



计算机工程与应用
Computer Engineering and Applications
ISSN 1002-8331, CN 11-2127/TP

《计算机工程与应用》网络首发论文

题目: 结合多维度特征的病理图像病灶识别方法
作者: 胡伟岸, 邹俊忠, 郭玉成, 张见, 王蓓
网络首发日期: 2020-09-21
引用格式: 胡伟岸, 邹俊忠, 郭玉成, 张见, 王蓓. 结合多维度特征的病理图像病灶识别方法. 计算机工程与应用.
<https://kns.cnki.net/kcms/detail/11.2127.TP.20200921.1437.013.html>



网络首发: 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约, 在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

结合多维度特征的病理图像病灶识别方法

胡伟岸¹, 邹俊忠¹, 郭玉成², 张 见¹, 王 蓓¹

1. 华东理工大学 信息科学与工程学院, 上海 200237

2. 清影医疗科技(深圳)有限公司, 广东 深圳 518083

摘 要:长时间的病理图像人工诊断会使医生产生视觉疲劳, 误诊和漏诊情况容易发生。针对以上现象, 提出一种结合卷积神经网络中多维度特征的方法, 快速准确识别出病理图像中的病灶区域。首先, 使用感兴趣区提取及图像裁剪获得小尺寸图块数据; 其次, 使用染色校正的方法以解决图块染色不均, 对比度弱等问题; 最后, 搭建深度学习模型, 使用多组深度可分离卷积提取不同尺度的特征, 加入残差连接以避免梯度消失, 联合不同维度的特征信息以提高特征利用率。实验结果表明, 染色校正能够提高预测准确率, 上述模型具有参数少, 鲁棒性强的特点, 最终对病理图像病灶的识别均能达到较高的准确率, 假阳性及假阴性均较低, 未来将具有广泛的应用前景。

关键词: 多维度; 深度学习; 卷积神经网络; 深度可分离; 染色校正; 病理图像; 病灶

文献标志码: A **中图分类号:** TP391 **doi:** 10.3778/j.issn.1002-8331.2001-0126

胡伟岸, 邹俊忠, 郭玉成, 等. 结合多维度特征的病理图像病灶识别方法. 计算机工程与应用

HU Weian, ZHOU Junzhong, et al. Lesion Recognition Method of Pathological Images Based on Multidimensional Features. Computer Engineering and Applications

Lesion Recognition Method of Pathological Images Based on Multidimensional Features

HU Weian¹, ZHOU Junzhong¹, GUO Yucheng², ZHANG Jian¹, WANG Bei¹

1. School of Information Science and Engineering, East China University of Science and Technology, Shanghai 200237, China

2. Tsimage Medical Technology, Shenzhen, Guangdong 518083, China

Abstract: Time-consuming artificial diagnosis of pathological images will cause visual fatigue of doctors, while both misdiagnosis and missed diagnosis are easy to occur. In response to the above phenomena, a method combining multidimensional features of convolutional neural network is proposed to quickly and accurately identify lesion in pathological images. Firstly, we use ROI extraction and image cutting to obtain small-scale block data; secondly, the method of stain correction is used to solve the problems of uneven staining and weak contrast in block data; finally, a deep learning model is built, using several depthwise separable convolution to extract features of different dimensions, adding residual connection to avoid gradient disappearance, combining the feature information of different dimensions to improve feature utilization. The experimental results show that stain correction can improve prediction accuracy and the above model has the characteristics of few parameters and strong robustness. At the same time, the accuracy of lesion recognition in pathological images can reach a high level, while both false positive rate and false negative rate are low, so it will have a broad application prospect in the future.

基金项目: 国家自然科学基金(No.61773164); 上海市自然科学基金(No.16ZR1407500)。

作者简介: 胡伟岸(1994-),男,硕士研究生,研究领域为深度学习,医学图像处理,E-mail: 857295087@qq.com;邹俊忠(1960-),男,博士,教授/博士生导师,研究领域为模式识别与智能系统应用;郭玉成(1989-),男,博士,研究领域为生物信息学、深度学习;张见(1976-),男,博士,工程师;王蓓(1976-),女,博士,副研究员。

Key words: multidimensional; deep learning; convolutional neural network; depthwise separable; stain correction; pathological image; lesion

1 引言

目前,通过病理检查对癌症进行诊断是一种常用的方法,它能够提供更明确的疾病诊断,指导病人的治疗。对病理图像进行人工分析本身是一件非常有挑战性的工作,一张病理切片通常包含数百万个细胞^[1],一名病理科医生一天需要分析许多病理图像,这给他们带来很大的工作负担,疲劳阅片现象时有发生^{[2] [3]}。同时,该领域内专家的培养速度赶不上病例的增加速度,将有限的、珍贵的人力资源大量投入到重复的病理图像的识别诊断中是非常可惜的。通过卷积神经网络^[4](CNN, Convolutional neural network)快速识别病理图像中的病变区域是本文的主要研究内容。

CNN 是一种高效的学习方法,局部连接和权值共享的特点降低了网络模型的复杂程度,减少了需要学习的参数。将计算机辅助诊断应用于数字病理图像已经得到了研究:在 Cossatto 等人^[5]的研究中,采用人工提取特征,多实例学习的半监督胃癌检测与诊断方法;Spanhol^[6]等人的研究中,在 BreakHis 乳腺病理数据集上对图像提取了 6 种特征描述,之后使用机器学习方法对图像进行分类,最终整体准确率达到 85%;Ciresan^[7]等人使用有监督的深度神经网络作为分类器,以 88%的查准率及 70%的查全率赢得了 ICPR 2012 有丝分裂检测竞赛;Wang^[8]等人使用 Alexnet 及 GoogLeNet 对临床结直肠病理图像进行图块分类,获得 91%的整体准确率。在 Sharma^[9]等人的研究中,使用了卷积神经网络,对小样本胃癌数据集的三个恶性级别进行了分类,并取得了良好的效果;Li^[10]等人设计了一种基于 CNN 的胃癌自动识别框架,获得了 97%的图块级别准确率;基于多种结构的融合模型在胃癌病理图像上取得了较好的分类结果^[11];Coudray^[12]等人使用基于 Inception 网络的架构,在对肺病理图像中肺组织及肺癌进行区分任务中获得良好的效果;迁移学习的应用在乳腺病理图像的自动分类中获得 91%的图块识别准确率^[13];以上的研究均证明了卷积神经网络在数字病理图像研究中的实用性。然而,在上述研究中仍存在问题,影响了计算机辅助治疗在实际中的应用,包括:手动提取图像特征的过程复杂;

网络模型参数多,训练过程耗时;准确率有待提高。

另外,出于多种原因,如染色原料差异、制备方法差异、图像保存时间长短等,病理图像所呈现的颜色都会有所不同。为使得网络在学习时更加关注细胞的形态、结构和排列等特征,我们期望图像在输入网络前表现出较一致的颜色。使用标准化染色方案和自动染色机可以通过产生更准确和一致的染色来改善染色质量。但是,很难避免所有的外在因素,使最终的染色完全一致^[14]。

针对上述问题,本文在预处理阶段采用了染色校正的策略,在不改变原有形态特征的情况下,使图像表现出更加一致的颜色;受文献^[15]启发,本文提出一种结合多维度特征的卷积神经网络模型以自动提取病理图像特征,将 CNN 中包含纹理等信息的低维特征与包含丰富抽象语义信息的高维特征相结合,使用深度可分离卷积 (Depthwise separable convolution) 降低模型参数量,减小了计算复杂度,并使用批量规范化 (Batchnormalization, BN) 及残差连接 (Residual connection) 避免梯度消失,改善输出结果。实验结果表明:采用染色校正的处理方法能够提高预测准确率;本文所提出的 CNN 模型能够自动提取图像特征,且参数少训练快,最终能够达到较好的识别性能。在胃病理病灶识别任务中准确率达到 98.67%,TCGA 肝癌病理病灶识别任务中准确率达到 97.4%,且假阳性及假阴性均较低,在病灶识别任务中能够媲美主流 CNN 模型所达到的效果。

2 方法

传统的机器学习方法需要以人工提取特征为前提,特征的提取需要设计者具有深厚的专业背景及实际操作经验。然而,使用深度学习的方法能够直接从原始数据中提取特征,LeCun 等人^[4]提出了卷积神经网络,使用卷积核对输入图像进行多层卷积操作,逐层学习并在更高层直接对抽象的特征进行整合,最终通过图像的高层表征对其做出分类。目前优秀的深度学习分类模型有 Vgg16^[16]、

Inception V3^[17]、InceptionResnet V2^[18]、Resnet50^[19]、Densenet169^[20]等。以上模型大多使用最后一个卷积层的 feature map 来进行分类，它们的缺点是往往在关注网络最后一层特征时忽略了其它层的特征。同时，上述模型在分类任务中能够取得较好的结果，通常是以消耗计算量和时间为代价的，这样以计算资源换取效果的策略在实际应用时常会受到局限。

有研究证明，深度学习模型的各层特征图侧重于不同类型信息：由于所经过的卷积网络层较少，模型的低层特征图包含较多边缘、纹理等信息；在经过更多的卷积操作后，高层特征包含更多抽象的语义信息；结合浅层特征可以在一定程度上提升模型精度，将不同特征高效融合，是改善模型性能的关键^[21]。因此，本文选择将多个维度的特征从 top-down 方向上采样为同一尺寸后进行融合。

2.1 模型架构

本文所提出的结合多维特征的病灶识别模型结构如图 1 所示：

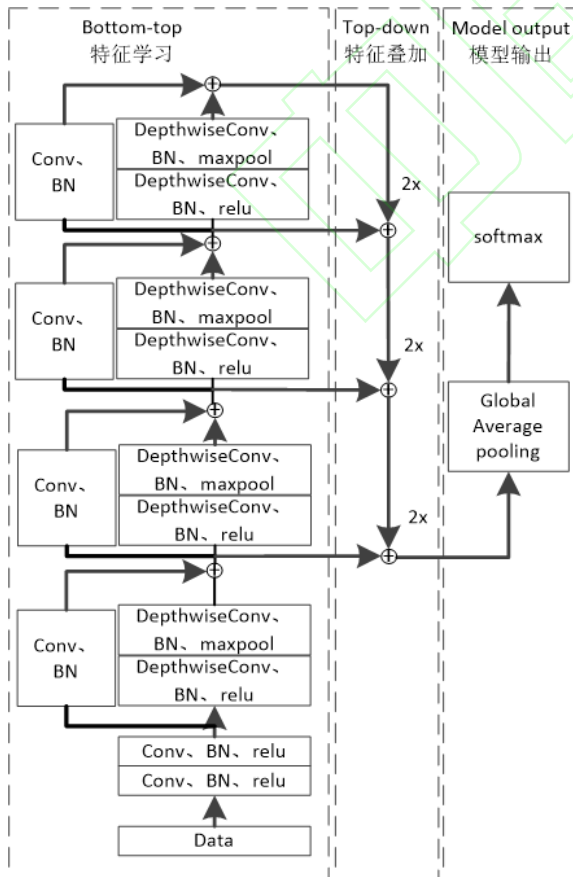


图 1 结合多维特征的病理图像病灶识别模型

图 1 中包含三个模块：Bottom-top 特征学习模块、Top-down 特征叠加模块及 Modeloutput 模型输出模块。

Bottom-top 特征学习模块主要由普通卷积、深度可分离卷积及残差学习组成，通过多层卷积自动提取多维度特征。在普通卷积后依次添加 BN 层及线性整流函数（Rectified Linear Unit, ReLU）。

在对输入图像进行两次普通卷积后，改用深度可分离卷积继续提取特征。在一般卷积中，卷积核对特征图的三个维度进行学习。其中两个维度处于特征图的平面（高和宽），另一个维度处于特征图的通道方向（深度），卷积核能同时学习到特征图的平面相关性和通道相关性。将平面相关性和通道相关性分离开来，是 Inception 模块的重要思想^[13]。在深度可分离卷积中，将卷积操作分为两个部分：首先将所有 2 维的卷积作用在特征图的每个单层通道上，平面相关性和通道相关性即达到了完全分离的效果，此步骤称为 depth-wise 卷积；然后使用 1*1 的卷积核对新特征空间进行映射，此步骤被称为 point-wise 卷积，如图 2a 所示。使用深度可分离卷积能够降低模型参数。假设输入通道数为 N ，长宽均为 D_i ，且输出通道数为 M ，长宽均为 D_o ，卷积核的尺寸设置为 k ，使用普通卷积时参数数量为：

$$M * N * k * k \quad (1)$$

在使用深度可分离卷积过程中，depth-wise 卷积参数数量与 point-wise 卷积参数数量分别为式 (2) (3) 所示：

$$N * k * k \quad (2)$$

$$M * N \quad (3)$$

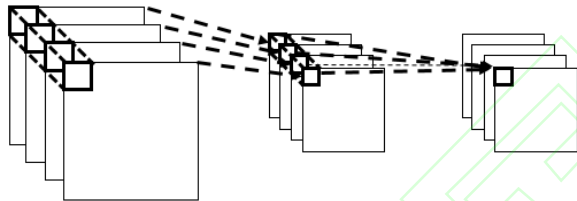
深度可分离卷积与普通卷积的参数数量之比为：

$$\frac{N * k * k + M * N}{M * N * k * k} = \frac{1}{M} + \frac{1}{k * k} \quad (4)$$

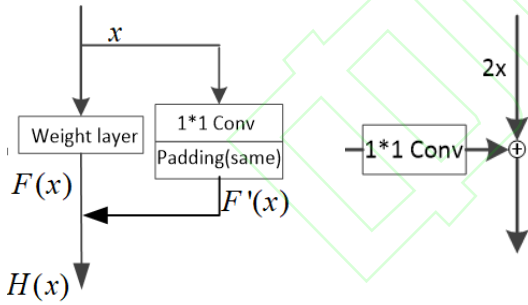
由上式可以看出，两者的比值与输出通道及卷积核大小相关，在本文所使用模型中， M 的取值包括 128、256、512 及 1024，卷积核尺寸 k 均设置为 3。因此，所使用的深度可分离卷积参数数量是普通卷积参数数量的 0.118、0.115、0.113 及 0.112 倍。较少的参数数量节省了模型训练及测试时间。

对于卷积神经网络,如果只是简单地堆叠网络层以增加网络深度,容易导致梯度弥散和梯度爆炸。为避免该问题的发生,在本文网络模型中增加了BN层及Residual connection。文中所使用的Residual connection结构如图2b所示。为保证 $F'(x)$ 与 $F(x)$ 在相加时尺寸相同,使用简单的 1×1 大小卷积层(步长为2)及全零填充(padding=same)对 $F'(x)$ 的通道数进行匹配。在Residual connection中, $F(x)$ 是求和前网络映射, $F'(x)$ 是输入 x 的简单映射,保留了 x 的大部分原有信息,三者关系如下:

$$H(x)=F(x)+F'(x) \quad (5)$$



(a) 深度可分离卷积



(b) 残差学习

(c) 上采样连接

图2 部分模型结构

Top-down 特征叠加模块中,在自顶向下的方向对特征进行上采样,并与左侧相邻的残差输出进行连接,如图2c中的 1×1 卷积被使用以匹配特征图的通道数量,本文中设置卷积核个数 $d=512$ 。之所以选择将每个残差块的输出作为特征结合的因子是因为这些输出具有每个阶段最具代表性的特征,同时,邻近的特征较为相似并且重复使用这些特征会使得计算量成倍增长。由于网络的前两

层卷积所包含浅层特征与原始图像较为相似,且尺寸过大,故不将其作为特征结合的因子。

特征图上采样能够使卷积核细致学习病理图像中的小特征。特征图上采样的处理可以简单理解为增加了特征映射的分辨率,使用相同卷积核在放大后的特征图进行操作时可以获取更多关于小目标的信息。

特征相加的方式采用主元素相加(add)而非特征图堆叠(concat)。add方式更加节省参数和计算量,如果使用concat,特征图通道数会是add方法的两倍,如此计算量会是一笔较大的开销。

在Modeloutput模块中,添加了全局平均池化层(Global Average Pooling, GAP),使用GAP可以对前一层的每一个特征图求平均值,保留全局信息,同时避免了全连接层(Fully Connected layer, FC layer)训练参数过多,耗时较长的缺点。在GAP层后添加2维的输出层,激活函数使用softmax,softmax的计算公式如下:

$$S_i = \frac{e^{V_i}}{\sum_{j=1}^k e^{V_j}} \quad (6)$$

其中 V_i 表示前级第 i 个单元的输出, i 表示类别索引,总的类别个数为 k 。 S_i 表示的是当前元素的指数与所有元素指数和的比值。通过这种方法,可以将神经元的输出映射到 $(0,1)$ 区间内,当成概率来理解。对应于softmax,采用交叉熵(CrossEntropy)作为损失函数,当输出为两类时,交叉熵公式如下:

$$L = -[y \ln y + (1-y) \ln(1-y)] \quad (7)$$

其中 y 代表真实样本标签, y 代表样本标签为1的概率。

表1给出了本文深度神经网络主要层的参

数设定以及输出。

表 1 深度神经网络每层参数以及输出

Layer	Type	Units	Size	Output Size
input	*	*	*	(256,256,3)
Conv1	卷积层	32	(3,3)	(128,128,3)
Conv2	卷积层	64	(3,3)	(128,128,3)
残差块 1	卷积层	128	(1,1)	(64,64,128)
*	深度可分离	128	(3,3)	(128,128,128)
*	深度可分离	128	(3,3)	(64,64,128)
*	Add	*	*	(64,64,128)
残差块 2	卷积层	256	(1,1)	(32,32,256)
*	深度可分离	256	(3,3)	(64,64,256)
*	深度可分离	256	(3,3)	(32,32,256)
*	Add	*	*	(32,32,256)
残差块 3	卷积层	512	(1,1)	(16,16,512)
*	深度可分离	512	(3,3)	(32,32,512)
*	深度可分离	512	(3,3)	(16,16,512)
*	Add	*	*	(16,16,512)
残差块 4	卷积层	1024	(1,1)	(8,8,1024)
*	深度可分离	1024	(3,3)	(16,16,1024)
*	深度可分离	1024	(3,3)	(8,8,1024)
*	Add	*	*	(8,8,1024)
特征结合 1	卷积层	512	(1,1)	(8,8,512)
*	上采样	*	*	(16,16,512)
*	卷积层	512	(1,1)	(16,16,512)
*	Add	*	*	(16,16,,512)
特征结合 2	上采样	*	*	(32,32,512)
*	卷积层	512	(1,1)	(32,32,512)
*	Add	*	*	(32,32,512)
特征结合 3	上采样	*	*	(64,64,512)
*	卷积层	512	(1,1)	(64,64,512)
*	Add	*	*	(64,64,512)
GAP	GAP	*	*	512

softmax softmax * * 2

其中在每层卷积层后分别包含 BN 层，ReLU 及最大池化层。

2.2 整体流程

本文所提出的病理图像病变区域检测的主要流程如图 3 所示：

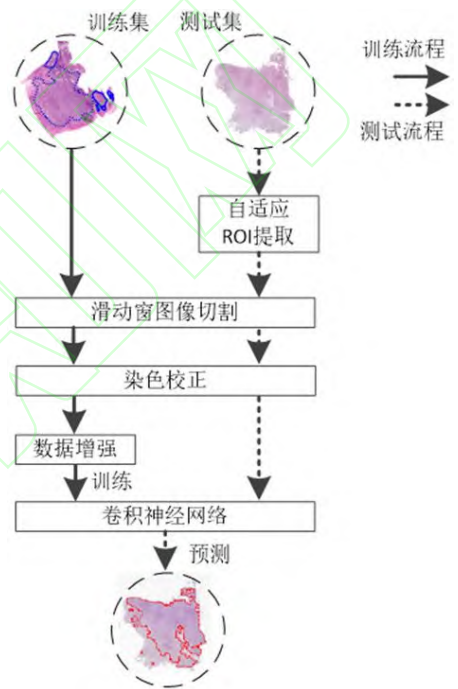


图 3 检测方法流程图

模型训练阶段，计算机接收数字病理图像，在专家标注内进行图像切割，随后对切割出的图块进行染色校正的操作，再对整个图块数据集进行数据增强，以提高模型的泛化能力，最后将图块数据输入所建 CNN 模型进行训练。

在对数字病理图像进行预测的过程中，首先自适应提取感兴趣区(Region of interests, ROI),在 ROI 内进行图像切割；再进行染色校正操作；然后使用 CNN 对各图块进行预测,计算图块属于病变区域的概率；最后根据图块坐标和预测概率在原始数字病理图像中进行标记，辅助医生进行最终诊断。

自适应 ROI 提取：使用掩膜进行分离是提取 ROI 的一种可靠方法。在本研究中，观察到不同病理图像呈现的主体颜色存在差异，因此无法直接使

用固定阈值对所有图像进行前、背景分离。本研究采用大津阈值法^[22] (Otsu) 自适应计算各图的阈值 T ，Otsu 也称最大类间方差法，方差是灰度分布均匀性的一种度量，背景和前景之间的类间方差越大，说明构成图像的两部分的差别越大，当部分前景错分为背景或部分背景错分为前景都会导致两部分差别变小。因此，使类间方差最大的分割意味着错分概率最小。以阈值 T 作为界限将灰度图像分为前景和背景。自适应计算方法公式如下：

$$w_0 + w_1 = 1 \quad (8)$$

$$\mu = w_0 * \mu_0 + w_1 * \mu_1 \quad (9)$$

$$g = w_0 * (\mu_0 - \mu)^2 + w_1 * (\mu_1 - \mu)^2 \quad (10)$$

其中 w_0 、 w_1 分别代表前景和背景像素占整幅图像的比例； μ_0 、 μ_1 、 μ 分别代表前景、背景和整幅图像的平均灰度； g 表示类间方差。将公式 (9) 代入公式 (10)，得到以下公式：

$$g = w_0 * w_1 * (\mu_0 - \mu_1)^2 \quad (11)$$

通过不断调整阈值 T ，使类间方差 g 获得最大值，对图像进行前、背景分离。

自适应提取 ROI 过程：计算原始图像的灰度图；使用大津阈值法获得掩膜；使用形态学闭运算（先膨胀再腐蚀），祛除掩膜中小黑点；剔除面积占整图比例较小区域；获得 ROI。

滑动窗图像切割：在标注及 ROI 内，从上向下依次滑动，保存滑动窗口内的图块作为后续 CNN 的输入数据，滑动窗口的边长与步长相同，为 1024px。如果图块处于标注的病变区域内，则将其标记为阳性，标签为 1；如果图块处于标注的正常区域内，则将其标注为阴性，标签为 0。

数据增强：由于期待卷积神经网络在图块中提取的特征应该是方位不变的，数据增强包括对图块进行旋转、水平或垂直翻转、边缘裁剪和等比例缩放。染色校正会对图像进行色彩上的调整，为保留图像原本的纹理特征，数据增强过程未对图像的亮度和对比度做任何额外的调整。

染色校正：染色校正的目的是降低各图像之间的颜色差异，使得卷积神经网络着重学习病理图像

中细胞的排列和结构，组织的纹理等特征。本研究中采取 Reinhard 颜色迁移方法^[23] 将模板图像的内部颜色特征应用于目标图像。RGB 各通道之间存在着一定的相关性，这意味着，如果要对图像的色彩进行处理，常常需要对像素的三个通道同时进行修改才不会影响图像的真实感，这将大大增加颜色调整过程的复杂性。

因此，在 RGB 色彩空间下进行色彩变化会比较复杂，得到的视觉效果也不自然。本研究中，首先将各图像转移到 Lab 色彩空间中，Lab 色彩空间不仅基本消除了颜色分量之间的强相关性，而且有效地将图像的灰度信息和颜色信息分离开来。所以可以分别对三个通道图像进行独立的运算，而不需要修改另外两个通道的信息，从而不会影响原图像的自然效果。色彩空间转换过程如下：

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.3811 & 0.5783 & 0.0402 \\ 0.1967 & 0.7244 & 0.0782 \\ 0.0241 & 0.1288 & 0.8444 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (12)$$

$$\begin{bmatrix} L' \\ M' \\ S' \end{bmatrix} = \begin{bmatrix} lgL \\ lgM \\ lgS \end{bmatrix} \quad (13)$$

$$\begin{bmatrix} l \\ \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{6}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -2 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} L' \\ M' \\ S' \end{bmatrix} \quad (14)$$

其中 R 、 G 、 B 代表像素点 RGB 通道分量； L 、 M 、 S 代表像素点 LMS 通道分量； l 、 α 、 β 代表像素点 Lab 通道分量。

此方法的主要思想是根据图像的统计分析确定一个线性变换，使得目标图像和模板图像在 Lab 色彩空间中有相同的均值和方差，计算公式如下：

$$\begin{bmatrix} L \\ A \\ B \end{bmatrix} = \begin{bmatrix} \frac{nl'}{nl}(rl - ml) \\ \frac{na'}{na}(ra - ma) \\ \frac{nb'}{nb}(rb - mb) \end{bmatrix} + \begin{bmatrix} ml' \\ ma' \\ mb' \end{bmatrix} \quad (15)$$

其中 rl 、 ra 、 $r\beta$ 分别是源图像 Lab 通道原有的数据， L 、 A 、 B 分别是变换后得到新的源图像 Lab 通道的值， ml 、 ma 、 $m\beta$ 和 ml' 、 ma' 、 $m\beta'$ 分别是源图像和着色图像的三个颜色通道的均值， nl 、 na 、 $n\beta$ 和 nl' 、 na' 、 $n\beta'$ 表示它们的标准方差。在获得 L 、 A 、 B 后，通过公式 (12) - (14) 获得染色校正后新图像的 RGB 通道数值。

经过染色后的图像在不改变原有纹理特征的情况下，使得 RGB 各通道有着更为相似的分布，有着类数据归一化的作用。图 4 所示为染色前后的图像对比图。

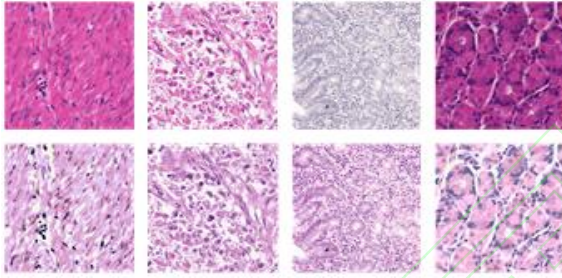


图 4 染色效果图

3 实验结果与分析

实验时，所使用的 CPU 型号为 3.4GHz、16 核 AMD1950x，内存容量为 64GB，GPU 为 12G 显存 NVIDIA TITAN-V。软件方面，使用 Python 作为设计语言，配合 openslide、opencv 等视觉库进行代码编写。深度学习模型框架使用 Keras 以完成对模型的构建，训练与预测。

实验过程中使用了不同数据集对本研究中的网络模型进行了验证，包括胃病理数据集和 TCGA 肝癌公开数据集。为了定量的评估染色效果及不同算法的性能，采用了以下评价指标：分类准确率 Accuracy，敏感性 Sensitivity，特异性 Specificity。

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} * 100\% \quad (16)$$

$$Sensitivity = \frac{TP}{TP + FN} * 100\% \quad (17)$$

$$Specificity = \frac{TN}{TN + FP} * 100\% \quad (18)$$

其中， TP 为真阳性图块数量， TN 为真阴性图块数量， FP 为假阳性图块数量， FN 为假阴性图块数量。

3.1 特征可视化

本次研究中，为探究不同维度特征在病灶识别中的意义，将 Bottom-top 模块中常规卷积层及各残差块中最后深度可分离卷积层的输出特征进行可视化，如图 5 所示。

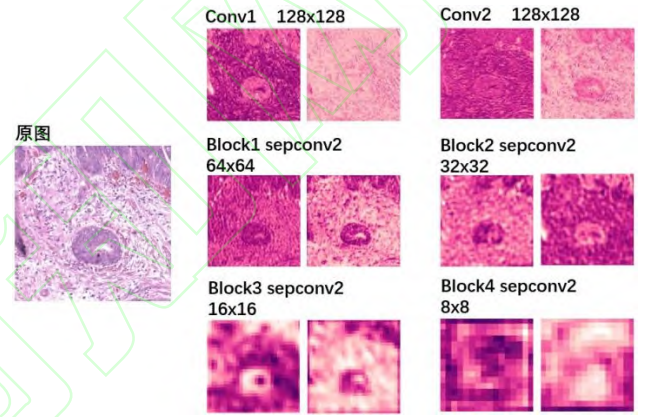


图 5 特征图可视化

图 5 中，左侧为进行染色校正后的病理图块；右侧为对当前图块进行不同卷积后的可视化特征图，每次卷积取两层特征图作为示例。为突出特征图的细节，将其颜色风格设置为紫红色以凸显对比度。可以观察到，在 Conv1，Conv2 中，由于卷积处于浅层位置，特征图中仍能保留原始图像中的较多信息，大量纹理特征被学习；随着网络加深，Block1 中特征图经过下采样，分辨率降低，但与上层特征仍然较为相似，这也是本文网络选择将此处特征图作为特征结合最后一项因子的原因；Block2 中，纹理信息淡化，边缘、轮廓等特征逐渐增强；Block3 中基本无法观察到纹理信息，细胞及组织的边缘、轮廓信息愈加明显；网络不断加深使得特征更加抽象化，Block5 中生成对分类有较强决策意义的语义特征。

病理医生在对图像进行诊断时需要不断地调整视野及关注区域，需对细胞的轮廓及大小，组织边缘的光滑程度，纹理的排列特征等进行观察，最后

整合信息形成判断。从特征可视化图像可以观察到，CNN 对图像进行学习的方式与人工判断的方式有着相似之处，不同维度特征在病灶识别任务中具有实际意义。

3.2 胃病理数据集

本次研究中的胃病理数据集来自于国内某三甲医院,共包含 355 张 tif 格式的全视野数字病理切片(Whole Slide Images, WSIs),所有切片均使用 HE(Hematoxylin-Eosin)染色法进行着色,且以 40X 的倍率进行数字化。专业病理科医生在 ASAP 标注软件中进行标注,所有标注经过另一名医生复核。

以 8:2 的比率将 WSI 分为两部分,其中 284 张作为训练集的原始数据,另外独立的 71 张作为验证集原始数据。经过切割后的图块数据分布如表 2 所示:

表 2 数据分布表		
类别	训练集(张)	验证集(张)
正常区域	73990	7123
病变区域	70970	7258
总量	144960	14318

训练过程中,初始学习率设置为 0.0001,使用 Adam 优化器优化训练,与其他自适应学习率算法相比,Adam 的收敛速度更快,且能够避免损失函数波动大的问题。Batchsize 设置为 64,即每轮随机输入图 64 张。设置迭代次数 epoch 值为 20。刚开始训练的时候使用较大的学习率,能够避免模型陷入到局部最优;在模型的效果趋于平缓后,使用衰减后较小的学习率,使模型能够收敛到最优值。

将图 1 本文模型的 Top-down 叠加模块去除(即 Bottom-top 输出直接连接到 Modeloutput),记为 Modelb。在染色校正后的数据集上与本文模型进行对比,对比效果如表 3 所示:

表 3 特征叠加效果对比			
模型	Accuracy	Sensitivity	Specificity

染色校正后	Modelb	96.80%	96.10%	97.52%
	Ours	98.67%	98.65%	98.69%

由表可知,在添加特征叠加模块后,预测准确率提高了 1.87%,这得益于多维度特征结合时对各层次特征的联合学习以及上采样后对细小特征的精细学习。

另选取多种主流 CNN 模型及其它方法进行实验对比,各模型对独立的验证集图像进行预测的信息如表 4 所示:

表 4 染色校正前后各模型效果对比					
模型	染色	Accuracy	Sensitivity	Specificity	参数(10 ⁶)
Vgg16	×	96.90%	97.26%	96.53%	14.71
	√	98.33%	98.53%	98.13%	
Inception V3	×	97.43%	97.26%	97.60%	21.81
	√	97.58%	97.81%	97.35%	
Incep-tion-Resnet V2	×	95.24%	94.10%	96.39%	54.33
	√	98.66%	98.54%	98.79%	
Resnet50	×	96.59%	96.91%	96.27%	23.51
	√	97.01%	97.26%	96.76%	
Densenet169	×	92.18%	94.43%	89.89%	12.65
	√	95.21%	96.98%	93.41%	
Ref a ^[11]	×	96.21%	96.94%	95.45%	9.98
	√	96.40%	97.41%	95.37%	
Ref b ^[24]	×	96.70%	97.27%	96.11%	25.68
	√	97.10%	97.55%	96.63%	
Ours	×	97.60%	97.91%	97.29%	3.9
	√	98.67%	98.65%	98.69%	

可以发现,在对病理图像的图块进行预测时,各深度学习方法都能够获得较高的准确率,均高于 95%。在同等条件下,本文所建立的模型能够获得更好的预测效果,InceptionResnet V2 所达到的准确率与本文模型相似,但是训练参数量却高达 14 倍之多。表 4 中,使用相同的网络模型且加入染色校正的预处理方法后,各性能指标均有提升,各模型预测准确率最多提高了 3.42%。染色校正能够降低各图像之间的颜色差异,使模型能够着重于学习细胞的形态、结构和排列等特征。在对图块进行预测

前仍进行了染色校正的预处理操作,所以网络查找相关特征不会受到颜色差异的影响。使用染色校正方法后,本文模型在验证集上可以获得最高 98.65% 的灵敏度,同时特异性达到 98.69%。虽然特异性没有达到所有模型中的最高,但是灵敏度作为一个非常重要的指标,能够反映模型检测出病变区域的能力。越高的灵敏度值说明漏诊率的下降,这对于病人早发现病情,接收早期治疗有着重要意义。特异性高说明模型未将正常区域错分为病变区域。本文模型对染色校正后的图块预测准确率最终为 98.67%,这已经达到了相当高的准确率。结合各模型中最高敏感性 98.65%,本次研究中所应用的方法能够为病理科医生提供可靠的诊断基础。

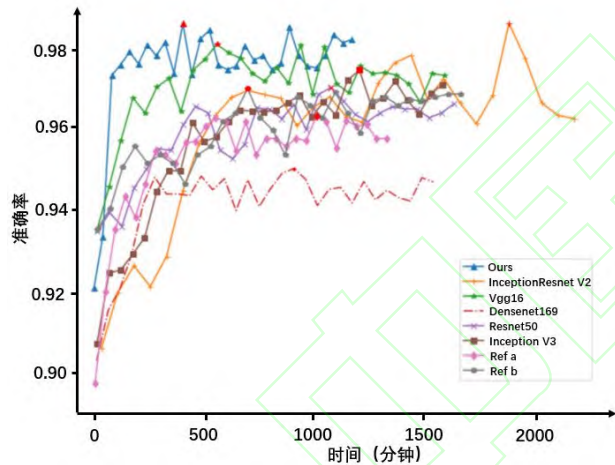


图 6 染色后图像验证集准确率曲线

本次研究中,在完成每轮 (epoch) 训练后使用当次训练所得参数对验证集图像进行测试。图 6 所示即为染色后的图块验证集在各模型上的准确率-时间曲线。为直观体现各模型在耗时上的对比,横轴设置为时间轴,代表在整个过程中获得对应坐标点所经历的训练与验证时间总和。对图 6 观察可知,对于相同规模的图像数据集,本文模型在训练及验证时长上远低于其它模型。本文模型每训练 1 个 epoch 约需 40 分钟,其余模型每训练 1 个 epoch 的时间在 45~73 分钟内。由于每个模型都需要训练多次,因此本文模型能够减少大量训练时间。由图 6 可以发现,本文模型在第 11 个 epoch,约 440 分钟时达到最佳效果,其余模型中最快在 600 分钟时才

能达到最高的准确率。

综上所述,本文模型在病灶识别任务中准确率高,并且较其它方法有更强的病灶检出能力,同时,模型参数少极大节省了训练时间。在图像不断累积,模型参数需要不断更新的现实情况下具有实际应用意义。

3.3 TCGA 肝癌公开数据集

为进一步验证本文模型在不同种类病理上的适用性,对肝癌数据集进行了测试。TCGA (The cancer genome atlas, 癌症基因组图谱) 收录了各种人类癌症的临床数据,选择其中的肝癌病理图像作为实验对象。TCGA 肝癌数据集共包含 379 张 svf 格式的 WSI,所有图像标注由专家提供并复核。选择 302 张 WSI 作为训练集,另外 77 张 WSI 作为验证集。在标注内对图像进行切割后,数据分布如表 5 所示:

表 5 数据分布表

类别	训练集(张)	验证集(张)
正常区域	84428	11245
病变区域	108024	12304
总量	192452	23549

经图像切割后,使用染色校正方法对数据进行处理,将染色后的图块输入网络进行训练和预测。模型参数设置与上文相同。对验证集进行预测后的混淆矩阵如表 6 所示:

表 6 验证集混淆矩阵

		预测标签	
		正常	病变
真实标签	正常	10963	282
	病变	329	11975

由混淆矩阵可以得出,本文模型在 TCGA 肝癌数据集上的准确率可以达到 97.4%,敏感性达到 97.3%。数据证明,本文方法可以在不同的病理图像上获得良好效果。

4 结论

为了提高病理图像病灶识别的准确率及效率,提出了一种结合多维度特征的病理图像病灶识别方法。本文中的模型以深度可分离卷积为基础,降

低了模型的训练参数量;在 Bottom-top 提取特征后,将多维度特征在 Top-down 方向进行上采样并进行叠加,使低维特征与高维语义特征进行关联,特征上采样能够使卷积学习到图像中更精细的特征,叠加特征可以提高特征利用率;最终获得的模型在训练参数量上较主流模型得到大幅降低。同时,本研究在预处理过程中采用染色校正的方法,降低了不同图像的颜色差异。本文在胃病理数据集及 TCGA 肝癌公开数据集上验证了所提模型的性能。实验结果表明,所提出的模型与其他方法相比,具有参数少,训练快,准确率高的优点,能够更准更快的实现病理图像病灶标注。

参考文献:

- [1] 许燕,汤烨,闫雯. 病理人工智能的现状和展望[J]. 中华病理学杂志, 2017, 46(9):593-595
- [2] Webster J D, Dunstan R W. Whole-slide imaging and automated image analysis: considerations and opportunities in the practice of pathology[J]. Veterinary Pathology, 2014, 51(1):211-223.
- [3] Gurcan M N, Boucheron L E, Can A, et al. Histopathological image analysis: a review[J]. IEEE Reviews in Biomedical Engineering, 2009, 2:147-171.
- [4] Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11):2278-2324.
- [5] Cosatto E, Malon C, Kiyuna T, et al. Automated gastric cancer diagnosis on H&E stained sections; ltraining a classifier on a large scale with multiple instance machine learning[J]. Medical Imaging 2013: Digital Pathology.
- [6] Spanhol F A, Oliveira L S, Petitjean C, et al. A dataset for breast cancer histopathological image classification[J]. IEEE Transactions on Biomedical Engineering, 2015, 63(7): 1455-1462.
- [7] Cireşan D C, Giusti A, Gambardella L M, et al. Mitosis detection in breast cancer histology images with deep neural networks[C].International conference on medical image computing and computer-assisted intervention. Berlin, Heidelberg, 2013: 411-418.
- [8] 王莹, 段佳佳, 祝闯等. 临床结直肠病理切片图像的自动辅助诊断[J]. 临床与实验病理学杂志, 2018, 34(10): 18-21.
- [9] Sharma H, Zerbe N, Klempert I, et al. Deep convolutional neural networks for automatic classification of gastric carcinoma using whole slide images in digital histopathology[J]. Computerized Medical Imaging and Graphics, 2017,61:2-13
- [10] Li Y, Li X, Xie X, et al, Deep learning based gastric cancer identification[C]. 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, 2018, pp. 182-185.
- [11] 张泽中,高敬阳,吕纲,赵地. 基于深度学习的胃癌病理图像分类方法[J]. 计算机科学,2018,45(S2):263-268.
- [12] Coudray N, Ocampo P S, Sakellaropoulos T, et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning[J]. Nature medicine, 2018, 24(10): 1559-1567.
- [13] 何雪英,韩忠义,魏本征. 基于深度学习的乳腺癌病理图像自动分类[J]. 计算机工程与应用,2018,54(12):121-125.
- [14] Llewellyn H. Observer Variation, dysplasia grading, and hpv typing a review[J]. American Journal of Clinical Pathology, 2000, 114 Suppl (Suppl 1):S21-35.
- [15] Lin T Y, Dollar P, Girshick R, et al. Feature pyramid networks for object detection[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu,2017: 2117-2125.
- [16] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[C]. International Conferenceon Learning Representations (ICLR), San Diego, 2015.
- [17] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for computer vision[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, 2016: 2818-2826.
- [18] Szegedy C, Ioffe S, Vanhoucke V, et al. Inception-v4, inception-resnet and the impact of residual connections on Learning[C]. 31st AAAI Conference on Artificial Intelligence(AAAI2017), San Francisco, 2017: 4278-4284.
- [19] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, 2016: 770-778
- [20] Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks[C].Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, 2017: 4700-4708.
- [21] Hariharan B, Arbelaez P, Girshick R, et al. Hypercolumns for object segmentation and fine-grained localization[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, 2015: 447-456.
- [22] Ostu N. A threshold selection method from gray-histogram[J]. IEEE Transactions on Systems, Man, and Cybernetics, 2007, 9(1):62-66.
- [23] Reinhard E, Ashikhmin M, Gooch B, et al. Color transfer between Images[J]. IEEE Computer Graphics and Applica-

tions, 2001, 21(5):34-41.

[24] 王恒, 李霞, 刘晓芳, 等. 基于 ResNet50 网络的乳腺癌

病理图像分类研究[J]. 中国计量大学学报, 2019, 30(1): 72-77.

