

直肠癌淋巴结转移的智能诊断

April 2019

Contents

1	摘要	3
2	简介	4
2.1	挖掘背景	4
2.2	挖掘目标	4
3	数据预处理	4
3.1	DCM格式简介	4
3.2	数据集简介	5
3.3	数据预处理	5
4	实现方法	6
4.1	肿瘤区域分割	6
4.1.1	相关工作	6
4.1.2	U-net	8
4.1.3	实验细节	9
4.2	肿瘤区域特征提取	9
4.2.1	强度特征	10
4.2.2	面积特征	11
4.3	直肠癌淋巴结转移诊断	13
4.3.1	全连接神经网络	13
4.3.2	k-Nearest Neighbor	14
5	模型优化	15

6 实验平台	15
7 参考文献	16

1 摘要

本文为解决直肠肿瘤 [4]淋巴结转移 [5]的预测首先利用卷积神经网络对病人的肿瘤CT影像进行划分，通过端对端的U-net我们实现了对每一个像素点是否属于肿瘤区域的二分类任务，在获得肿瘤区域之后，利用现有数据集提供的信息，在对病人CT影像PixelSpacing、KVP、mAS值进行了统一后，对病人肿瘤区域的大小特征、强度特征进行了提取，并发现肿瘤区域的像素强度值明显高于非肿瘤区域，这为肿瘤区域的识别提供了帮助。最后分别利用对MNIST网络进行修改得到的全连接神经网络和K-最近邻方法对直肠肿瘤淋巴结转移情况进行了预测。

关键词：CNN Sobel算子 k-Nearest Neighbor MNIST [2] 回归分析

2 简介

直肠癌是指齿状线到直肠乙状结肠交界处的一种癌症，它是我国最高发的癌症之一。直肠癌的治愈率根据肿瘤所处阶段的不同而定，早期和中期的直肠癌患者通过合理的治疗大概率可以达到临床治愈；一旦直肠癌向肠外浸润并发生向淋巴结和其他部位的转移，即进入直肠癌晚期，这一阶段的患者即使通过辅助放化疗，手术机会的获取率和术后的治愈率都较低。在我国，结直肠癌的发病率和死亡率都高于世界平均水平，搭建出可以检测出结直肠癌存在和判断癌症转移情况的图像识别模型成为一种迫切的需要。在临床上，它帮助减轻影像科医生的负担，并显著降低因为医生的疏忽或劳累而产生的误诊的概率，并在一定程度上提高各期直肠癌患者的存活率。

2.1 挖掘背景

2.2 挖掘目标

- 以数据集提供的CT影像和医生标记出的直肠肿瘤掩模为训练样本，设计图像分割算法，分割出直肠肿瘤所在的区域
- 提取CT图像中的直肠肿瘤区域的影像特征
- 根据数据集提供的数据分析直肠肿瘤区域影像特征与是否淋巴结转移之间的关系

3 数据预处理

3.1 DCM格式简介

DCM文件是遵循DICOM (医疗数位影像传输协定) [6]标准的一种文件，而DICOM标准支持的设备包括心电图、核磁共振成像、血管镜、超声心动图等多种医疗设备，因而DCM文件被广泛应用于医疗行业。DICOM文件可以大致分为两部分，与图像相关的元信息，包括患者信息，检查信息，序列信息，图像信息等等。另一部分包含图像的像素数据。在解析DICOM文件中的像素数据的时候，我们先需要读取以下图像相关信息：

- Samples per Pixel: 一个像素的取样数
- Rows: 图像的高度

- Columns:图像的宽度
- Pixel Spacing:图像像素间距，主要用于长度测量
- Window Center:窗宽，指CT图像所显示的CT 值范围，窗宽能反映该组织或病变的CT 值变化范围
- Window Width: 窗位，指窗宽范围内均值或中心值，窗位应该等于或接近需要观察的CT 值
- KVP: 管电压(这里所有病人CT影像文件KVP值均为120)
- mAS: 管电流

3.2 数据集简介

数据集包括临床数据和CT影像数据。临床数据包含病人ID、性别、年龄、是否淋巴结转移，CT影像提供各个病人的腹部横断位动脉期和门脉期两种增强CT影像数据以对应的由医生标记出的直肠肿瘤的掩模图像。掩模图像是一个与CT图像具有相同分辨率的二值灰度图像，其中像素值为1表示该像素属于直肠肿瘤区域，像素值为0 表示该像素属于非直肠肿瘤区域。

3.3 数据预处理

- 原数据集所提供尺寸为 512×512 的CT影像与掩模图片，而结肠是大肠的末端位于肛门上方，结肠横截面与CT扫描平面基本平行，所以可能的直肠肿瘤区域只位于CT 影像中的一特定区域。为了减小模型的计算复杂度，首先我们将原CT 影像与掩模裁剪为大小为 200×200 的图像，如图Fig.1
- 由于直肠肿瘤是贴附于直肠内壁，所以肿瘤的横断面面积应该是一个从小到大再变小的连续变化，而根据对数据集的观察我们发现原数据集所提供的病人CT 影像第一次标注出肿瘤区域时，肿瘤区域便已比较明显，所以我们猜测原数据集提供的标注并不完整，针对这个问题，为了提高之后的学习效果，我们仅将距离包含肿瘤区域的影像较近的数据进行了剔除。
- 实际应用中会根据患者年龄、性别等因素对CT管电压(KVP)、管电流(mAS)进行调整，从而导致不同患者CT影像的像素值明显不同，为消除因管电压、管电流引起的差异，我们将患者的CT影像针对管电压、管电流做了标准化处理，使得不同患者的CT影像不存在像素值强度方面的差异。

- 为了提高模型的“Robust”我们对经过上两步操作得到的影像进行了旋转变换，得到了原影像经 0° ， 120° ， 240° 旋转后的影像如图2



Figure 1: 裁剪

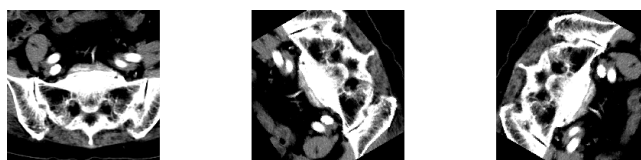


Figure 2: 旋转 0° 、 120° 、 240°

4 实现方法

4.1 肿瘤区域分割

结直肠癌在CT上可表现为局限性腔内软组织肿块影、肠壁局限性或全周性增生。肿瘤密度一般均匀，较大肿瘤可因缺血坏死而出现局灶性低密度区。肿瘤常呈分叶状，不对称。管壁呈环形增厚，在横断面上呈“炸环”样。CT强化后检查病变，增强程度决定于造影剂的量及浓度、注药速度及扫描速度。

在此我们做出如下假设：

- 数据集中提供的病人CT影像扫描、增强条件相同，且CT扫描切片厚度相同。
- 数据集提供的CT影像采用的比例尺相同，即1个像素点所体现的实际面积相同

4.1.1 相关工作

卷积神经网络(Convolutional Neural Networks, CNN) Fig.3是多层感知机(MLP)的变种。由生物学家休博尔和维瑟尔在早期关于视觉皮层的研究发展

而来。卷积神经网络是一种多层神经网络，擅长处理图像特别是大图像的相关机器学习问题。一个典型的卷积神经网络包括卷积层、池化层、全连接层，其中卷积层与池化层配合，组成多个卷积组，逐层提取特征，最终通过若干个全连接层完成分类。CNN最大的特点就是通过稀疏连接、各层权值共享降低了网络参数的数量级，从而能在当前条件下得到令人满意的结果。

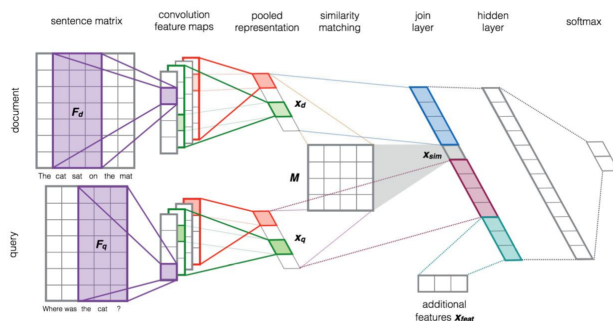


Figure 3: CNN

全卷积网络(Fully Convolutional Networks for Semantic Segmentation, FCN) [3] Fig.4第一次将深度学习与图像分割结合起来。在此之前深度学习一般用在分类和检测问题上。但是由于CNN各层的feature map都在不断缩小，所以最后提取的特征的尺度是变小的，无法得到与输入图像尺寸相同的输出结果。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 [9]将feature map还原到原图像大小，然后做像素级的分类，也就是说将分割问题变成分类问题，而分类问题正好是深度学习的强项。

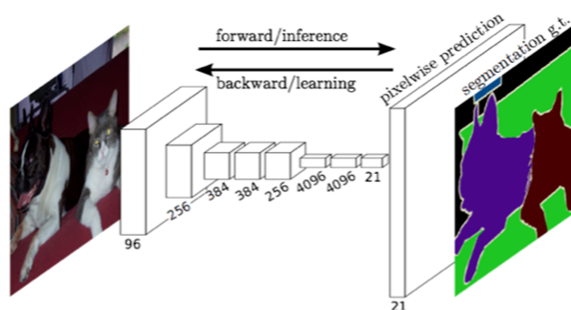


Figure 4: FCN

但单纯的将feature map上采样或者反卷积，明显会丢失很多特征信息。针对这一问题FCN采取解决方法是将池化层与得到的feature map融合起来，由于

池化层与feature map大小尺寸是不一样的，所以融合前应该将二者上采样到同一尺寸，这里的融合是拼接在一起，不是对应元素相加。

U-net网络采用交叉熵作为损失函数

$$L = \sum_{i=1}^N y^i \log \hat{y}^i + (1 - y^i) \log (1 - \hat{y}^i)$$

上式中N为CT影像像素总个数， y^i 为U-net输出的第*i*个像素点的类别， \hat{y}^i 为掩模图像第*i*个像素点的类别(像素值为1表示该像素属于肿瘤区域，像素值为0表示该像素不属于肿瘤区域)

4.1.2 U-net

[8] 本文采用的方法是在FCN基础上改进得到主要针对生物医学图片的分割的U-net网络。相较于FCN，U-Net是完全对称的，且对解码器（即将图像→语义信息的过程看成编码器，语义信息→像素级别的分类的过程看作解码器）进行了加卷积加深处理，而FCN只是单纯的进行了上采样。

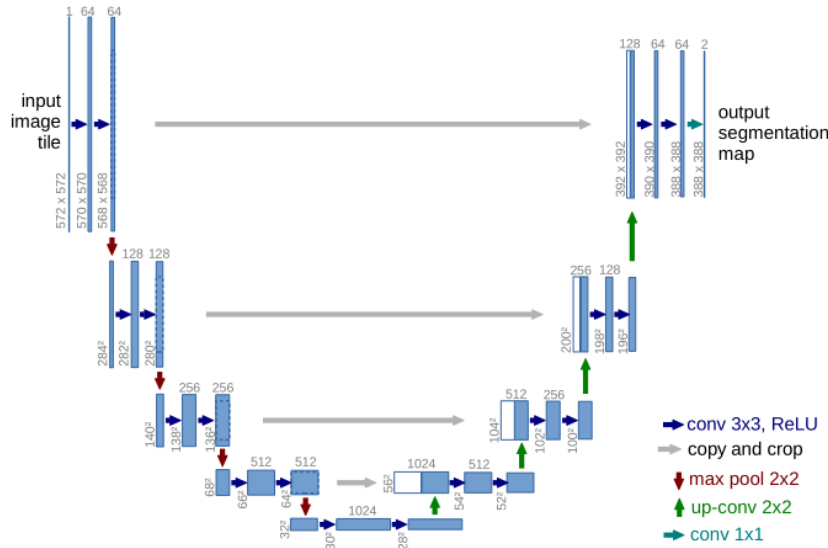


Figure 5: Unet

此外由Fig.5，为了解决特征信息丢失问题U-net提出了一种新的解决方法。不同于FCN种将池化层与feature map结合以减少特征信息的丢失，U-net 则是将卷积过程与反卷积过程中同一层次的feature map进行了结合(feature map 的叠加而不是元素的相加)，从而降低了在卷积过程中低层feature map中的信息随卷积与池化而丢失，同时又由于拼接的两部分来自同一层次，所以两部分的尺寸完全相同，免去了将池化层与feature map上采样到同一尺寸这一过程，网络拓扑结构更加简洁。

4.1.3 实验细节

将预处理过的CT影像文件与掩模文件合并至同一文件，并将其中的80%作为训练集，剩余的20%作为验证集，同时U-net网络超参数设置如Table. 1 所示

Table 1: 超参数	
epoch	30
batch_size	30
dropout	0.5
momentum	0.9
学习率	\exp^{-4}

为避免出现过拟合现象我们在U-net中设置 $dropout = 0.5$ (Srivastava N, Hinton G, Krizhevsky A, 2014), dropout的作用是在神经网络的训练过程中，按一定概率暂时从神经网络中丢弃神经元，对于随机梯度下降(SGD)，由于各个mini-batch丢弃的神经元不同，所以各个mini-batch训练的网络也不同，进而避免了网络的过拟合，但在引入dropout的同时也造成了网络训练时间的增加。

U-net网络loss值随迭代次数的变化如下所示，可以看到在第600个batch 时网络便趋于收敛，为避免过拟合，我们在此时停止训练过程。

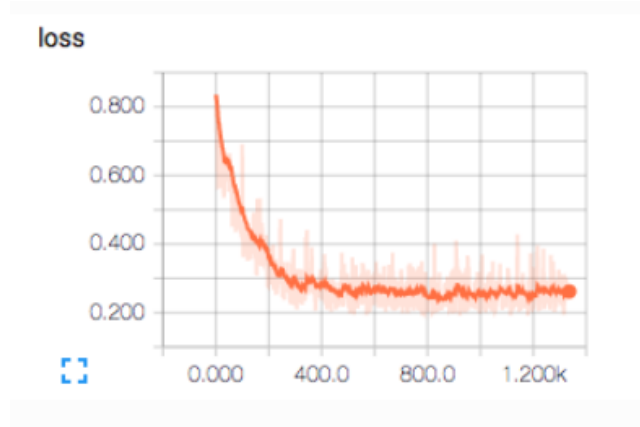


Figure 6: Loss

4.2 肿瘤区域特征提取

针对直肠肿瘤的影像识别特征，我们分别对肿瘤区域的强度信息与大小进

行信息的提取。我们采用影像文件的像素灰度值(0-255) [7] 代表强度,使用掩模区域内的像素点个数CT影像文件内含的Pixel Spacing表示。

4.2.1 强度特征

通过对CT影像文件与相应的掩模合并,我们获取到数据集中的肿瘤区域。区域灰度等级为8,所以我们按照 2^8 个灰度值分别对各个病人的肿瘤区域灰度信息进行了统计分析,由于不同病人的KVP、mAS值并不相同,为了获得同一标准的强度信息,我们将 \tilde{y} 份DCM文件的KVP、mAS值与CT影像的灰度信息进行统一与归一化处理,计算公式如下所示

$$CT_{new}^i = \frac{CT_{old}^i}{N * KVP * mAS}$$

其中 CT_{new}^i 为处理后的CT影像文件的像素值, CT_{old}^i 为原始CT影像文件的灰度值, KVP 、 mAS 分别为该CT影像文件的管电压、管电流,为防止灰度整体过于压缩,我们添加 N 从而调节灰度值变化范围。

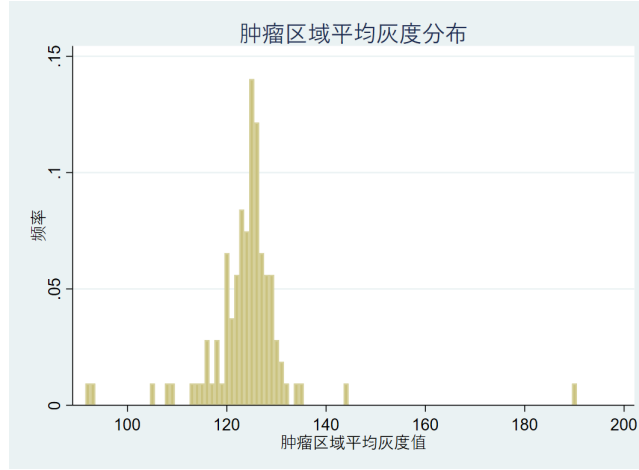


Figure 7: 肿瘤区域灰度值分布

首先我们对肿瘤区域的灰度信息进行了统计,以1灰度值为统计区间我们绘制得到肿瘤区域灰度值分布信息如图.7。粗略观察肿瘤区域的灰度值主要分布在120 ~ 130之间。为了获得肿瘤区域与非肿瘤区域的强度特征区别,我们在裁剪得到的 200×200 CT影像图片中对非肿瘤区域也做了相同的统计计算,统计结果如图.8

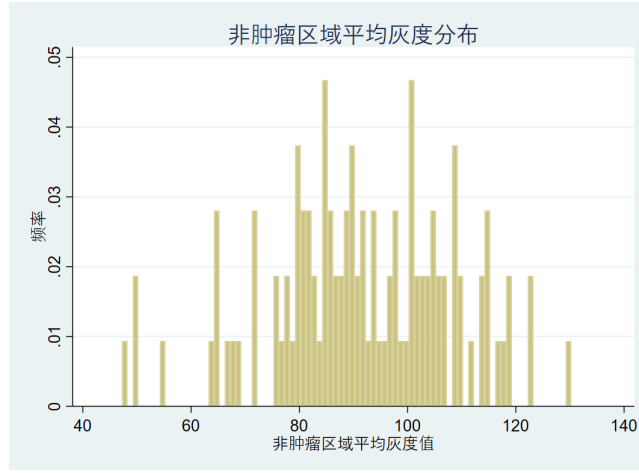


Figure 8: 肿瘤区域灰度值分布

可以看到在直肠周围的非肿瘤区域，灰度主要分布在80 ~ 110之间，从整体上说明了肿瘤区域与非肿瘤区域在组织强度上是存在明显区别的，并且肿瘤区域的灰度值较非肿瘤区域整体偏高，这为我们预测肿瘤区域提供了有力的依据。

4.2.2 面积特征

根据DCM格式要求，CT影像文件中两个像素点间距离为1Pixel Spacing，我们近似认为一个像素点所代表的实际面积为 $PixelSpacin^2$ ，通过对数据集中的掩模文件中肿瘤区域像素点进行计数并按照面积计算公式 $S = Pixel_count * PixelSpacin^2$ 我们近似的得到病人肿瘤区域的面积。由于直肠肿瘤呈圆球状或椭球状，经CT断层扫描的得到的肿瘤影像面积并非均匀变化，根据DCM文件我们得知CT文件扫描切片距离为5mm，直肠肿瘤根据时期从7 ~ 10mm到5 ~ 7cm不等，所以我们使用同一病人不同CT切片肿瘤面积的平均值来代表该病患的肿瘤面积。首先，我们对所有病人的肿瘤区域平均面积做核密度估计，得到107位病患的肿瘤面积分布图如9所示，统计时以10单位面积为同一区间，得到不同面积区间所占百分比。

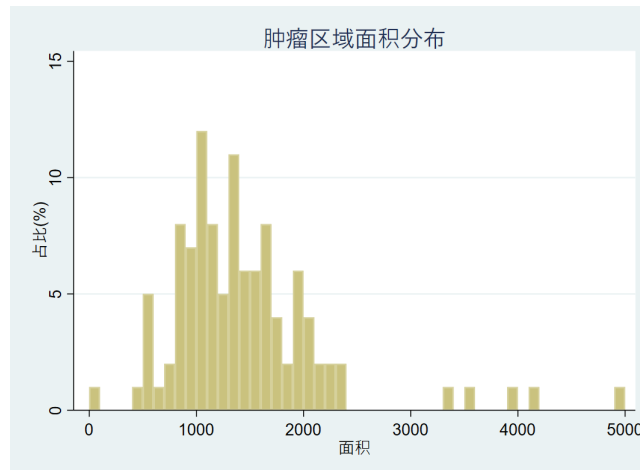


Figure 9: 面积特征

通过以上分析我们得到肿瘤区域与非肿瘤区域的一个区别时在于肿瘤区域的灰度值更高，并且我们大致了解了肿瘤区域的平均面积，所以我们首先对肿瘤区域面积与强度信息之间的关联进行了研究。最后我们想探究下肿瘤区域面积与肿瘤强度之间是否存在关联，基于这种想法，我们绘制了病人肿瘤面积与强度间的散点图Fig.10

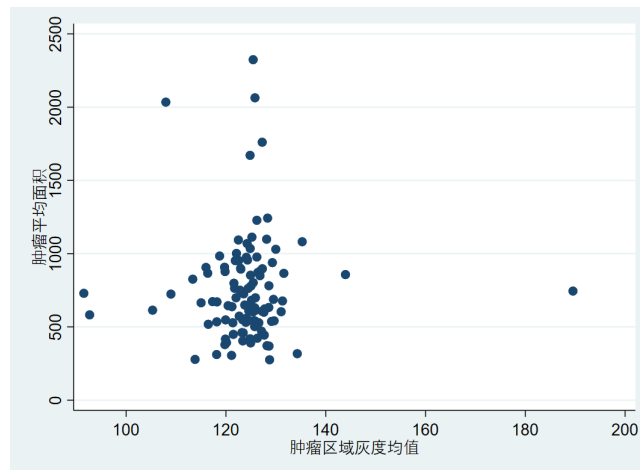


Figure 10: 强度、面积关联图

为了做对比我们同时绘制了非肿瘤区域的面积强度散点图Fig.11

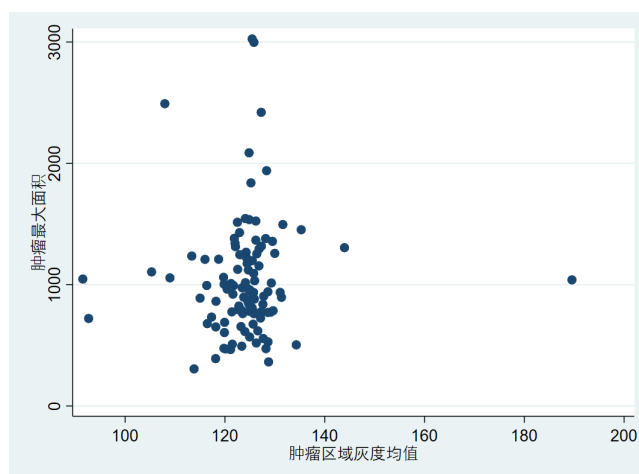


Figure 11: 强度、面积关联图

由Fig.10, Fig.11可以看到这两者间在二维空间并不存在明显的关联(尽管在更高维度可能存在关系),同时可以看到由于病人年龄、性别、体重等因素存在差异,并且肿瘤不同阶段的面积存在着些许差异,除个别数据波动较大外,大部分数据均在一定范围内均匀分布。

4.3 直肠癌淋巴结转移诊断

对于淋巴结转移的诊断,由于不能单一的从CT影像获得确切证据,所以我们采用了两种方式对其进行预测

4.3.1 全连接神经网络

全连接神经网络为最基础的神经网络,通过将数据进行编码作为输入层输入,经过若干隐层,最后得到输出结果,实例图如图Fig.12,通过对自定义的损失函数进行优化,最终可以获得网络隐层的参数,在一些分类问题上具有良好表现。

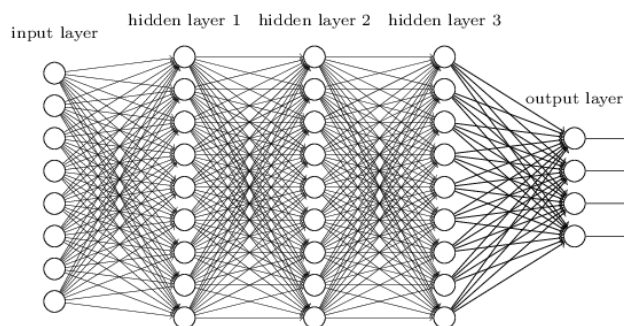


Figure 12: 全连接网络

我们将病人的性别、年龄、平均肿瘤面积、平均强度作为输入，淋巴结转移与否作为输出对数据进行预测。为将数据输入网络，我们做出如下约定。

- 性别使用0、1表示，其中男性使用数值1，女性使用数值0
- 使用0、1表示是否存在淋巴结转移，其中1表示存在淋巴结转移，0表示不存在转移
- \tilde{y} 个病人仅采用包含标注肿瘤的CT图片进行预测

我们通过对mnist手写数字识别网络进行修改得到适用于本问题的全连接网络，但由于数据数量仅有107例，最终的效果并不理想，故我们采用了另一种方式对淋巴结转移进行预测。

4.3.2 k-Nearest Neighbor

k-Nearest Neighbor(KNN) [1]是通过测量不同特征值之间的距离进行分类。它的思想是：如果一个样本在特征空间中的k个最相似(即特征空间中最邻近)的样本中的大多数属于某一个类别，则该样本也属于这个类别，其中K通常是不大于20的整数。KNN算法中，所选择的邻居都是已经正确分类的对象。该方法在定类决策上只依据最邻近的一个或者几个样本的类别来决定待分样本所属的类别。

同样采用上面的数据表示方法，我们在(性别、年龄、肿瘤区域平均面积、强度)四维空间中对107组数据进行表示，但由于无法将四维空间直接映射到二维平面上，在这里我们就不给出可视化结果。

5 模型优化

通过对模型的优化过程中，我们深切体会到对于小规模医疗图像识别的困难所在，由于数据量的缺少和标注的准确性，始终无法得到一个较高的准确率，尽管我们对图片进行了旋转以增强数据的多样性，但仍无法满足模型对数据的需求。在未来的工作中我们认为可以从两个方面对模型进行优化调整

- 增加可利用的数据量，并对数据进行增强处理以便提高准确率
- 通过对CNN网络所获得的feature map 进行可视化表示以获得经由卷积神经网络所学习得到的图像特征，并采用这些特征对之后的任务进行分类。由于现阶段的工作均是采用人工总结出的特征，较之一般仍存在较大程度的主观性，无法保证预测的准确率

6 实验平台

实验环境的软硬件配置如Table.2所示

Table 2: 实验环境配置	
CPU	Intel i7
GPU	GTX 1050
内存	8GB
操作系统	Ubuntu 16.04
Python	3.6
CUDA	Intel i7
cuDNN	Intel i7

用到的Python库有：

- tensorflow
- opencv
- numpy
- matplotlib

7 参考文献

References

- [1] T. M. Cover, P. E. Hart, et al. Nearest neighbor pattern classification. *IEEE transactions on information theory*, 13(1):21–27, 1967.
- [2] Y. LeCun, C. Cortes, and C. Burges. Mnist handwritten digit database. *AT&T Labs [Online]*. Available: <http://yann.lecun.com/exdb/mnist>, 2:18, 2010.
- [3] J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440, 2015.
- [4] T. I. Murphy. Line spacing in latex documents. https://en.wikipedia.org/wiki/Colorectal_cancer.
- [5] T. I. Murphy. Line spacing in latex documents. <https://zh.wikiversity.org/zh-hans/%E7%99%8C%E7%97%87%E8%A8%BA%E6%96%B7r>.
- [6] T. I. Murphy. Line spacing in latex documents. <https://en.wikipedia.org/wiki/DICOM>.
- [7] T. I. Murphy. Line spacing in latex documents. <https://en.wikipedia.org/wiki/Grayscale>.
- [8] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.
- [9] F. Yu and V. Koltun. Multi-scale context aggregation by dilated convolutions. *arXiv preprint arXiv:1511.07122*, 2015.