cross_entropy=tf.nn.softmax_cross_entory_with_logits(y,y_)其中 y_表示真实值，y 表示模型输出。

3.3 CNN 模型训练

根据我们数据预处理后的数据，我们传入神经网络的数据大小格式为 190 x 160 x 1 的三维数组，我们利用 TensorFlow 框架，先创建两个占位符 placeholder，

大小与输入数据格式匹配，用来存放训练数据（测试时存放测试数据）和训练数据标签。然后是三个卷积层 C1、C2 和 C3 和三个池化层 S1、S2 和 S3，定义卷积层 C1： 20 个卷积核，卷积核大小为 5 x 5，用 ReLU 函数激活，卷积层 C2：40 个卷积核，卷积核大小为 4 4，用 ReLU 函数激活，卷积层 C3：80 个卷积核，卷积核大小为 2 x 2，用 ReLU 函数激活；三个池化层均采用最大池化操作 maxpooling，pooling 窗口为 2x2，步长为 2x2。全连接层为两层，首先我们将第三层池化层输出的三维特征向量平铺展开转换为一维特征向量，然后进入第一个全连接层，我们依然采用 ReLU 函数激活，同时提取长度为 400 特征向量，第二层全连接层即为输出层，输出层基于上一层全连接层的结果，进行 logits 分类判断，输出层的另一项任务是进行反向传播，依次向后进行梯度传递，计算相应的损失函数，并重新更新权重值和偏执量。我们采用 TensorBoard 可视化工具对训练过程进行记录，流程图如下图 3-3：

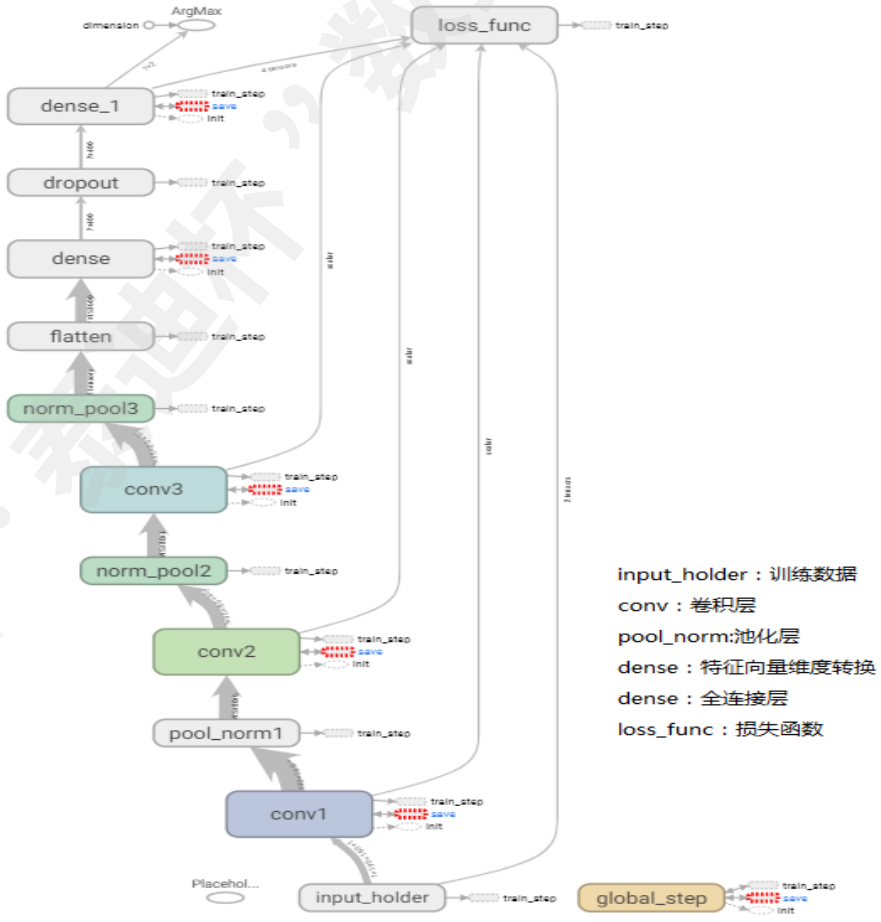

图 3-3 模型训练流程图

进行样本训练 20 次，每两步记录一次模型损失函数损失情况，同时计算三个卷积层的权重值和偏执量更新情况，统计及结果如下图 3-4 和 3-5：

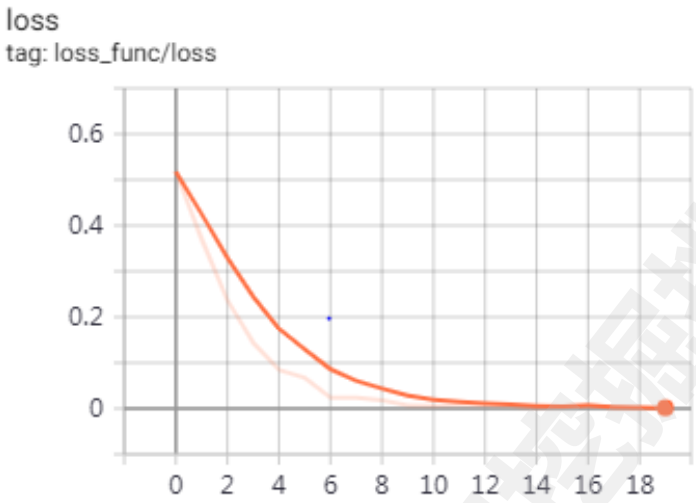

图 3-4 损失函数

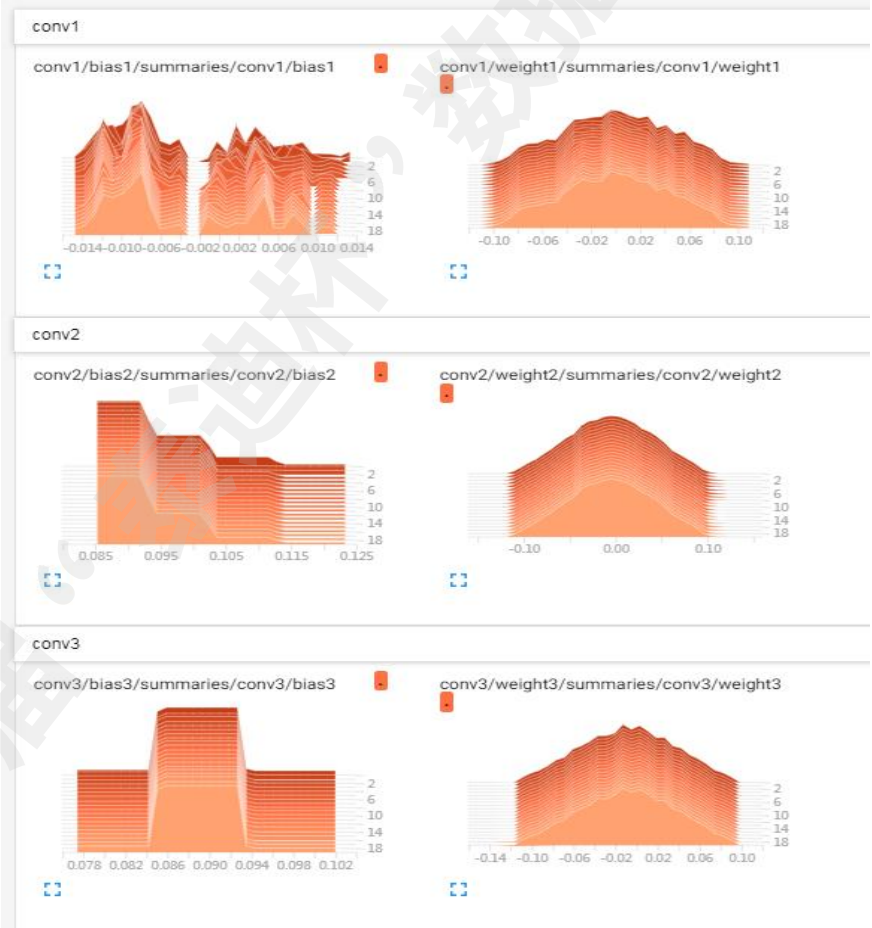

图 3-5 模型权重

通过损失函数记录情况可知，损失率梯度下降，最终趋近于 0，说明模型的

拟合比较强。但是在模型训练到第十次时，损失率已经接近于 0，继续训练模型怎会可能出现过拟合情况，则不利于模型的泛化，所以为了更多的训练模型次数，而不产生过拟合情况，我们引入防止过拟合方法，即在数据传入网络结构时，我们在创建数据和数据标签占位符 placeholder 时，在另外创建一个占位符 placeholder，用来存放 Dropout，数值大小为 0.2，Dropout 会在训练过程中随机选择百分之二十神经元的连接，从而使其不参与神经网络训练，降低神经元之间的耦合度。并且，由于减少了总的神经元数量，强制模型更多学习重要的特征。

此外学习率控制每次更新参数幅度，是比较重要的模型超参，过高过低的学习率都可能对模型结果带来不良影响，合适的学习率可以加快模型的训练速度。学习率太大会导致权重更新幅度太大可能会跨过损失函数极小值，导致参数值在极值点两边徘徊，即在极点值两端不断发散，或者是剧烈震荡，随着训练次数增大，而损失没有减小的趋势。如果学习率设置太小，参数更新速度太慢，导致无法过快的找到好的下降的方向，随着训练次数增加，而模型损失基本不变，需要消耗更多的训练资源来保证获取到参数的最优值。对于学习率的设置，刚开始更新的时候，学习率尽可能大，当参数快接近最优值时，学习率逐渐减小，保证参数最后能够达到最优值。而且希望训练的次数足够少，这样不仅加快训练速度，还可以减少资源的消耗。这里我们采用 Adam 算法自动调整模型的学习率，Adam 算法根据损失函数对每个参数的梯度的一阶矩估计和二阶矩估计动态调整针对于每个参数的学习速率。TensorFlow 提供的 tf.train.AdamOptimizer 可控制学习速度。Adam 也是基于梯度下降的方法，但是每次迭代参数的学习步长都有一个确定的范围，不会因为很大的梯度导致很大的学习步长，参数的值比较稳定。

4 基于 U-Net 的肿瘤图像切割

在图像分割中，机器必须将图像分割成不同的 segments，每个 segment 代表不同的实体。如下图 4-1、4-2 实例：

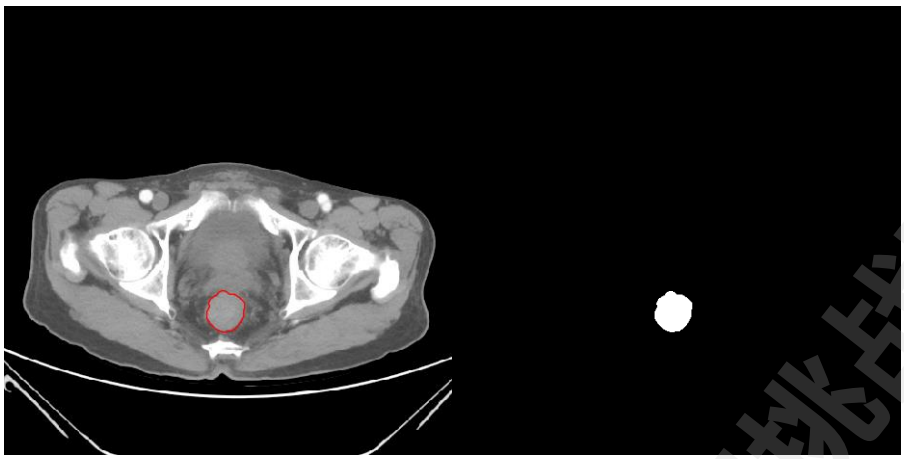

图 4-1　　图 4-2

如上图，图 4-1 中红色区域则为肿瘤实体，而图 4-2 中的掩模图则是我们在识别出肿瘤实体之后，对图像进行二值化处理，肿瘤区域像素全部改为 255，其他区域全 0 填充。

卷积神经网络在简单的图像分割问题上取得了不错的效果，但在复杂的图像分割问题上却没有取得任何进展。这就是 UNet 的作用。UNet 是在全卷积神经网络的基础上进行改进，最初是专门为医学图像分割而设计的。该方法取得了良好的效果，并在以后的许多领域得到了应用。

4.1 FCN 与 CNN 对比

U-Net 框架是在全卷积神经网络（FCN）的基础上进行改进。在此之前我们对于肿瘤的用在分类和检测问题上，我们使用卷积神经网络，由于用到 CNN，所以最后提取的特征的尺度是变小的。和我们要求的图像切割要求不一样，我们要求的图像是输入多大，输出有多大。为了让 CNN 提取出来的尺度能到原图大小，FCN 网络利用上采样和反卷积到原图像大小。然后做像素级的端到端的语义分割，所谓的“全卷积神经网络”是指该网络中没有全连接层，全部用全卷积和池化操作来代替，全卷积化也是深度学习模型的一个趋势，除了分类网络最后保留最后一个全连接层用来分类外，其他领域都有去掉全连接层的趋势。全卷积网络结构的重点是卷积化，在传统的神经网络中，为了输出分类标签，会使用全连接层将带有空间信息的二维图像（上一章中我们用的三维图片，只不过 z 轴数值为 1）压缩为一维，但是图像分割输出的结果要求是二维的，为了避免空间信息丢失，我们将全连接层替换为卷积层，即卷积化。

4.2 U-Net 神经网络结构

这里我们对应的 U-Net 模型结构如图 4-3：

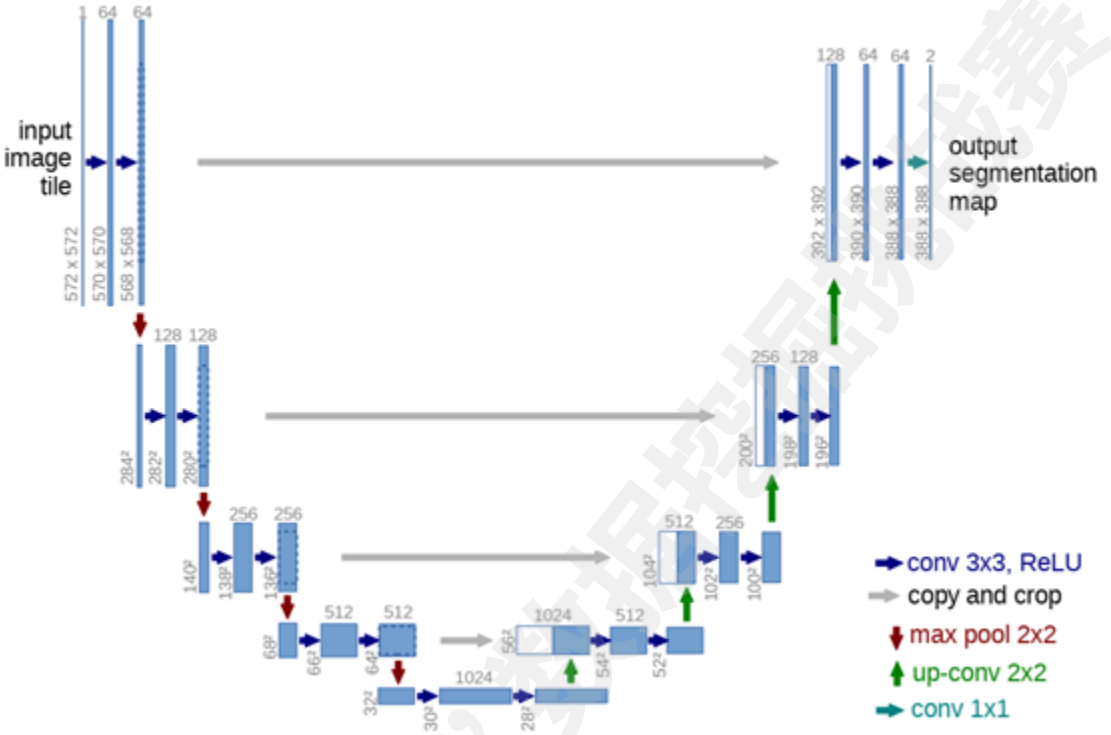

图 4-3 U-Net 模型

网络分为四个主要部分：preprocessing（预处理）、down convolution（下卷积）、up convolution（上卷积）、Output Map（输出映射）。

4.2.1 下卷积层

有三次下卷积层，一个下卷积层实际包括两次下卷积和一次 pooling 池化，在这里卷积和池化的意义与上述 CNN 结构操作一样。第一层卷积层设置 32 个卷积核，卷积核大小 3 x3；第二层卷积层设置 64 个卷积核，卷积核大小 3 x3；第三层卷积层设置 128 个卷积核，卷积核大小 3 x3；三层卷积层均采用最大池化操作 Maxpooling，pooling 窗口大小 2x2。三层下卷积层的激活函数同样采用 ReLu 函数激活。

经过三次下采样之后，为了加强模型训练，防止过拟合，继续进行两次卷积展开，卷积核数分别设置为 256 个和 512 个，卷积核大小 3 x3，但同时引入 Dropout，数值大小为 0.5，Dropout 意义与 CNN 中一样，会在训练过程中随机选择百分之五十神经元的连接，从而使其不参与神经网络训练，降低神经元之间的耦合度。

4.2.2 上采样层

上采样层又叫做反卷积层，顾名思义，反卷积层就是卷积层的逆操作。有三次反卷积层，一个反卷积层实际包括一个反卷积，一个连接操作和两次下卷积。我们这里的反卷积操作使用 Keras 框架提供的 UpSampling2D 方法，可以看作是 Pooling 的反向操作，就是采用 Nearest Neighbor interpolation（最近邻插值）来进行放大，就是复制行和列的数据来扩充特征映射的大小。

这里我们演示反卷积的操作过程，如图 4-4：

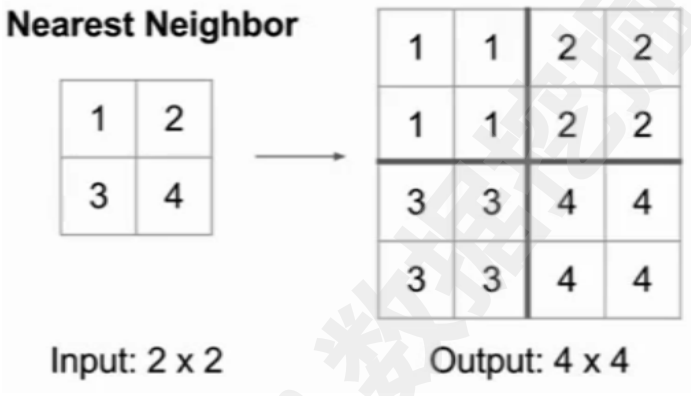

图 4-4 反卷积示意图

第一个反卷积层反卷积设置 256 个卷积核，pooling 大小 2 x 2，两次下卷积设置 256 个卷积核，卷积核大小设置 3 x3；第二个反卷积层反卷积设置 128 个卷积核，pooling 大小 2 2，两次下卷积设置 128 个卷积核，卷积核大小设置 3 x3；第一个反卷积层反卷积设置 64 个卷积核，pooling 大小 2 x 2，两次下卷积设置 64 个卷积核，卷积核大小设置 3 x3；三次反卷积连接操作均使用 Keras 框架提供的 Concatenate 函数，激活函数使用 ReLU 函数激活。

4.2.3 输出层

这里我们依然采用 Adam 算法自动调整模型的学习率， 在 Keras 中我们使用 compile 优化器来使用 Adam 算法自动调整模型学习率。

在这里我们需要注意，在数据预处理阶段，由于数据规约，我们使用的图片大小格式 256 x256，所以我们在保存肿瘤掩模图是，要将图像扩充到 512 x512 大小，扩充区域全 0 填充。

4.3 U-Net 模型训练

经过 CNN 神经网络对给定的患者 CT 影像处理，识别出有肿瘤的 CT 影像，将对应的 DICOM 文件和肿瘤掩模图一块传入 U-Net 神经网络中，进行模型训练，识别肿瘤位置，并给出对应肿瘤掩模图。

为了增加网络结构对肿瘤位置的准确识别和精准切割，我们对传入的训练数据进行数据集增强。数据集增强主要是为了减少网络的过拟合现象，通过对训练图片进行变换可以得到泛化能力更强的网络，更好的适应应用场景。常用的数据增强方法有：旋转，反射变换、翻转变换、缩放变换、平移变换、尺度变换、噪声扰动、颜色变化等。在这里我们自定义图片数据增强，我们拿其中一张肿瘤图像演示增强效果，原图如图 4-4，效果图如 4-5：

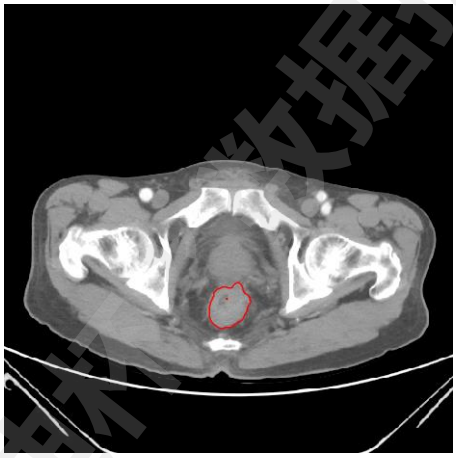

图 4-4 原始图像

增强效果图（这里做了缩放大小，并未改变像素大小）：

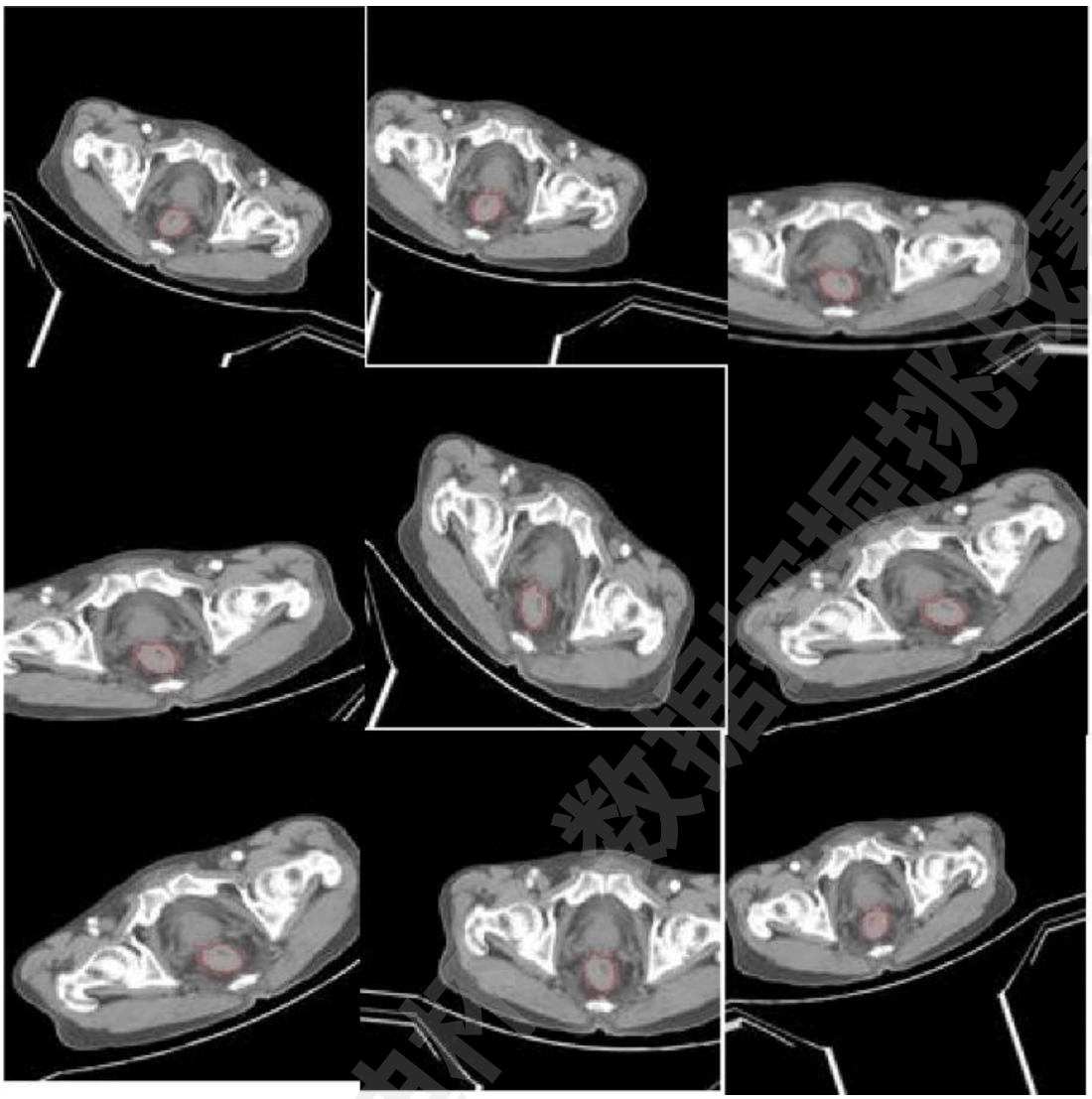

图 4-5 图像增强后示意图

我们对训练数据集进行数据增强后，我们将增强数据集传入神级网络，然后开始定义训练参数，我们这里基于 Keras 框架开始训练模型，定义模型训练 10 次，每次训练时，传入 4 张图片，防止训练数据集过大，导致硬件资源压力过大。并且每训练一次，则随机抽取百分之五数据，做样本测试，验证损失函数变化，自动调节学习率。

5 肿瘤特征提取

图像特征提取是计算机视觉和图像处理中的一个概念，是指采用计算机提取图像特征，主要任务便是提取出区别其他属性的特征。图像特征是标识一个图像最基本的概念，利用图像的属性差异可以与其它图像进行区分。对于第二问直肠

肿瘤部分特征的提取，主要提取了两大类特征：肿瘤部分卷积特征和肿瘤部分放射学特征。

5.1 卷积特征提取

卷积特征是采用卷积神经网络（CNN）进行的特征提取，由于图像具有一定得“不变性”，即图片中得一个部分与另一个部分经常使相同得。因此可以利用卷积神经网络中卷积和池化的计算性质，使得图像中平移部分对最后得特征向量没有影像，使提取的特征不容易过拟合。又由于不同得卷积、池化和最后输出得特征向量的大小可以控制整体模型的拟合能力，使得图片特征的拟合度可以控制，特征提取方法更加灵活可变。

在进行特征提取时将 CT 影像和对应的掩模图像进行合并，提取图像肿瘤区域，并将图像切割为 100x100 像素大小（如图 5-1 所示）这样可以最大限度的提取肿瘤区域的特征减少干扰项。在进行特征提取时建立了三层的网络结构，提取平均池化特征（average_pool2D），输出一个一定长度的特征向量（部分数据如表 5-1 所示）。

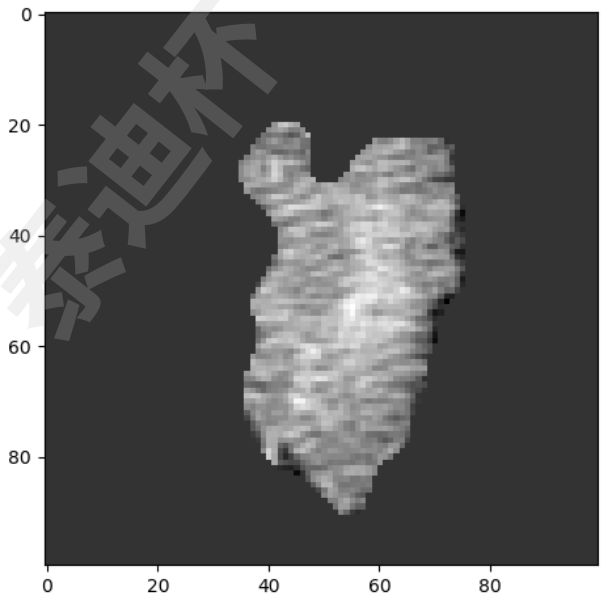

图 8-1 进行特征提取图片格式

对应图片	特征数据					
1008-10008	19.05	0	0	…	0	0
1010-10011	0	0	13.72	…	0	1.366

1012-10015	0.41	0	0	…	0.32	0
1053-10007	0	0	0	…	5.09	0

表 5-1 卷积特征数据

5.2 放射学特征提取

进行放射学特征提取是使用 Pyradiomics 包来进行特征提取，Pyradiomics 是一个开源的 python 包，用于医学图像的影像组学特征提取。采用这个包来进行 CT 影像的放射学特征提取，主要是提取了以下几方面的特征：一阶功能、形状特征（3D）、灰度共生矩阵（GLCM）特征、灰度级区域矩阵（GLSZM）功能、灰度级运行长度矩阵（GLRLM）功能、相邻灰度差分矩阵（NGTDM）特征、灰度依赖矩阵（GLDM）特征，共记提取了 110 中放射学特征。

在进行放射学特征提取时需要输入 CT 影像和对应的掩模图像，并且所输入的两张图片都是 dicom 格式。对于每张图片的放射学特征我们都在后面将其对应患者的年龄和性别（0/1：男/女）附加上去（部分特征数据如表 5-2 所示）。

特征名称	特征值
能量值	8548146.0
总能量	4488274.176724408
熵	1.8512256893637309
…	…
小区域高灰度重点（SAHGLE）	9.655042093144393
大面积低灰度强调（LALGLE）	289.5786185639757
大面积高灰度强调（LAHGLE)	195523.97902097902

表 5-2 放射学特征数据

6 预测是否发生淋巴结转移算法设计

淋巴结转移是肿瘤最常见的转移方式，是指浸润的肿瘤细胞穿过淋巴管壁，脱落后随淋巴液被带到汇流区淋巴结，并且以此为中心生长出同样肿瘤的现象。由于发生淋巴结转移对病人的治疗会产生巨大的影响并且不利于术后康复，因而

我们设计的算法便是通过对当前所给的患者信息和患者的一系列 CT 影像进行训练，建立合适的模型，可以直接通过对患者 CT 影像的识别来判断患者是否发生淋巴结转移。

6.1 算法设计

根据患者的 CT 影像来预测患者是否发生淋巴结转移，最终便是建立一个分类模型，根据患者 CT 影像的卷积特征和放射学特征建立一个二分类模型。在机器学习中有多种分类算法：决策树、KNN、支持向量机（SVM）、朴素贝叶斯、随机森林等，但由于所提取的数据集中含有大量无关数据且数据量较大、特征数据参杂在大量数据中特征不明显，因此采用随机森林这一适合于高纬度（feature 很多）的数据、不需要进行特征选择的分类方法。

6.2Random-Forest 简介

随机森林最早由 Leo Breiman 和 Adele Cutler 提出的，是一种统计学习理论，它是利用 bootsrap 重抽样方法从原始样本中抽取多个样本，对每个 bootsrap 样本进行决策树建模，然后组合多棵决策树的预测，通过投票得出最终预测结果。算法对异常值和噪声具有很好的容忍度，且不容易出现过拟合。

随机森林分类是由很多决策树分类模型组成的组合分类模型，其基本思想：首先，利用 bootstrap 抽样从原始训练集抽取 k 个样本，且每个样本容量都与原始训练集一样；其次，对 k 个样本分别建立 k 个决策树模型，得到 k 种分类结果；最后，根据 k 种分类结果对每个记录进行投票表决决定其最终分类（如图 6-1）。

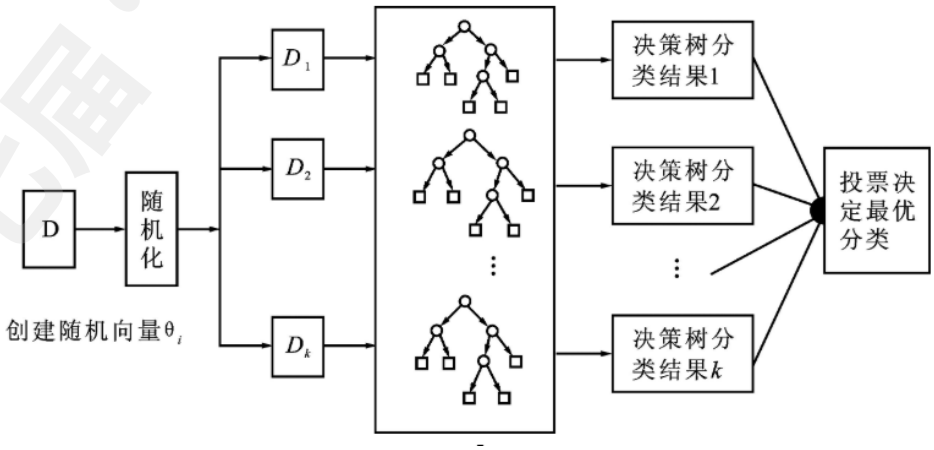

图 6-1 随机森林示意图

随机森林算法通过构造不同的训练集增加模型间的差异，从而提高组合分类模型的外推预测能力。通过 k 轮训练，得到模型序列$\{h_1(x), h_2(x), \cdots h_k(x)\}$，再用它们构成一个多分类模型系统，该系统的最终分类结果采用简单多数投票法。最终的分类决策：

$$\mathrm{H(x)} = \arg\ max \sum_{i=1}^{k} I(h_i(x) = Y) \qquad (3)$$

其中，H(x)表示组合分类模型，hi 是单个决策树分类模型，Y 表示输出变量(或称目标变量)，I(.)为示性函数。式（3）说明了使用多数投票决策的方式来确定最终的分类。

随机森林是由多棵决策树组合而成的，决策树算法含有 ID3 算法、C4.5 算法，CART 算法因此在随机森林的决策选取上也有多种随机特征选取的方法。

6.3 模型训练

所采用的是 Tensorflow 开源框架和 Scikit-Learn 框架进行搭建的，先将获取的卷积特征和放射学特征进行合并，并输入随机森林中进行预测，选取的决策树为‘aiout’，通过不断修改随机森林中决策树的数量来获取一个最佳模型，折线图如图所示 6-1

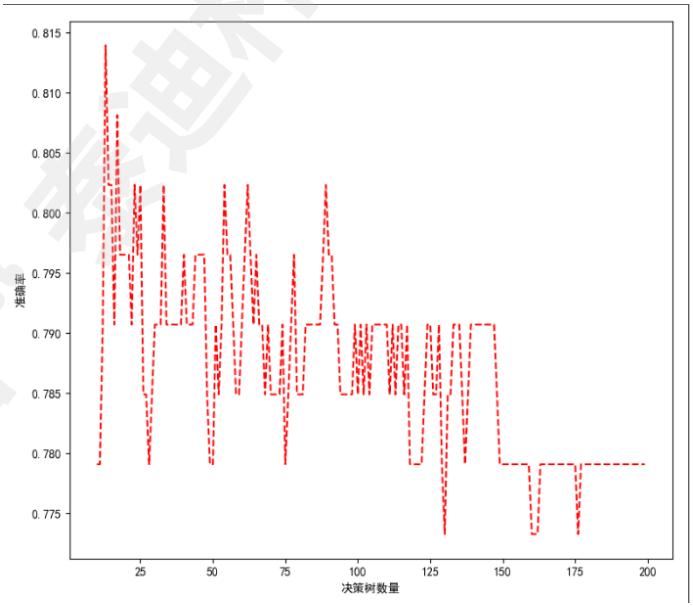

图 6-1 决策树数量于准确率图

7 实验设计与分析

对分类模型效果的评估是是整个系统重要的一部分，通常给定一个数据集 R，将其分为训练数据集R_{train}和测试数据集R_{test}，而且$R_{train} \cup R_{test} = R$。采用训练数据集建立分类模型，将测试数据集中每一张图片都定义为一个患者，测试出每一张图片对应的结果后，在采用投票的方式根据一个患者的多张图片来判断患者的分类。建立分类模型依据的便是提取的特征值，而提取特征值关键在于对图像进行精准切割。对图像进行切割时由于为了模型的准确度，先对所有的图片进行分类，将图片分为可画出肿瘤区域的和画不出肿瘤区域的，然后在根据分类结果对图像进行切割。在建立图像切割模型时为了加强拟合度需要对数据经增强处理，选取前 80%的图片为测试数据集进行增加处理，图像经过平移、选择变换后每一张图片扩充为原图片的 10 倍，然后采用这些图片进行模型的建立，使用后 80%的图片进行测试，算法具体流程如图 7-1。

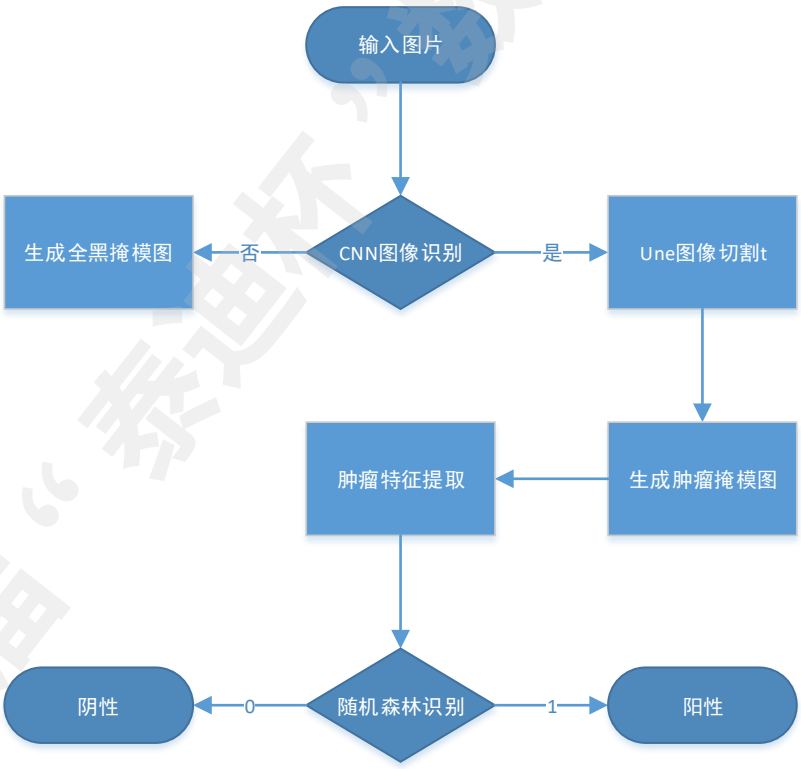

图 7-1 算法流程图

7.1 算法评价指标

对算法的评价包含对图像切割的评价、CT 影像分类模型和预测淋巴结转移的

评价。

图像分割评价采用 Dice 系数进行评价，它是一种集合相似度度量函数，通常用于计算两个样本的相似度：

$$\mathrm{Dice(A,B)} = \frac{2|A \cap B|}{|A| + |B|} \qquad (4)$$

其中A是医生勾画的直肠肿瘤区域，B是算法分割得到的直肠肿瘤区域。Dice 系数的取值范 围是[0,1]，取值越接近 1 表明直肠肿瘤分割的结果与医生给出的结果越接近。

对 CT 影像识别和淋巴结转移的评价指标是使用 F-Score 对分类结果进行评价：

$$\mathrm{F} = \frac{2PR}{P + R} \qquad (5)$$

其中P为查准率（Precision），R为查全率（Recall）。

7.2 实验结果与分析

本文算法的核心系统采用用 JetBrains PyCharm 2018.2.4 x64 编写，程序运行环境为 Windows 10 操作系统，文中的卷积神经网络采用 Google 开源框架 Tensorflow 编写的，U-Net 采用 Kreas 进行编写的，随机森林采用 Scikit-Learn 进行编写。算法的训练测试数据集都是进行规约处理后的图片，在使用 CNN 进行测试时采用平均损失函数 reduce_mean()，在采用 U-Nent 进行图像切割时定义损失函数 Adam(lr=1e-4)，使用随机森林进行测试时定义决策树的数量为 13 进行测试。通过对问题的合并于分解，我们重新将问题分解为三部分：图像识别、图像切割、图像特征提取于分类。

问题一 CT 影像识别，我们随机选取所有图片的 80%进行模型训练，20%进行模型测试，通过测试出来的结果采用公式（5）计算出 F-Score=0.924，模型基本满足技术需求。

问题二图像切割，图像肿瘤切割的最后结果便是对所有图片绘制出掩模图，其中无法标记处肿瘤区域的掩模图为全黑，可以标记出肿瘤区域的掩模图中肿瘤区域为白色，部分测试结果如图 7-2 所示。其中 1087 为患者 id，10017 为患者对应的掩模图，含有’_pred’的是通过算法切割出来的掩模图。采用 dice 系数来计算肿瘤区域切割的准确度，通过对 172 张测试图片的计算，取其平均值

dice=0.853。

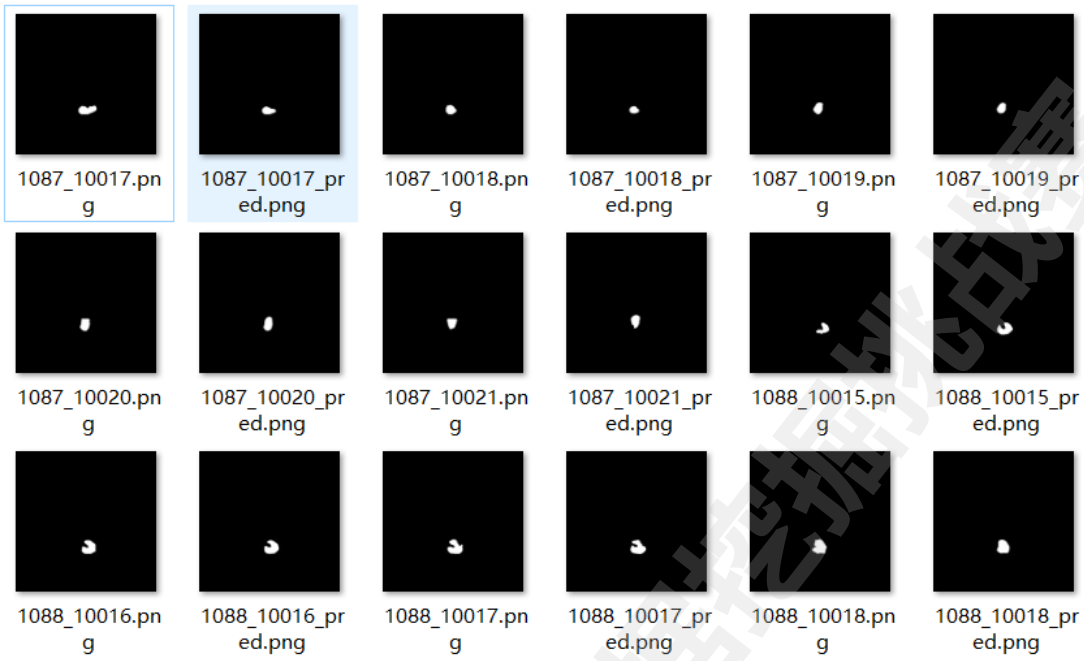

图 7-2 切割图片于原始图片

对于问题三，我们将肿瘤特征提取于图像识别结合到一起，我们对每一张图片进行肿瘤特征的提取，采用随机森林对 20%的患者进行预测，通过计算得出评价预测淋巴结转移的评价指数 F-Score=0.864。

8 总结展望

对淋巴结转移预测的关键是肿瘤区域的切割，但是由于所建立模型的缺陷，因此肿瘤区域的切割分了两部分：对图片的识别、图片切割。在图像识别模型只有 90%左右的成功率因此在对图像切割和淋巴结转移预测方面会有较大影像，由于这些影像会使模型的准确率大幅度下降。

针对此问题我们计划设计一个新的模型直接对 CT 影像进行切割，避免因为分类带来的较大误差，同时针对预测淋巴结转移方面也计划直接对患者进行预测，将患者的所有信息进行归一化处理，不考虑患者的 CT 影像数量，减少投票表决带来的误差。对于模型的预测仍然采用 dice 系数和 F-Score 进行评价。

在神经网络方面，卷积神经网络仍是为了发展的趋势，我们在本算法中所采用的 CNN 和 U-Net 是借鉴已经成熟的架构，我们所作的改进只是对原始数据以及模型参数进行的改动，希望在往后可以搭建一个根据适合于 CT 影像切割的模型。

参考文献

[1] 魏炜、刘振宇、王硕、田捷，影像组学技术研究进展及其在结直肠癌中的临床应用， 中国生物医学工程学报，2018，37:513-520.

[2] https://tensorflow.google.cn/

[3] https://keras.io/zh/

[4] Aurelien Geron. 机器学习实战：基于 Scikit-Learn 和 TensorFlow[M] 机械工业出版社，2018.7.

[5] 赵卫东,董亮. 机器学习[M]，人们邮电出版社，2018.8

[6] 杨云，杜飞. 深度学习实战[M]，清华大学出版社，2018

[7] 张平. OpenCV 算法精解：基于 Python 和 C++[M]，电子工业出版社，2017.10

[8] 韩彦芳，施鹏飞.基于多层小波和共生矩阵的纹理表面缺 损检测[J]. 上海交通大学学报，2006，40（3）：425-430.

[9] 付增良，陈晓军，叶铭，等.心脏 CT 图像分割方法[J].计算机工程，2009，35（12）：189-191.

[10] 赵一凡，夏良正.基于轮廓波特征的纹理图像识别方法[J]. 东南大学学报：自然科学版，2008，38（2）

[11] 刘吉，孙仁诚，乔松林. 深度学习在医学图像识别 中的应用研究[J]. 青岛大学学报(自然科学版)，2018(01)：69-74.

[12] 吕鸿蒙，赵地，迟学斌. 基于增强 AlexNet 的深度学 习的阿尔茨海默病的早期诊断[J]. 计算机科学，2017，44(S1):50-60.

[13] H. Z, B. Y, T. F, et al. Learning Deep Features for Classification of Typical Ecological Environmental Elements in High-Resolution Remote Sensing Images[C]// The Proceedings of 10th International Symposium on Computational Intelligence and Design (ISCID). Hangzhou: IEEE, 2017: 1.223-227.

[14] Z. H, G. C, H. W, et al. Building extraction from multi-source remote sensing images via deep deconvolution neural networks[C]// The Proceedings of IEEE International Geoscience and Remote Sensing Symposium

(IGARSS). Beijing: IEEE, 2016:1835-1838.