



桂林电子科技大学
GUILIN UNIVERSITY OF ELECTRONIC TECHNOLOGY

硕 士 学 位 论 文

(全日制专业学位硕士)

题目 基于弱监督学习的组织病理学图像分割算法研究

(英文) Research on histopathology image segmentation

algorithm based on weakly supervised learning

研 究 生 学 号: 20032303054

研 究 生 姓 名: 蓝小斌

指导教师姓名、职称: 张艳菊 研究员

申 请 学 位 类 别: 电子信息硕士

领 域: 软件工程

论 文 答 辩 日 期: 2023 年 5 月 26 日

摘要

区分病理学图像中的不同组织对于病理医生确定病变类型和程度至关重要。全监督语义分割算法通过密集的像素级标签已经实现了出色的性能。然而，在组织病理学图像领域，像素级标签的获取非常困难。为了缓解算法对像素级标签的依赖，一种流行的做法是利用图像级标签训练多标签分类网络来生成类激活图，进而区分不同组织。尽管类激活图能够反映组织在图像中的位置，但该方法有两个不足：其一，随着训练的深入，类激活图会越来越集中在最具有判别性的部分而无法全部覆盖组织；其二，所采用二元交叉熵损失函数独立地计算每种组织的分类损失，无法对类激活图进行监督，导致出现错误激活的问题。针对以上不足，本文以肺腺癌组织分割数据集为对象，研究如何利用图像级标签来更加准确地区分不同组织，并设计了一款集成深度学习模型的组织病理学图像分析与标注软件。主要研究工作如下：

(1) 在训练多标签分类网络过程中，本文结合通道注意力和渐进失活注意力来扩大类激活图并引导网络对重要特征的关注。我们通过对特征图进行全局池化操作得到平均池化特征和最大池化特征，并将两个一维特征输入多层感知机生成通道注意力图来捕获不同通道的重要程度，从而使网络关注重要特征。渐进失活注意力根据训练轮次动态丢弃特征图中最具有判别性的部分，训练越深入，被丢弃的区域就越大，网络就需要探索更多的区域以维持分类预测。

(2) 本文设计并实现了类重新激活映射方法。在多标签分类网络收敛后，我们生成单标签特征，并额外训练一个全连接层，通过 softmax 交叉熵损失函数的类别互斥特性，让网络对错误激活的区域进行惩罚，提高网络分割性能。

(3) 我们利用 PyQt5、OpenCV、Pytorch 等框架开发了一款集成深度学习模型的组织病理学图像分析与标注软件。该软件除了基本图像处理功能以外，还集成了深度学习模型。标注人员可以将模型生成的分割结果转化为可编辑的多边形图元，再通过微调多边形图元来完成数据标注，减少数据标注成本。

本文提出的分割方法在肺腺癌组织病理学图像数据集进行了测试，最终在 MIoU 指标上达到了 76.53%，性能超过了现有方法。同时通过消融实验证明了结合通道注意力和渐进失活注意力以及类重新激活映射的有效性。设计的算法和软件有助于促进自动化分析技术在病理学上的研究与应用。

关键词：弱监督语义分割；组织病理学图像；注意力机制；类激活映射

Abstract

Distinguishing different tissues in histopathological images is crucial for pathologists to determine the type and degree of lesions. The fully-supervised semantic segmentation algorithm has achieved excellent performance through dense pixel-level labels. However, in the field of histopathology images, it is very difficult to obtain pixel-level labels. To alleviate the dependence of algorithms on pixel-level labels, a popular approach is to use image-level labels to train multi-label classification networks to generate class activation maps and distinguish different tissues. Although class activation maps can reflect the position of tissues in the image, this method has two shortcomings: firstly, as the training deepens, the activated regions will become increasingly concentrated in the most discriminative parts and cannot fully cover the tissues; Secondly, the binary cross-entropy is used to calculate the classification loss of each kind of tissue independently, and different tissue may be activated at the same pixel, resulting in segmentation errors. Given the above shortcomings, this thesis takes the LUAD-HistoSeg dataset as the object to study how to use the image-level label to distinguish different tissues accurately and develops a histopathology image analysis and annotation software. The main research work is as follows:

(1) In the process of training multi-label classification networks, this thesis combines channel attention and progressive dropout attention to expand the activation region. Progressive dropout attention dynamically drop the most discriminative regions in the feature map based on the training epochs. The deeper the training, the larger the discarded regions and the network needs to explore more regions to maintain classification prediction. At the same time, we add channel attention before channel attention, obtain the average-pooled feature and max-pooled feature through global pooling operation on the feature map, input two one-dimensional features into the multi-layer perceptron to generate channel attention map to capture the importance of different channels, and finally multiply the channel attention map and the input feature map channel by channel to guide the network to pay attention to important features.

(2) This thesis designs and implements a class reactivation mapping method. After the multi-label classification network converges, we generate a single-label feature and train an additional fully connection layer. Through the class mutual exclusion feature of softmax cross-entropy loss function, the network punishes the wrongly activated regions and improves the network segmentation performance.

(3) Using PyQt5, OpenCV, Pytorch and other frameworks, we have developed a histopathology image analysis and annotation software that integrates deep learning models. In addition to basic image processing functions, the software also integrates deep learning models. Annotators can convert the segmentation results generated by the model into editable polygon primitives, and then complete data annotation by fine-tuning the polygon primitives to reduce data annotation costs.

The segmentation method proposed in this thesis has been tested in the LUAD-HistoSeg dataset, and reached 76.53% MIoU, which outperformed the existing methods. At the same time, the effectiveness of combining channel attention with progressive dropout attention and class reactivation mapping was demonstrated through ablation experiments. The designed algorithms and software help promote the automation of medical image analysis technology in pathological research and application.

Keywords: weakly supervised semantic segmentation; histopathology image; attention mechanism; class activation mapping

目 录

第一章 绪论.....	1
§1.1 研究背景与意义.....	1
§1.2 国内外研究现状.....	3
§1.3 本文研究内容、方法以及创新点	5
§1.3.1 研究内容.....	5
§1.3.2 研究方法.....	5
§1.3.3 创新点.....	6
§1.4 本文组织结构.....	6
第二章 相关理论知识	8
§2.1 卷积神经网络.....	8
§2.1.1 卷积层.....	8
§2.1.2 池化层.....	9
§2.1.3 全连接层.....	9
§2.1.4 激活函数.....	10
§2.1.5 类激活映射	10
§2.2 残差网络.....	11
§2.3 视觉注意力机制.....	12
§2.3.1 通道注意力.....	12
§2.3.2 空间注意力.....	12
§2.4 网络训练以及性能评估	13
§2.4.1 损失函数介绍.....	13
§2.4.2 分割算法性能评估指标.....	14
§2.5 本章小结.....	15
第三章 基于弱监督学习的肺腺癌组织分割算法	16
§3.1 弱监督语义分割模型的构建	16
§3.1.1 特征提取.....	17
§3.1.2 扩大类激活图.....	17
§3.1.3 类重新激活映射.....	18
§3.2 数据集以及实验设置.....	19
§3.2.1 数据集介绍.....	19
§3.2.2 实验参数设置.....	20
§3.3 实验结果分析.....	20
§3.3.1 消融实验结果与分析.....	20
§3.3.2 对比实验结果与分析.....	22

§3.3.3 局限性分析	23
§3.4 本章小结	24
第四章 组织病理学图像分析与标注软件	25
§4.1 软件需求分析	25
§4.2 软件概要设计	27
§4.2.1 软件层次结构设计	27
§4.2.2 软件开发技术介绍	27
§4.3 软件详细设计	28
§4.3.1 软件界面设计	28
§4.3.2 图像分割功能	29
§4.3.3 图像几何变换功能	29
§4.3.4 颜色处理功能	30
§4.3.5 图像标注功能	31
§4.4 软件功能测试	31
§4.4.1 图像分割功能功能测试	31
§4.4.2 图像标注功能测试	32
§4.4.3 图像几何变换功能测试	32
§4.4.4 颜色处理功能测试	32
§4.5 本章小结	33
第五章 总结与展望	34
§5.1 本文总结	34
§5.2 研究展望	34
参考文献	36

第一章 绪论

§ 1.1 研究背景与意义

癌症是一种致死率极高的疾病,已经成为严重威胁人类健康的公共卫生问题之一^[1,2]。随着人们寿命的延长,癌症的影响还在继续扩大,根据国际癌症研究机构(IARC)在2020年发布的最新数据^[3],全球每年新诊断的癌症病例已经超过1900万,每年死于癌症的人数也超过1000万。随着诊断技术的发展,如果癌症能够早期发现和准确诊断,患者就更有可能得到有效治疗,提高存活率。

数字病理图像能够为癌症的筛查和诊断提供可靠依据。通常情况下,病理医生对患者的组织、细胞以及体液进行取样,然后使用苏木精-伊红染色剂染色制作成病理切片,最后利用扫描仪对病理切片进行扫描生成超高分辨率的数字图像。根据样本类型的不同,数字病理图像可以进一步细分成细胞病理学图像和组织病理学图像等。与细胞病理学图像不同,除了显示细胞形态以外,组织病理学图像还能反映病变组织的结构以及其与周边组织的关系。因此,为了确定患者的病情,病理医生通常需要对组织病理学图像中的各种组织进行分析,并结合自身从业经验做出诊断结论。无论是组织取样还是图像分析,都需要病理医生的深度参与,而根据发表在《柳叶刀》杂志的调查结果^[4-6]可知,由于病理学人才培养周期长、工作风险大等原因,具有执业资格的在册病理医生数量不足,无法满足患者的诊断需求。为了缓解这一问题,除了加强病理学教育的投入以外,探索如何实现组织病理学图像自动化分析也是一个紧迫且重要的方向,它有助于提高诊断效率,减轻病理医生的工作负担。

组织病理学图像自动化分析的一个关键是利用计算机视觉技术区分不同组织,这是一个语义分割任务。正如图1-1所示,(a)表示源图像,(b)展示了算法的分割结果,(c)则展示了语义分割任务的本质,即图像中每个像素都被算法准确分类,从而达到区分不同组织的目的。

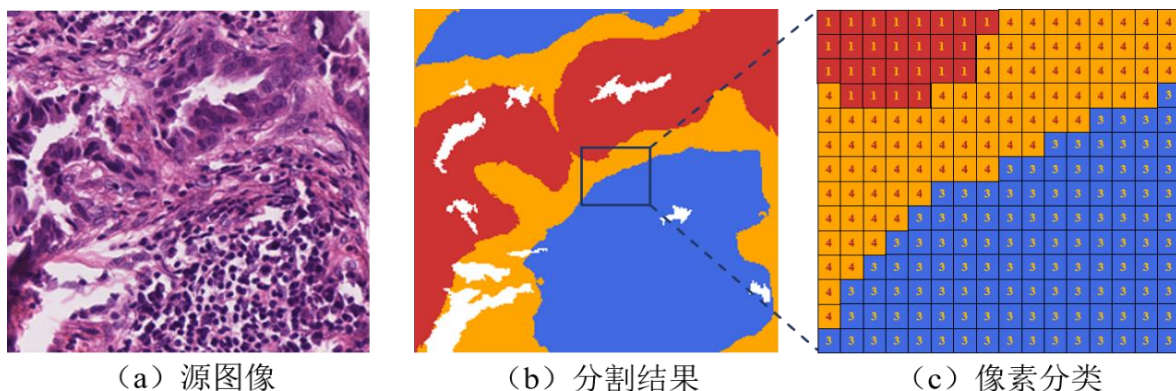


图 1-1 语义分割

近些年来，随着硬件设备的快速发展，硬件计算能力大幅提高，深度学习技术在语义分割任务上展现出强大的优势^[7,8]，其原因在于深度学习技术可以从数据中学习高层次和抽象的特征表示，并且能够利用大量标注数据进行端到端的训练和优化。全监督学习是语义分割任务中最常见的深度学习方法，在各个领域都有成功的应用，例如医学图像领域、遥感图像领域和自动驾驶领域等，常用的网络架构有 UNet 和 DeepLab 两种。UNet 以及衍生网络^[9-11]，由一个编码器和一个解码器组成，形成了一个 U 形结构。编码器负责提取高级别特征，解码器负责恢复细节信息，并且使用跳跃连接将对应层级的特征进行融合，其在医学图像分割任务中表现出色；DeepLab 系列网络^[12-14]是一系列不断改进和发展的全监督分割算法，采用空洞卷积来增加感受野，并使用空洞空间金字塔池化提取多尺度特征，在遥感图像分割和自动驾驶领域中得到了实际应用。这些分割算法之所以能取得很好的效果，除了优秀的网络结构设计之外，还有一个重要原因就是它们采用像素级标签作为监督信息，能够为网络提供准确的位置和精细的边界信息。然而，像素级标签的制作需要标注人员利用标注工具拖动鼠标描绘每一个对象的准确边界，这需要大量的标注时间。更重要的是，对于组织病理学图像而言，标注工作只能依赖于有相关知识背景的病理医生，由于病理医生的稀缺以及繁重的工作量，组织病理学图像分割领域缺乏大量、高质量标注的图像，进而阻碍了全监督学习算法在组织病理学图像分析上的应用。

为了减少对像素级标签的依赖，弱监督学习算法孕育而生，其算法致力于利用弱标签来完成不同组织的分割。如图 1-2 所示，根据弱标签形式的不同，可以把弱监督语义分割算法分为包围框、点、涂鸦以及图像级标签四种^[15-18]。其中，图像级标签仅需标注人员标注类别是否存在即可，1 表示存在，0 表示不存在，其标注成本最低，应用价值最大，是推动智能医疗发展的研究热点。

因此，本文重点研究如何利用图像级标签来准确区分图像中的不同组织。算法只需要根据图像中存在组织的种类，就能够分割图像中的不同组织，极大程度减少了对标注精度的依赖，具有更广泛的实际应用价值，有助于推动医学图像的自动化分析的进展和应用。

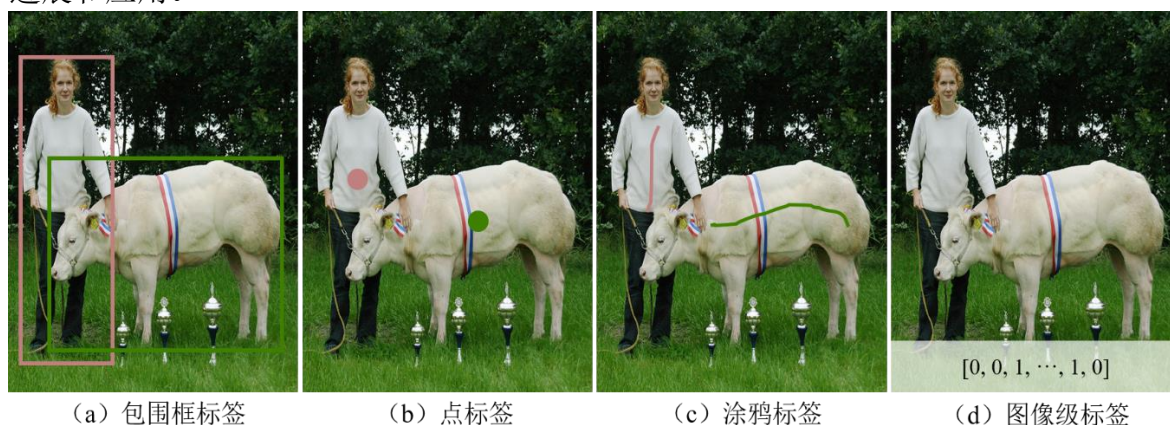


图 1-2 常见的弱标签形式（示例来自于 PASCAL VOC 数据集）

§ 1.2 国内外研究现状

近年来,随着人工智能和数字病理成像技术的发展,组织病理学图像自动化分析具有广阔的发展前景,引起了国内外科研人员的高度重视^[19-21]。

深度学习技术在组织病理学图像自动化分割领域取得了许多成就。Chen 等人^[22]提出了深度轮廓感知网络,该网络包括两个主要部分:全卷积网络(Fully Convolution Network, FCN)和轮廓感知网络(Contour-Aware Network, CAN)。全卷积网络用于提取图像特征,并产生初步的分割结果,轮廓感知网络则专注于分割结果的轮廓信息,以提高分割的准确性。该方法不仅可以准确地分割出腺体区域,而且还可以清晰地描绘出轮廓,但该方法需要为每个层次设计合适的边界权重参数,这些参数可能需要根据不同数据集进行调整,缺乏泛化性。Qu 等人^[23]提出全分辨率卷积神经网络(Full Resolution Neural Network),用于分割结肠腺癌,该方法通过保持全分辨率的特征图来提高边界定位的准确性,并且提出了一种方差约束交叉熵损失,指导网络学习像素之间的空间关系,但由于保持全分辨率特征图,所以该方法需要耗费大量的计算资源。Tokunaga 等人^[24]提出了自适应加权多视场卷积神经网络(Adaptive Weighting Multi-Field-Of-View CNN),该网络由多个分支组成,每个分支负责提取不同视场下的图像特征,并根据特征质量和相关性自适应地进行加权融合。同时,该研究还提出了均衡交叉熵损失函数(Balanced Cross-Entropy Loss),用于避免训练过程中类别不平衡问题对模型性能的影响。可以接收不同分辨率的组织病理学图像并识别更多的癌症亚型。

这些分割方法虽然取得较好的分割效果,但却需要精细的像素级标签,为了减少分割任务对像素级标签的依赖,研究人员提出了许多弱监督分割算法,可以将其分为多示例学习(Multiple Instance Learning)和类激活映射(Class Activation Mapping)两个方向。

多示例学习是一种弱监督学习方法,它假设训练数据是由一些数据包组成的,每个数据包里有多个样本。对于正类数据包来说,至少有一个样本是正类,对于负类数据包来说,则所有样本都是负类。在图像分割领域,多示例学习可以用来解决只有图像级别的分割任务。例如,在语义分割任务中,可以将一张图像看作一个数据包,图像级别标签看作数据包的标签。Xu 等人^[25]将聚类引入多示例学习中,充分利用图像级标签的监督信息,完成高精度的分类和分割任务,该方法还通过引入上下文约束作为先验知识,增加了网络对不同尺度和形态特征的感知能力,并且减少了多个示例之间相互干扰或重叠的可能性,在结肠癌数据集中表现优异。Jia 等人^[26]提出的 DWS-MIL 方法通过结合弱监督学习和约束优化技术,实现对组织病理学图像自动化分割。该方法利用弱监督学习技术从训练数据中自动标记区域,并引用条件随机场模型,提高分割算法的准确性和鲁棒性。Lin 等人^[27]提出的 ScanNet 使用卷积神经网络和多任

务学习技术来提取特征,并能够在组织病理学图像中快速地检测出乳腺癌转移,除此之外,该方法还提出多级阈值分类方法,对检测结果进行筛选和优化。多示例学习算法通常忽略包内或包间的关系,导致无法充分利用数据中潜在的结构和语义信息。除此之外,多示例学习需要不断迭代优化,交替地更新样本标签和分类模型,这可能导致网络收敛速度慢或陷入局部最优。

类激活映射是利用卷积神经网络定位目标对象的一种方法,它通过将全连接层的权重与特征图相乘得到类激活图,类激活图能够反映目标对象在图像中的大致位置,是利用图像级标签来完成分割任务的主要方法。由于组织病理学图像中组织的分布较为分散,且组织结构复杂多变,区分组织病理学图像中的不同组织是一项极具挑战性的视觉任务。目前,利用图像级标签实现组织病理学图像分割的研究较少,Chan 等人^[34]提出了一个以 VGG16^[35]为主干网络的弱监督分割方法。该方法利用卷积神经网络和梯度加权类激活映射来自动地为图像中的不同组织类型生成类激活图,然后利用一系列后处理步骤修正类激活图,并得到最精确和完整的分割结果,最终完成了超过 10 种组织类别的分割。FGNet^[36]是一个联合全卷积神经网络和图卷积神经网络的端到端分割流程,利用前景比例和其置信区间来指导模型学习,该方法使用全卷积神经网络提取图像特征,然后通过动态超像素技术划分特征图,并将结果输入到图卷积神经网络进行聚类。其方法的有效性在 HER2, KI67 以及 H&E 三个数据集中得到验证^[37-39],由于该方法将前景与背景比例作为监督信息来训练网络,因此,该方法只能用于区分前景与背景。Han 等人^[40]提出了多层伪标签监督训练的分割流程,其主要有两步:在第一步,利用图像级标签训练一个多标签分类网络,待多标签分类网络收敛后利用梯度类激活映射生成不同卷积层的类激活图作为伪标签,受 Srivastava 等人的启发^[41],Han 等人^[40]提出渐进失活注意力(Progressive Dropout Attention, PDA)来丢弃判别性特征,迫使网络探索更多区域来维持分类预测。在第二步中,将第一步生成的多层伪标签作为监督信息来训练全监督分割网络,提高模型的分割性能,最终在肺腺癌组织和乳腺癌组织分割数据集中表现最优。

综上,虽然先前的研究能够成功地利用图像级标签实现不同组织的分割,但仍然存在不足之处。基于多示例学习的方法无法充分利用图像中潜在的结构信息,且在训练过程中需要迭代更新样本标签和优化模型,导致模型性能有限。基于类激活映射的方法通过卷积操作学习目标对象更多的结构信息,生成的类激活图能够反映目标对象的大致位置,但其无法全部覆盖目标对象。丢弃判别性特征是扩大类激活图的有效方法,但该方法仅作用在空间维度上,缺少对通道重要性的关注。此外,基于类激活映射的弱监督语义分割方法通常采用二元交叉熵损失函数来独立地计算分类损失,该损失函数无法对类激活图进行监督,导致出现错误激活区域。

§ 1.3 本文研究内容、方法以及创新点

通过对国内外研究现状和发展趋势的分析,针对当前技术存在的不足,本文提出了一个利用图像级标签来准确区分不同肺腺癌组织的算法,并在此基础上开发了一款集成深度学习模型的组织病理学图像分析与标注软件,进一步推动组织病理学图像研究,不仅有技术创新也有实际的应用价值。

§ 1.3.1 研究内容

本文采取类激活映射方法作为基础方案,并对如何生成更高质量的分割结果展开深入研究,其主要研究内容包括:

(1) 扩大类激活图范围,使其全面覆盖目标对象。如何引导网络生成更大范围的类激活图是本文的研究重点之一。

(2) 减少无关类别的激活。基于类激活映射方法通常需要二元交叉熵损失函数来计算分类损失,无法对位置信息进行监督。如何在仅有图像级标签的监督信息条件下对错误激活区域进行惩罚,优化类激活图,也是本文研究内容之一。

(3) 提高组织病理学图像的标注效率。相对于其他场景的图像数据而言,组织病理学图像的标注工作依赖于专业的病理医生,本文将研究如何利用深度学习模型来为病理医生提供客观参考,并减少标注成本。

§ 1.3.2 研究方法

如何设计算法来引导网络生成更准确的类激活图是利用图像级标签区分不同组织的关键问题。其主要研究方法如下:

(1) 数据预处理。本文以肺腺癌组织为对象展开研究。在训练过程中,输入图像以 50% 的概率水平、垂直翻转,以增强数据的多样性。

(2) 设计弱监督语义分割算法。本文研究如何利用图像级标签来区分不同组织,其方法主要包括三个部分:第一,特征提取。我们将 WiderResNet38^[42]作为主干网络来提取特征;第二,生成类激活图。在这一部分,我们利用通道注意力捕捉不同通道之间的重要程度,引导网络关注重要特征。同时,我们通过丢弃判别性特征,引导网络探索更多目标区域;第三,类重新激活映射。待网络收敛后,我们额外训练一个全连接层,通过利用 softmax 交叉熵损失函数的类别互斥特性,在仅有图像级标签的监督信息条件下,对错误激活的区域进行惩罚,优化类激活图。

(3) 算法分析与评估。本文通过计算语义分割任务中常用的指标 MIoU 来评估

算法的有效性，并通过比较分割结果和类激活图来体现本文方法的优越性。

§ 1.3.3 创新点

本文的主要目的是利用图像级标签更加准确地区分不同组织，并通过已收敛模型提高标注工作的效率。主要创新如下：

(1) 本文提出了一种结合通道注意力和渐进失活注意力的方法在扩大类激活图范围的同时让网络保持对重要特征的关注。我们通过通道注意力为不同通道分配权重，以提高重要特征的激活程度。同时，我们根据训练轮次来丢弃特征图中最具有判别性的部分，驱使网络探索更多目标区域。

(2) 本文实现了一个类重新激活映射方法，在仅有图像级标签的监督信息条件下对错误激活的区域进行惩罚，优化类激活图。具体做法是：待网络收敛后，我们额外训练一个全连接层，通过 softmax 交叉熵损失函数的类别互斥特性来减少无关类别的激活，重新生成类激活图。

(3) 为了减少病理医生的工作负荷，我们开发了一款集成深度学习模型的组织病理学图像分析和标注软件。该软件为用户提供便捷的几何变换、颜色变换、图像分割以及标注等功能，能帮助医生快速准确地分析组织病理学图像，并通过模型生成的初步标注结果来提高病理医生的标注效率。

§ 1.4 本文组织结构

第一章：绪论。在这一章中，我们介绍了组织病理学图像自动化分割对于提高病理诊断效率、减轻病理医生工作负担的重要意义，还介绍了弱监督分割方法在组织病理学图像分割的应用价值，然后我们总结并概括了基于图像级标签的弱监督语义分割方法在组织病理学图像分割方法的国内外研究现状。最后我们提出本文的研究内容以及创新点。

第二章：相关理论知识。简要介绍了与本文相关的理论知识，包括卷积神经网络相关技术、残差网络、视觉注意力机制以及网络训练和性能评估指标。其中，在卷积神经网络这一小节中，我们重点介绍了 sigmoid 与 softmax 两种激活函数在计算上的区别，同时，我们还重点介绍了类激活映射这一卷积神经网络可视化技术，这是许多弱监督语义分割方法的基础。在视觉注意力机制这一小节，我们主要介绍了通道注意力和空间注意力机制，这两种注意力分别从通道和空间两个角度来引导网络训练。最后，在网络训练以及性能评估这一小节，我们介绍了二元交叉熵损失和 softmax 交叉熵损失函数，并讨论了它们对本文方法的启示作用。此外，我们还介绍了分割算法中常用的性能评估指标。

第三章：基于弱监督学习的肺腺癌组织分割算法。简要介绍了肺腺癌病理诊断对于临床诊断以及治疗方案的意义，以及着重介绍了如何利用图像级标签来实现肺腺癌组织的分割。为了提高网络性能，我们还结合了通道注意力和渐进失活注意力的方法来扩大类激活图范围，并引导网络对重要特征的关注。在网络收敛后，我们额外训练一个全连接层，通过 softmax 交叉熵损失函数的类别互斥特性对错误激活的区域进行惩罚，优化类激活图质量。最后我们通过消融实验和对比实验验证了本文方法的有效性。

第四章：组织病理学图像分析与标注软件。主要介绍组织病理学图像分析与标注软件的层次结构设计、界面设计、功能设计、功能实现以及功能测试等。

第五章：总结与展望。对实验结果以及结论进行分析和总结，并基于现有方法的不足明确未来的研究方向。

第二章 相关理论知识

语义分割是计算机三大视觉任务之一。语义分割算法的结果能够提供图像中目标对象的类别、位置以及精细的边界信息，因此语义分割对于机器理解图像具有非凡的意义，是研究组织病理学图像自动化分析的重要组成部分。随着硬件更新换代，硬件计算速度大幅度提高，计算机已经能够承担大规模的计算任务，这为深度学习算法在视觉任务中的研究和应用创造了条件。为了方便读者对后续章节的理解，本章介绍相关理论知识。2.1 节介绍卷积神经网络的相关知识，2.2 节介绍残差网络结构，2.3 节介绍视觉任务中的注意力机制，2.4 节介绍网络训练以及性能评估的知识。

§ 2.1 卷积神经网络

卷积神经网络（Convolutional Neural Networks, CNNs）是一种广泛应用于图像和语音等信号处理的深度神经网络。卷积神经网络的基本结构包含卷积层、池化层和全连接层等。其中，卷积层是卷积神经网络的核心，用于提取特征。池化层的作用在于下采样，降低特征维度。全连接层则是将网络汇总的特征输入到分类器进行分类。由这些基本组件组成的卷积神经网络具有以下几个特点：

（1）局部连接：卷积神经网络通过使用卷积层来检测输入的结构化特征，卷积层的每个神经元只与输入张量的一小块相连，这使得卷积神经网络能够有效地捕捉输入中的局部模式。

（2）参数共享：卷积神经网络中的卷积核在整个输入中共享参数，这减少了网络的参数量，提高了训练效率和泛化能力。参数共享还使得卷积神经网络能够适应输入的平移不变性。

（3）多层次抽象：卷积神经网络通过多层卷积层和池化层来逐层提取输入张量的特征，并逐渐合成更好层次的抽象特征，这种逐层抽象的过程使得卷积神经网络能够适应输入中多个层次的抽象特征。

§ 2.1.1 卷积层

卷积层是卷积神经网络中最基本的组件之一，其主要作用是从输入数据中提取出特征。在卷积层中，使用卷积核（也称为滤波器）对输入数据进行卷积操作，将其转换为特征图。卷积操作的本质是一种滑动窗口的运算，将卷积核在输入数据上进行滑动，并将每个位置上的数据与卷积核对应位置的权重进行乘积并求和，得到一个输出

值。卷积层的优点在于它可以自动学习输入数据的特征表示，减少了手工提取特征的需求，同时也具有参数共享的特点，减少了模型的参数量，提高了模型的泛化能力。

§ 2.1.2 池化层

池化层也称为下采样层，它的主要作用是对特征图进行降维和压缩，减少模型的参数量和计算量。与卷积层类似，池化层通过在特征图上移动滑动窗口并对每个窗口内的数据进行计算来实现下采样，这是针对特征图空间维度的操作。根据计算的不同，池化操作可以分为最大池化和平均池化两种。其中，最大池化是将滑动窗口内最大值作为输出，而平均池化则是将滑动窗口内的平均值作为输出。如图 2-1 所示，如果把滑动窗口的尺寸与特征图尺寸相同，那么池化操作也被称为全局池化操作。除了对特征图空间维度进行池化以外，还可以沿着特征图通道维度进行池化操作，这种池化被称为跨通道池化操作。跨通道池化可以将每个通道的特征压缩成一个标量，考虑整个特征图通道的信息。

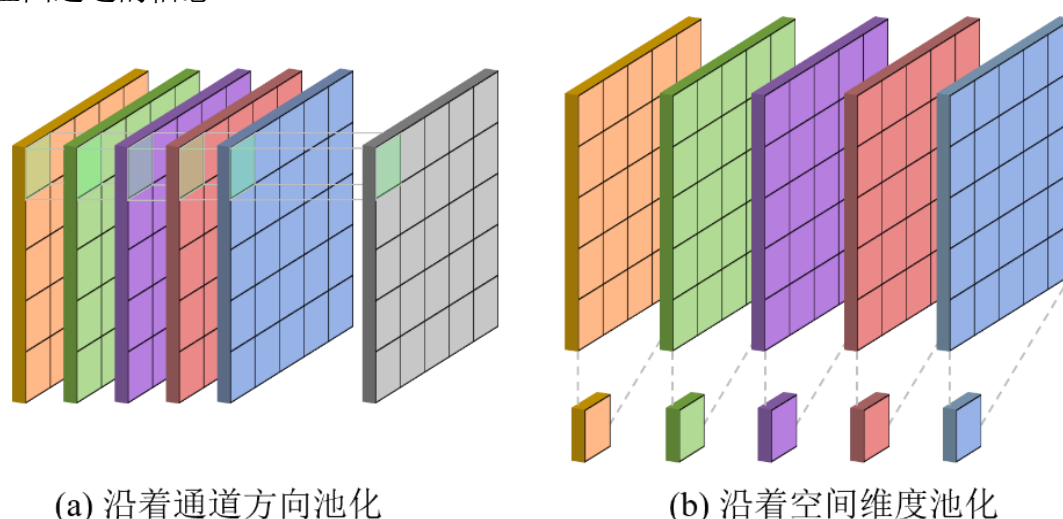


图 2-1 全局池化操作

§ 2.1.3 全连接层

全连接层也称为密集连接层，它将前一层中所有神经元的输出连接到当前层的每个神经元上，是一种常见的神经网络层。在全连接层中，每个输入特征与当前层的每个神经元都有一定的权重连接，并通过激活函数进行非线性转换。全连接层的作用是将卷积层和池化层输出的特征图转化为最终的分类结果，将高维的特征向量映射到类别标签上，通过多个全连接层的组合，可以提取出更复杂的特征，从而提高网络的分类性能。需要注意的是，在深度卷积神经网络中，全连接层的参数数量非常大，容易导致过拟合问题，可以使用卷积层和全局平均池化替代全连接层，减少网络参数量。

§ 2.1.4 激活函数

激活函数在卷积神经网络中起到非常重要的作用。通过引入激活函数能够增加卷积神经网络的非线性表达能力，从而使网络能够拟合更多的函数。除此之外，激活函数常用在全连接层中，将全连接层的输出映射为一个 0 到 1 的实数，作为类别的分类得分。常用的激活函数有 **relu**、**sigmoid**、**softmax** 等。对于 **relu** 激活函数，它是目前最常用的卷积神经网络激活函数之一，其计算公式如下：

$$f(x) = \max(0, x) \quad (2-1)$$

relu 激活函数的特点是计算速度快，能够缓解梯度消失问题。在本文中，除了在网络结构中应用 **relu** 激活函数增加网络的非线性表示以外，在生成类激活图时也使用 **relu** 激活函数将负值置 0。

在本文中，我们还将 **sigmoid** 和 **softmax** 激活函数应用在全连接层的输出上。**sigmoid** 激活函数将网络输出转化为一个概率值，表示该输出属于正类的概率，常用在二分类任务中，其计算公式如下：

$$y_i = \sigma(z_i) = \frac{1}{1+e^{-z_i}} \quad (2-2)$$

其中， z_i 是全连接层的输出。除此之外，**sigmoid** 函数还可应用在多标签分类任务中，即独立地对每个输出进行计算，得到每个类别的输出概率。而 **softmax** 函数主要用于多分类任务中，将输出限制在 0 到 1 之间，并且使得所有类别输出的概率之和为 1，其计算公式如下：

$$y_i = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}} \quad (2-3)$$

从公式 (2-3) 中可以看出，在计算第 i 类的输出概率时，会考虑所有类别的输出情况，因此，**softmax** 函数是类别互斥的。

§ 2.1.5 类激活映射

在卷积神经网络发展初期，研究人员通过大量的实验证明了卷积神经网络在各种视觉任务中的有效性，然而却对其缺乏解释。Zhou 等人^[43]通过利用预测类别的权重与卷积网络最后一层输出的特征图相乘能够得到类别在位置上的激活图。类激活图能够反映了网络的分类依据，即能够定位目标对象在图像的大致位置。Zhou 等人^[43]将热力图称为类激活图，并把这种技术称为类激活映射。生成类激活图的计算过程如 2-2 所示，卷积神经网络在最后一层输出 K 个特征图，对于目标类别 c ，分类器权重 W_k^c 表示第 k 个特征图对于预测目标类别 c 的贡献程度。我们用 f_k 代表卷积神经网络最后一层输出的第 k 个特征图，那么类激活映射可以通过公式得到目标类别 c 的类激活图：

$$CAM_c = \sum_k^K W_k^c f_k \quad (2-4)$$

这为弱监督语义分割算法开拓了新的思路,即通过分类任务也能够得到目标对象的大致位置。与全监督分割算法相比,利用图像级标签的分割算法能够大幅度减少标注成本。

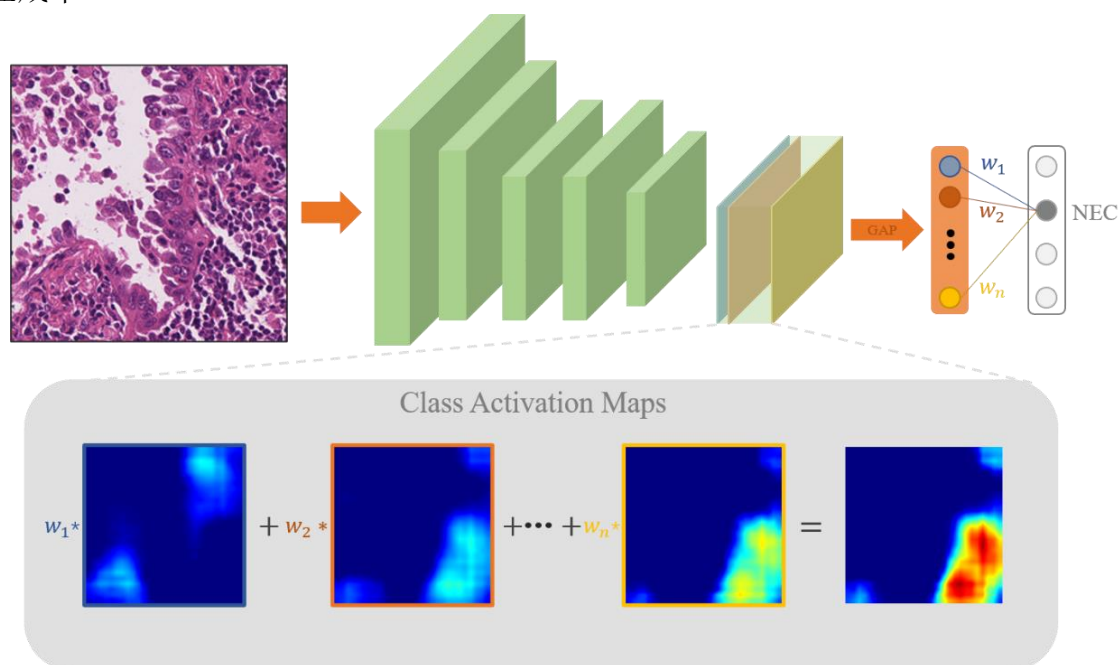


图 2-2 类激活映射

§ 2.2 残差网络

在传统的卷积神经网络中,每个卷积层的输出作为下一个卷积层的输入,通过不断叠加多个卷积层,神经网络就能够学习到更多抽象特征,研究人员认为随着网络层数的继续增加,神经网络能够提取的语义特征就越丰富,性能就越好。然而,实际上则相反,研究人员通过实验发现,网络在超过一定深度之后性能反而越来越差。为了缓解这一问题,来自微软实验室的 He 等人^[44]提出残差网络结构,利用跳层连接的方式,将基本网络结构的输入特征 x 与输出特征 $\mathcal{F}(x)$ 相加作为下一层的输入,从而缓解深层网络在前向传播过程中信息丢失的问题。前向传播过程可以用以下公式表示:

$$y = \mathcal{F}(x) + x \quad (2-5)$$

来表示,其中 x 表示输入特征, y 表示输出特征,由于输入特征 x 的加入,所以 $\mathcal{F}(\cdot)$ 只需要负责学习输出特征与输入特征之间的残差信息即可,即使后续的卷积层学习不到任何有用的特征,网络仍然可以学习到原始输入的标识性信息。跨层连接还可以使得梯度更加顺畅地传播,从而缓解梯度消失的问题。基于残差结构,研究人员可以设计更加深层次的网络结构,使得网络可以学习更加复杂的特征。与此同时,残差网络的设计思路被广泛应用于计算机视觉和自然语言处理等领域,激发了更多研究者对深度学习的探索和研究,更多的残差网络见参考文献^[42,45]。

§ 2.3 视觉注意力机制

对于复杂场景，人类大脑通常选择性地处理一些感兴趣的区域。类似地，在视觉任务中，我们也可以用注意力机制来使深度神经网络自动关注感兴趣的区域。在视觉任务中的注意力机制是指根据任务要求和场景特征，自动选择或加强输入图像中某些区域的特征表示，以提高模型的性能。在实际应用中，视觉注意力机制通常被分为通道注意力、空间注意力、时间注意和力分支注意力^[46]。在该小节，我们主要介绍通道注意力和空间注意力。

§ 2.3.1 通道注意力

通道注意力通过自适应地调整不同通道的权重来加强或削弱输入特征图的某些通道信息。众所周知，特征图的每一个通道都代表某种特征，为每一个通道赋予不同的权重能够引导网络关注有效特征，抑制无关特征。SENet^[47]是第一篇使用通道注意力的工作，这篇论文中的核心思想是通过 Squeeze 和 Excitation 两个操作来构建通道注意力模块。其结构如图 2-3 所示，Squeeze 操作是将对网络中每个通道的特征图进行空间维度的全局平均池化，得到包含全局信息的一维向量，用于表示各个通道的重要性。而 Excitation 操作则是对该向量进行非线性变换得到注意力图。将这个注意力图与输入特征图逐像素相乘，就能够得到加权后的特征图。通道注意力能够使网络自动学习对任务有意义的特征。

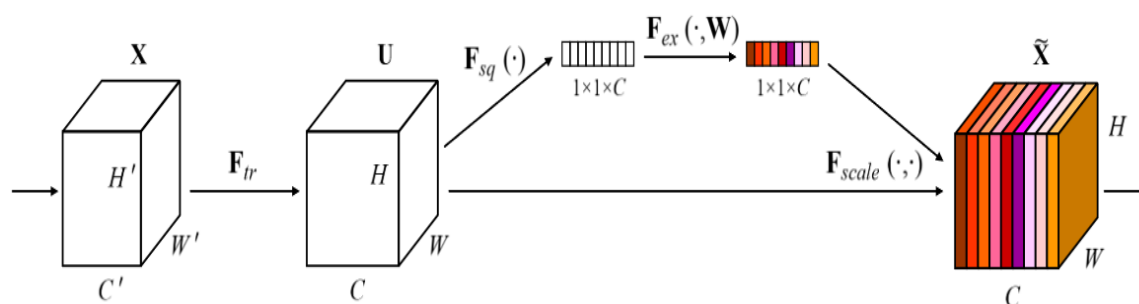


图 2-3 通道注意力

§ 2.3.2 空间注意力

除了需要考虑不同通道之间的关系，还需要考虑像素之间的关系。空间注意力是一种能够学习特征图中不同像素权重的方法，可以在像素级别上加强重要区域的显示。Woo 等人^[48]在论文中提出一种空间注意力模块，其结构如图 2-4 所示。他们在输入特征图执行通道方向的全局池化操作，再通过卷积操作计算特征图不同位置的重要性，其计算过程可以公式表示：

$$M_s(F) = \sigma(f^{7*7}([AvgPool(F), MaxPool(F)])) \quad (2-6)$$

其中, f^{7*7} 表示卷积核尺寸为 7×7 的卷积层, F 表示输入特征图, $AvgPool$ 和 $MaxPool$ 分别代表通道方向的最大平均池化和全局最大池化操作, $M_s(F)$ 代表空间注意力图。

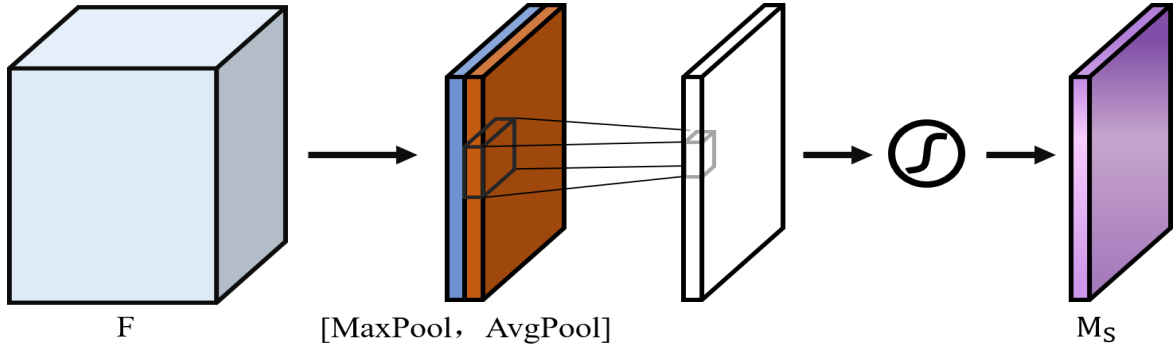


图 2-4 空间注意力

§ 2.4 网络训练以及性能评估

§ 2.4.1 损失函数介绍

损失函数是网络训练过程中的关键组成部分之一,它的作用是衡量网络预测结果和真实标签之间的差距。网络训练的目标是最小化损失函数的值,以使网络能够尽可能的预测新的样本。损失函数的重要性在于,它为网络提供了一个优化目标。通过在训练过程中最小化损失值,逐渐调整参数以提高网络的预测能力。在实际应用中,损失函数的选择通常取决于任务的类型。本文研究内容是通过训练一个多标签分类网络来区分四种肺腺癌组织,应用的损失函数主要有:

(1) 二元交叉熵损失函数: 二元交叉熵损失函数 (Binary Cross-Entropy Loss) 是结合 sigmoid 激活函数和交叉熵的函数,常用于多标签分类问题。二元交叉熵损失函数的计算方法如公式 2-7 所示,其中, K 表示类别总数, $\hat{y}_k \in \{0, 1\}$, 表示第 k 个组织的标签值,而 y_k 是 sigmoid 函数 (公式 2-2) 的输出。

$$\mathcal{L}_{bce} = -\frac{1}{K} \sum_{k=1}^K \hat{y}_k \log y_k + (1 - \hat{y}_k) \log (1 - y_k) \quad (2-7)$$

(2) softmax 交叉熵损失函数: softmax 交叉熵损失函数 (Softmax Cross-Entropy Loss) 是结合 softmax 激活函数和交叉熵的损失函数,常用于多分类任务中。在多分类任务中,类别之间是互斥的,即每一个样本只能对应一个标签。在多标签网络收敛后,我们利用 softmax 交叉熵损失函数额外训练一个全连接层,抑制非目标特征的表达。其计算方法如公式 2-8 所示。其中, y_k 是 softmax 激活函数 (公式 2-3) 的输出。

$$\mathcal{L}_{sce} = -\frac{1}{K} \sum_{k=1}^K \hat{y}_k \log y_k \quad (2-8)$$

§ 2.4.2 分割算法性能评估指标

分割算法的评估指标用于度量算法在像素级别的分割任务中的性能表现，评估指标的选择和评估方法的正确使用可以帮助我们更好地理解算法的优劣，从而更好地指导算法改进和应用。各个指标计算如下所示，其中 K 表示预测的类别总数，TP、FP、FN、TN 的含义如图 2-5 所示。

(1) 像素精度 (Pixel Accuracy, Acc)

像素精度指标指的是正确分类的数量与实际数量的比值，是图像分割中最为简单的评价指标，主要衡量模型对每种类别的分割性能，其计算公式如下：

$$Acc = \frac{TP}{TP+FP+TN+FN} \quad (2-9)$$

(2) 交并比 (Intersection over union, IoU)

交并比是语义分割任务中最常用的指标，它表示预测结果与实际结果重叠部分占两者并集的比例，是各种基准数据集最常用的指标之一，其计算公式如下：

$$IoU = \frac{TP}{TP+FP+FN} \quad (2-10)$$

(3) 频率权重交并比 (Frequency weight intersection over union, FwIoU)

频率权重交并比是另一种评价语义分割算法的指标，它会根据每个类别出现的频率分别赋予不同的权重，其计算公式如下：

$$FwIoU = \frac{TP+FN}{TP+FP+TN+FN} * IoU \quad (2-11)$$

(4) 均交并比 (Mean intersection over union, MIoU)

均交并比也是语义分割任务中最常用的指标之一，其计算公式如下：

$$MIoU = \frac{1}{K} \frac{TP}{TP+FP+FN} \quad (2-12)$$

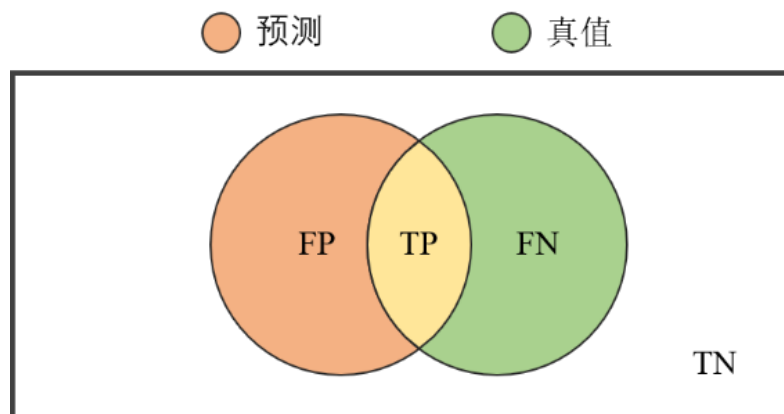


图 2-5 TP、FP、FN 以及 TN 的含义

§ 2.5 本章小结

本章简单介绍了相关理论知识，2.1 节介绍了卷积神经网络的基础知识，主要包括卷积层、池化层、全连接层以及类激活映射等相关知识，其中全局池化技术是 2.3 节视觉注意力机制的前置知识，而类激活映射是利用图像级标签进行分割算法研究的基础知识。2.2 节介绍了残差网络结构，许多网络结构都是由残差网络结构衍生而成的。2.3 节介绍了视觉注意力机制，视觉注意力机制是应用在视觉任务中重要的算法，选用视觉注意力机制能够提高模型的性能和鲁棒性。2.4 节介绍了网络训练和性能评估所涉及到的损失函数和评估指标。

第三章 基于弱监督学习的肺腺癌组织分割算法

肺腺癌是一种恶性肿瘤，通常起源于肺部中的异常组织，是肺癌中最常见的类型之一^[49]。目前，应对肺腺癌的有效手段是早期发现、早期诊断和早期治疗。当涉及到诊断和治疗肺腺癌时，不同组织对于治疗的反应和预后都有不同的影响，因此区分肺腺癌中的不同组织对于治疗和预后评估非常重要，同时也有助于推动肺癌研究和治疗的发展。

在本章节中，我们将介绍如何利用图像级标签训练网络并完成不同肺腺癌组织分割。首先，我们将介绍该方法的整体结构和各个部分的工作流程，其次说明数据集以及实验参数设置，最后对实验结果进行分析。

§ 3.1 弱监督语义分割模型的构建

在本节中，我们将介绍如何利用图像级标签训练一个弱监督语义分割模型，并生成类激活图来区分不同组织。本文方法的整体工作流程如图 3-1 所示，主要包含三个部分：第一个部分，采用 WiderResNet38^[42]作为主干网络提取特征；第二部分，结合通道注意力和渐进失活注意力生成更大范围的类激活图，并提高对重要特征的关注；第三部分，重新类激活映射，在仅有图像级标签的情况下对激活区域进行惩罚，优化类激活图。

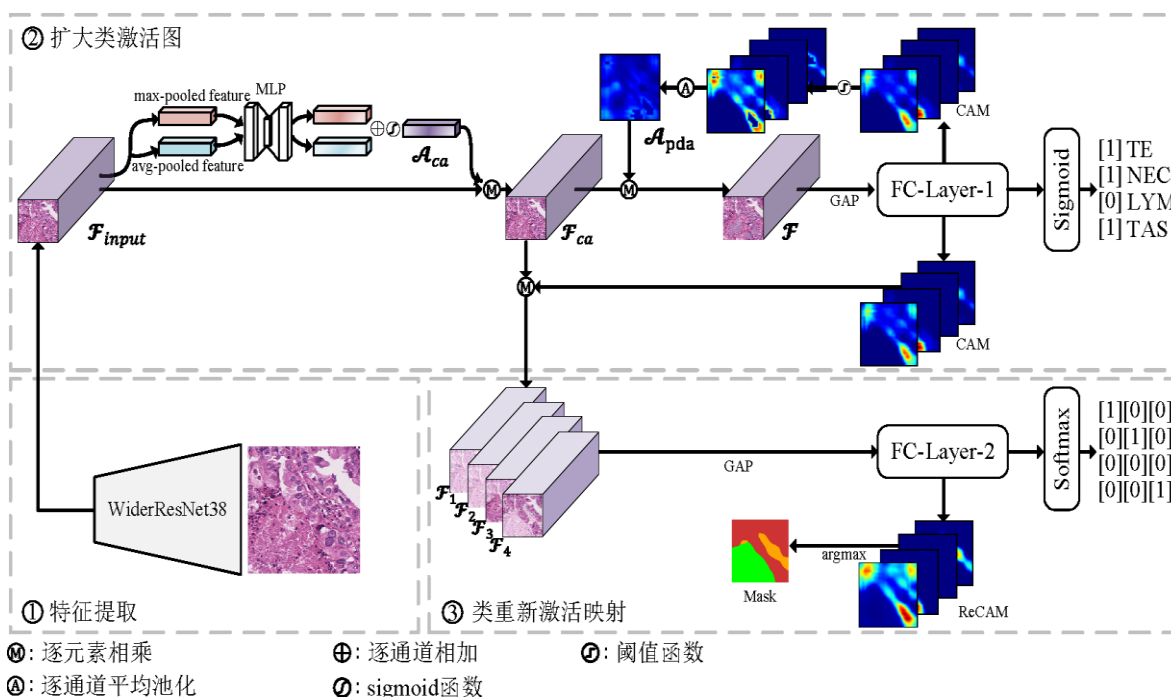


图 3-1 算法工作流程

§ 3.1.1 特征提取

主干网络 WiderResNet38 是基于残差结构而构建的，它通过增加每层的通道数来提升网络的性能，是弱监督分割算法中常用的网络之一。在本文中，我们利用 WiderResNet38 来提取组织病理学图像特征。为了方便描述算法流程，我们令数据集 $D = \{(x, \hat{y}) | x \in X, \hat{y} \in Y\}$ ， x 表示输入图像，而 \hat{y} 表示标签，是一个一维多标签向量，向量的每一个位置代表对应的组织类别是否存在，值为 1 表示存在，值为 0 表示不存在。 $y = f(x)$ 表示多标签分类网络的预测结果，在训练多标签分类网络过程中，我们使用二元交叉熵损失函数（公式 2-7）计算分类损失。

§ 3.1.2 扩大类激活图

正如之前所述，由类激活映射方法生成的类激活图只能关注目标对象最具有判别性的区域而无法覆盖全部对象，因此，我们结合通道注意力（Channel Attention, CA）和渐进失活注意力（Progressive Dropout Attention, PDA）来扩大类激活图范围并加强重要特征的激活。通道注意力的作用是让网络自动学习不同通道的重要性，为不同的通道赋予不同的权重，使得网络自适应地关注更为重要的特征，从而加强重要特征的激活。在这一部分，我们首先通过全局最大池化和全局平均池化来聚合空间维度的信息，生成平均池化特征和最大池化特征。然后将两个一维特征输入到具有一层隐藏层的多层感知机来学习不同通道的重要性。为了减少参数量，多层感知机中的隐藏层神经元个数为 C/R ，其中 C 是输入层和输出层神经元个数，而 R 表示缩减因子，在本文中，我们将 C 设置为 4096，而 R 设置为 1024。我们将多层感知机应用于每个一维向量之后，将结果逐元素求和并输出特征向量，再将其送入 sigmoid 函数生成通道注意力图 \mathcal{A}_{ca} ，sigmoid 函数的作用是将特征向量中的数值归一化到 0 到 1 之间来代表不同通道的重要性程度。具体计算公式如下：

$$\mathcal{A}_{ca} = \sigma(MLP(GAP(\mathcal{F}_{input})) + MLP(GMP(\mathcal{F}_{input}))) \quad (3-1)$$

其中， \mathcal{F}_{input} 表示主干网络提取的特征图， GAP 和 GMP 分别代表全局平均池化和全局最大池化， MLP 表示多层感知机运算， σ 表示 sigmoid 函数。通道注意力图 \mathcal{A}_{ca} 与输入特征图逐元素相乘得到新的特征图 \mathcal{F}_{ca} ：

$$\mathcal{F}_{ca} = \mathcal{A}_{ca} \otimes \mathcal{F}_{input} \quad (3-2)$$

除此之外，我们加入渐进失活注意力扩大类激活图范围。网络在丢弃判别性特征之后，会探索更多的区域来维持分类预测，进而达到扩大类激活图的目的。首先，我们提取类激活图：

$$CAM_k = \frac{ReLU(A_k)}{\max(ReLU(A_k))}, A_k = w_k * F_{ca} \quad (3-3)$$

其中, w_k 表示全连接层的权重, 代表了特征图对预测类别 k 的贡献。然后, 我们设定一个阈值 th 用于计算类激活图丢弃的范围, 其阈值 th 的计算公式如下:

$$th = \max(CAM_k(i, j)) * u_t \quad (3-4)$$

其中, i 和 j 表示类激活图中的坐标, 参数 u_t 表示丢弃系数, t 表示训练轮次。如果类激活图中的值大于阈值, 那么就会设置为 0, 所有类激活图沿着通道方向进行全局平均池化后得到注意力图 \mathcal{A}_{pda} 。其计算公式如下所示:

$$CAM_k(i, j) = \begin{cases} CAM_k(i, j), & CAM_k(i, j) \leq th \\ 0, & CAM_k(i, j) > th \end{cases} \quad (3-5)$$

$$\mathcal{A}_{pda} = \frac{1}{K} \sum_k CAM_k \quad (3-6)$$

我们知道在训练初期, 网络的预测能力有限, 但随着训练轮次的增加, 网络提取特征能力越强, 预测值就越接近目标值, 而类激活图的响应则越强烈。因此, 为了适应类激活图的变化, 我们设置一个衰减率 d 来控制参数 u_t 的变化, 训练越深入, u_t 值就越小, 丢弃的判别性区域就越多, 直到 u_t 达到下限值 l , 阈值 u_t 的计算公式如下:

$$u_t = \max(l, u_{t-1} * d) \quad (3-7)$$

其中, 参数 t 和参数 d 分别表示训练的轮次和衰减率, 在实践过程中, 我们设置 $d = 0.985$ 和 $l = 0.65$, 最终的注意力图 \mathcal{A}_{pda} 与特征图 F_{ca} 逐像素相乘得到最终特征图

$$\mathcal{F} = \mathcal{A}_{pda} \otimes F_{ca} \quad (3-8)$$

其中, \otimes 表示逐像素的乘法。将 \mathcal{F} 进行全局平均池化, 后送入全连接层 (图 3-1 中的 FC-Layer-1) 得到网络的分类预测, 再利用公式 2-7 计算损失 \mathcal{L}_{bce} 。

§ 3.1.3 类重新激活映射

二元交叉熵损失函数在使用图像级标签的弱监督分割方法中被广泛使用, 但该函数并不能惩罚错误激活的区域。出现这种情况的本质原因是因为二元交叉熵损失函数并不是类别互斥的。受 Chen 等人的启发^[50], 我们利用类重新激活映射方法 (Class Re-Activation Mapping, ReCAM) 抑制错误激活区域, 即将二元交叉熵损失函数替换为 softmax 交叉熵损失函数, 并学习另一个全连接层 (即 FC-layer-2) 来重新生成类激活图。具体实现如下: 将收敛网络生成的类激活图 CAM_k 与 F_{ca} 相乘, 能够得到 K 个特征图 \mathcal{F}_k , 这样就能够保留每个标签对应的特征图信息。同时, 我们将多标签向量分解成多个单标签向量, 从而把多标签分类任务转化为单标签多分类任务。那么对非目标对象有贡献的特征因为 softmax 交叉熵损失函数的类别互斥特性得到惩罚, 从而能够使网络学习到更多与目标对象相关的特征信息。多标签特征转化为单标签特征的公式如下:

$$\mathcal{F}_k = CAM_k \otimes \mathcal{F}_{ca} \quad (3-9)$$

通过全局平均池化操作并送入 FC-Layer-2 得到每个单标签特征的分类预测，然后利用 softmax 交叉熵损失函数（公式 2-8）计算损失 \mathcal{L}_{sce} ，最后总的损失为：

$$\mathcal{L} = \mathcal{L}_{bce} + \mathcal{L}_{sce} \quad (3-10)$$

直到模型再次收敛，可以生成最后的重新类激活图，其公式如下：

$$ReCAM_k = \frac{ReLU(A_k)}{\max(ReLU(A_k))}, A_k = w_k * \mathcal{F} \quad (3-11)$$

由于本文使用的图像级标签作为监督信息，在判断组织是否存在时，分类结果更具有说服力。因此，对于第 k 个组织的重新激活图 $ReCAM_k$ ，如果该组织的分类预测低于阈值 ϵ ，则将重新类激活图 $ReCAM_k$ 置为 0，最后通过 \argmax 函数得到最终的分割结果。在实践中，我们将 ϵ 设置为 0.15，其计算公式如下：

$$ReCAM_k = \begin{cases} 0 * ReCAM_k, & y_k \leq \epsilon \\ ReCAM_k, & y_k > \epsilon \end{cases} \quad (3-12)$$

$$mask = \argmax(ReCAM) \quad (3-13)$$

§ 3.2 数据集以及实验设置

§ 3.2.1 数据集介绍

LUAD-HistoSeg 是由 Han 等人^[40]发布的肺腺癌组织病理学图像数据集，数据集包含了肿瘤上皮组织（Tumor Epithelial, TE）、坏死（Necrosis, NEC）、淋巴细胞（Lymphocyte, LYM）以及肿瘤相关基质（Tumor-associated stroma, TAS）四种不同的组织。肺腺癌的肿瘤上皮组织通常由多层柱状细胞构成，坏死指的是由于肿瘤细胞缺乏氧气和营养物质，导致组织坏死，通常显示为低密度区域，淋巴细胞则是对肺腺癌进行免疫攻击的一种细胞，通常呈圆形或者椭圆形，肺腺癌的肿瘤基质指的是肺腺癌细胞周围的非癌细胞成分。LUAD-HistoSeg 数据集来源于广东省人民医院 29 例和 TCGA 中 20 例肺腺癌患者的病理切片，三名经验丰富的病理学家检查所有的切片并选择最具有代表性的切片作为研究对象，每个切片都通过数字病理切片扫描仪进行扫描，然后以 10 倍物镜放大率随机裁剪了 800 个图像块，图像块尺寸为 224×224 。另外，由于肺中存在许多肺泡，因此图像中存在许多的白色区域，专家在筛选时丢弃了背景模糊、脏、大量白色背景的图像块，最后留下 17285 张图像块作为最终的数据集。为了验证模型性能，LUAD-HistoSeg 数据集被划分为训练集（16678 张图像块，图像级标签）、验证集（300 张图像块，像素级标签）和测试集（307 张图像块，像素级标签）。图像级标签是包含四个元素的多标签向量，四个元素按照顺序依次代表了肿瘤上皮组织、坏死组织、淋巴细胞以及肿瘤相关基质。图 3-2 显示了 LUAD-HistoSeg 训练集、验证集以及测试集中不同图像级别标签的数量，我们可以发现肿瘤上皮组织和

肿瘤相关基质组织经常伴随出现。同时,我们还统计了测试集中不同组织的像素占比,肿瘤上皮组织、坏死、淋巴细胞以及肿瘤相关基质分别占测试集总像素数量的 42%、7.5%、13.4%以及 37.1%。

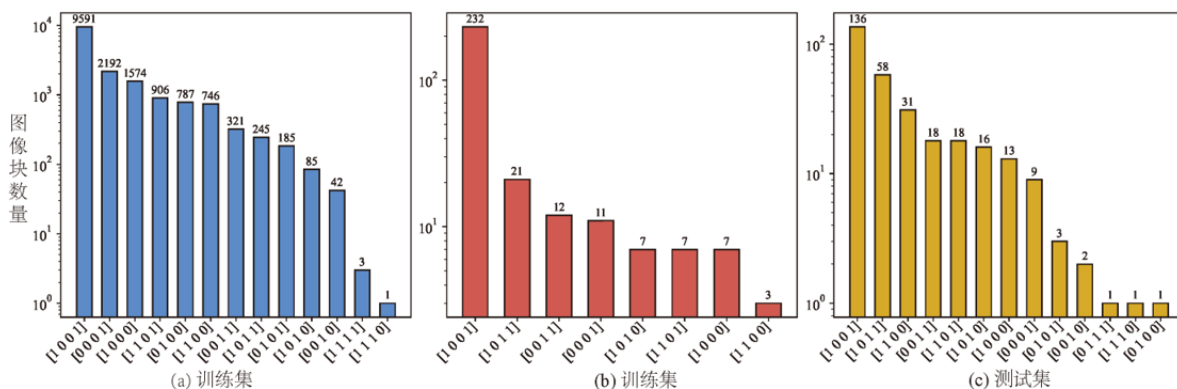


图 3-2 LUAD-HistoSeg 训练集、验证集以及测试集标签分布

§ 3.2.2 实验参数设置

在我们的实验中,所有的神经网络都是在 PyTorch 框架中实现的。实验均在在具有 NVIDIA RTX 1080Ti 的服务器上进行训练。具体实验设置如下:在多标签分类训练部分, batch size 设置为 20, epoch 设置为 20。输入图像以 50%的概率进行随机水平和垂直翻转数据增强。我们将 learning rate 设置为 0.01 并使用多项式衰减策略来更新参数。在重新类激活映射部分,三个超参数 batch size、epoch 和 learning rate 分别设置为 16、10 和 0.005。我们通过每种类别的交并比、频率权重交并比、平均交并比以及像素精度来评估我们提出的方法。在评估所有实验中的性能时,白色背景区域的计算被排除在外。

§ 3.3 实验结果分析

在这一小节中,我们开展了消融实验和对比实验,并从评估指标和分割结果两个方面说明本文方法的有效性。消融实验的目的是为了验证方法中各个模块对网络性能的提升作用,而对比实验的目的则是为了与现有先进方法进行比较,展现本文方法的优势,最后,我们还对方法的局限性进行讨论。

§ 3.3.1 消融实验结果与分析

我们设置了七组实验用于验证主干网络以及模块的有效性,首先是主干网络的选择。在弱监督语义分割领域,目前比较流行的网络是 ResNet 及其变体。在选择主干

网络时，我们通过对比 ResNet50、ResNet101、ResNet152 以及 WiderResNet38 网络的表现，发现 WiderResNet38 在所有网络中表现最好，其 MIoU 指标达到 71.45%。因此，我们最终选择 WiderResNet38 作为我们方法的主干网络并基于该网络进行模块性能的评估。实验具体涉及如下：（1）仅使用 WiderResNet38 网络训练数据集，并将其作为基准网络。（2）将渐进失活注意力加入基准网络（3）渐进失活注意力和通道注意力集成到基准网络中（4）渐进失活注意力、通道注意力以及重新类激活映射集成到基准网络中。消融实验结果如表 3-1 所示。在第四组实验中，基准网络在肿瘤上皮组织、坏死组织、淋巴细胞以及肿瘤相关基质中分别实现了 72.86%、72.69%、71.30% 以及 68.95% 的 IoU。在第六组实验中，我们观察到集成了通道注意力与渐进失活注意力的网络分别在 FwIoU、MIoU 以及 ACC 指标上分别达到 74.81%、76.15% 和 85.50%。与基准网络的效果性能相比，在每个指标中分别提高了 3.62%、4.7% 和 2.39%。可以看出，通道注意力与渐进失活注意力的加入使得网络性能有了显著提高。在第七组实验中，我们发现类重新激活映射的加入提升了模型对坏死组织、淋巴细胞以及肿瘤相关基质的预测性能，在 IoU 指标上的提升幅度分别为 0.88%、0.29% 以及 0.38%。从实验结果中，我们能够验证各个模块的有效性。

表 3-1 消融实验结果（以百分比为单位）

PDA	CA	ReCAM	Backbone	TE	NEC	LYM	TAS	FwIoU	MIoU	ACC
-	-	-	ResNet50	69.29	73.86	71.00	64.81	68.23	69.76	81.05
-	-	-	ResNet101	69.34	72.57	67.05	64.08	67.32	68.26	80.42
-	-	-	ResNet152	71.54	69.31	66.14	67.20	69.04	68.55	81.70
-	-	-	WiderResNet38	72.86	72.69	71.30	68.95	71.19	71.45	83.11
√	-	-	WiderResNet38	75.57	78.08	73.70	69.70	73.32	74.26	84.51
√	√	-	WiderResNet38	76.49	81.32	75.41	71.36	74.81	76.15	85.50
√	√	√	WiderResNet38	76.50	82.20	75.70	71.74	75.05	76.53	85.66

为了进一步说明本文方法的有效性，我们对不同模块生成的类激活图进行观察。我们把真值和类激活图与原图以 50% 的比例进行混合。为了更好的观察，我们用虚线框来特别注明需要关注的区域。其中，红色虚线框展示了目标组织的区域，白色虚线框展示了与目标组织无关的区域。如图 3-3（b）所示，尽管基准网络能够很好定位组织的位置，但生成的类激活图并没有覆盖全部范围；在图 3-3（c）中，我们可以观察到肿瘤相关基质和肿瘤上皮组织的类激活图范围有所扩大；从图 3-3（d）中可以看出重要特征的激活程度有所增加。观察图 3-3 中白色框区域可以得出结论，ReCAM 减少了类激活图中的错误激活区域，使激活区域的边界更为准确。

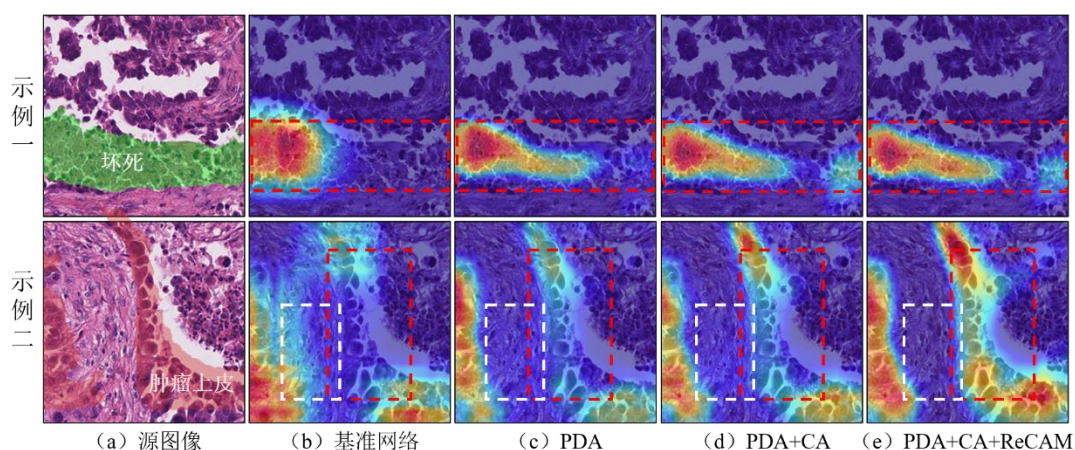


图 3-3 不同模块的类激活图

§ 3.3.2 对比实验结果与分析

表格 3-2 展示了本文方法与现有先进方法之间的性能比较。其中, HistoSegNet 和 MLPS 都是将类激活映射技术应用在组织病理学图像的弱监督分割方法^[34,40], 其余方法则是应用在自然场景图像领域的先进弱监督分割方法^[51-54]。从表格中可以看出, 本文方法在 MIoU 上达到 76.53%。相比较 MLPS 提升了 0.93%的性能, 实现了比现有方法更好的性能。特别是, 我们的方法在像素占比较少的坏死组织和淋巴细胞组织中实现了 82.20%和 75.70%的 IoU, 分别提高了 2.88%和 2.29%。同时, 我们也观察到, 本文方法在 FwIoU 与 ACC 指标上的性能与最先进方法存在细微的差距, 其主要原因在于本文方法在像素数量占比较大的肿瘤上皮组织和肿瘤相关基质上分割效果与先进方法比较还存在一定的差距。上述实验表明, 我们的方法在仅使用图像级别标签的条件下能够在整体上更好地区分肺腺癌组织。

表 3-2 对比实验结果 (以百分比为单位)

序号	方法	TE	NEC	LYM	TAS	FwIoU	MIoU	ACC
(1)	HistoSegNet	45.59	36.30	58.28	50.82	48.54	47.75	65.97
(2)	SC-CAM	68.29	64.28	62.06	61.79	64.74	64.10	78.70
(3)	OAA	69.56	53.56	67.18	62.90	65.58	63.30	79.25
(4)	Grad-CAM++	72.90	74.18	67.93	66.01	69.78	70.26	81.97
(5)	CGNet	71.85	73.30	69.09	67.26	68.89	70.38	82.21
(6)	MLPS	77.70	79.32	73.41	71.98	75.13	75.60	85.70
(7)	本文	76.50	82.20	75.70	71.74	75.05	76.53	85.66

如图 3-4 所示, 我们对比了先进方法于本文方法的分割结果。每一行代表着一个示例, 第一列表示源图像, 其余列依次代表真值、MLPS 以及本文方法的预测结果。为了方便观察, 真值、预测结果与源图像以 50%的比例进行混合。除此之外, 在图中,

我们以黑色虚线框来展示本文方法与先进方法的主要差异。在第一行中，源图像包括了肿瘤上皮组织、淋巴细胞、以及肿瘤相关基质，从 MLPS 方法预测结果（即图 3-4（c）示例一）中可以看出，MLPS 将淋巴细胞识别为肿瘤上皮组织。从图 3-4（d）示例一中的结果来看，本文方法能够在淋巴细胞的识别上具有优势，这与本文方法在淋巴细胞的评估指标结果相呼应。从图 3-4（c）示例二的结果来看，MLPS 方法对于坏死组织（其 IoU 指标为 79.32%）的识别能力还有待提高，而本文方法相比 MLPS 方法能够更好地识别坏死组织，这与本文方法在坏死组织上实现 82.20%IoU 相呼应。综上，我们通过比较发现本文方法在坏死组织和淋巴细胞上的分割要优于当前先进的方法，并在整体指标 MIoU 上达到最好的效果。

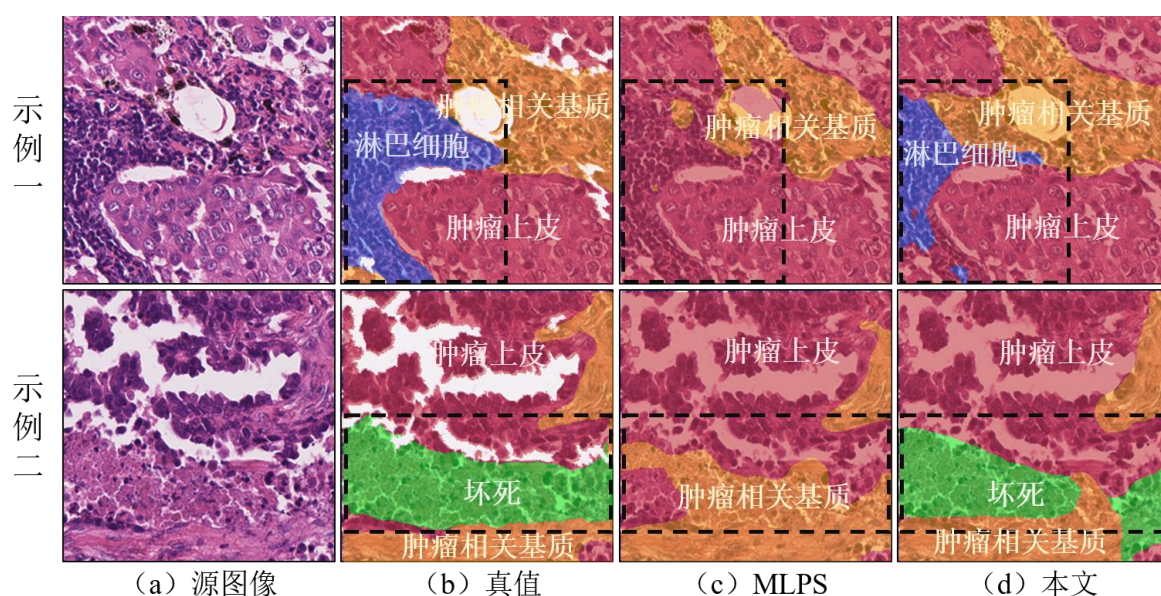


图 3-4 与先进方法比较

§ 3.3.3 局限性分析

我们挑选一些样本进行讨论，如图 3-5 所示，每一列代表一个示例，第一行表示源图像，第二行表示真值，第三行表示本文方法的预测结果。观察图 3-5 可知，示例一源图像中包括了淋巴细胞和肿瘤上皮组织，示例二源图像中包括了肿瘤上皮组织和肿瘤相关基质，在这两张图像中，组织与组织之间的边界比较清晰，且组织形状较为简单，因此本文方法在这两张图像中的分割效果较好。与示例二类似，示例三也包括了肿瘤相关基质与肿瘤上皮组织，主要区别在于示例三中肿瘤相关基质区域呈树枝状，范围较小且比较分散，在本文方法的预测结果中，我们只定位了肿瘤相关基质的部分区域，且组织之间的边界识别较差。此外，在示例四中，肿瘤相关基质与淋巴细胞的颜色、纹理相近，导致模型将全部淋巴细胞识别为肿瘤相关基质。综上，对于结构复杂、颜色纹理相近的组织，本文方法的区分能力还有待提高。

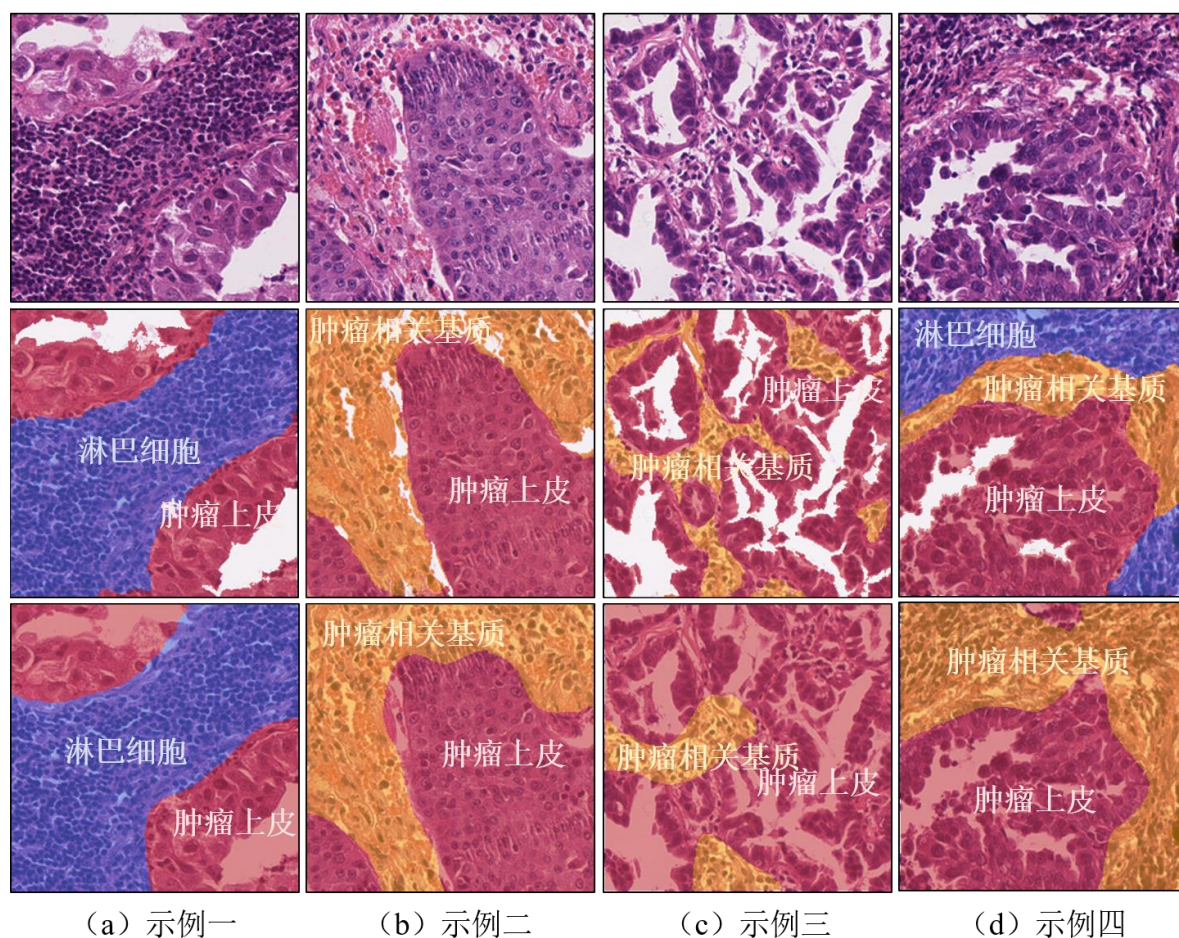


图 3-5 预测结果分析

§ 3.4 本章小结

使用图像级标签来获得准确的组织边界是一项挑战性但实际的任务。如前所述，类激活图只能精确定位目标组织的位置，不能覆盖目标组织整个区域。在此基础上，我们结合通道注意力和渐进失活注意力来扩大类激活图范围，渐进失活注意力通过丢弃判别性特征引导网络扩展类激活图区域。此外，本文根据特征图的平均池化特征和最大池化特征生成通道注意力图，为不同通道分配不同的权重，提升重要特征的激活程度。如图 3-3 示例所示，由于缺少位置信息，损失函数无法在像素级别惩罚错误的分类，因此同一像素可能同时响应不同的组织，我们使用 softmax 交叉熵损失函数将多标签分类任务转换为单标签分类任务，微调收敛模型，减少无关类别的激活。总之，上述实验表明，我们的方法可以通过仅使用图像级别标签更好地完成肺腺癌图像的语义分割。

第四章 组织病理学图像分析与标注软件

组织病理学图像是一种用于诊断疾病的重要医学图像，其特点是细节复杂、结构多样，需要经验丰富的病理医生进行解读和分析。利用深度学习算法辅助病理医生区分不同组织是目前的研究趋势。其中，全监督分割算法是语义分割任务中最常用的算法，利用全监督分割算法区分不同组织是一件顺其自然的事情。然而，全监督语义分割算法需要大量的标注数据支撑。在标注组织病理学图像分割数据集时，标注人员需要利用多边形标注工具仔细描绘组织的边界，且标注工作只能由病理医生进行，而全球病理医生稀缺，让病理医生标注数据是人力资源的浪费。为了缓解这一问题，许多研究人员提出了利用不完全监督信息来完成组织分割任务，但利用不完全监督信息的分割算法性能有限，如何在标注成本和模型性能之间取得平衡是一件非常困难的事情。

在第三章，本文利用图像级标签来区分肺腺癌中不同组织，并在 MIoU 指标上达到了 76.53%，证明使用不完全监督信息训练的模型也能够得到较好的分割效果。但本文认为图像级标签与像素级标签存在天然的信息鸿沟，经过海量数据训练的模型才能应用在实际的业务场景中，海量数据的需要与数据标注困难成为了一对不可调和的矛盾。对此，本文认为一个有效的解决方案是结合深度学习模型与人工标注来减少数据标注成本。因此，本文开发一款集成深度学习分割算法的组织病理学图像分析与标注软件，除了为用户提供组织病理学图像分析功能以外，还提供深度学习模型供用户使用。集成深度学习模型的图像分析与标注软件的意义如下：其一，标注人员不必手动描绘组织的边界，而是可以通过微调算法分割结果得到精细的像素级标签；其二，集成深度学习模型能够为标注人员提供客观依据，提高标注准确性。

总而言之，组织病理学图像分析与标注软件可以帮助用户提高标注效率和准确性，促进医学研究和医疗质量的提升，对推动智能医疗的发展具有重要的意义。本章从需求分析、概要设计、详细设计以及功能测试等方面介绍了组织病理学图像分析与标注软件开发的各个环节。

§ 4.1 软件需求分析

组织病理学图像分析与标注软件通过结合深度学习分割算法为标注人员提供初步标注结果和客观参考，提高标注效率和准确性。在本小节中，我们将对组织病理学图像分析与标注软件进行需求分析。一款合格软件可以分为文件操作、图像处理以及图像标注三大模块。其中，文件操作模块应该能够支持用户打开文件、打开目录、选择图像、保存标注结果以及退出功能。具体需求如表 4-1 所示。

表 4-1 文件操作模块

功能名称	功能描述
打开图像	通过文件对话框打开单张图像并加载在工作区，软件支持 jpg、bmp、png 等常用的图像数据格式，如果文件所在目录中存在对应的标注文件，需要将标注文件一并加载在工作区。
打开目录	通过文件对话框打开目录，读取目录下所有待标注图像的路径，并保存在文件列表中，供后续选择。
选择图像	通过选择文件列表的文件名，加载对应图像到工作区。如果图像被标注过，则需要将相应的标注文件读取到工作区，供用户操作。
保存标注结果	如果打开的是单张图像，则标注完一张图像后，用户可以点击菜单栏的保存菜单项保存标注结果；如果打开的是目录，则标注结果直接保存在目录所在路径中。标注结果以 JSON 文件格式存储。
退出	结束标注工作并退出软件。

在该模块中的打开图像过程中，如果图像所在目录存在相应的标注文件，则需要将标注结果加载至软件中以使用户继续完成标注工作。图像处理以及分析模块包括图像几何变换、图像分割、颜色处理等功能，为用户提供图像处理服务，方便用户在标注前对图像进行分析具体功能如表 4-2 所示。

表 4-2 图像处理以及分析模块

功能名称	功能描述
几何变换	支持对工作区图像进行水平翻转、垂直翻转、旋转等功能。
颜色处理	支持对工作区图像进行颜色归一化以及标准化处理。
图像分割	利用深度学习算法图像进行分割，生成初步的多边形标注结果

图像标注模块应该支持点标注、矩形框标注、多边形标注等多种标注方式选择，同时，还应支持用户对标注结果进行编辑，比如编辑图元、撤回、恢复等操作。用户在标注时可以通过点击图像分割算法生成初步标注结果，再进行微调。由分割算法生成的标注结果以可编辑的多边形图元显示，用户通过编辑多边形图元的顶点来微调结果。具体功能需求如表 4-3 所示。

表 4-3 图像标注模块

功能名称	功能描述
矩形框标注工具	通过鼠标点击以及拖动画出目标对象的矩形框图元。
多边形标注工具	通过鼠标点击多个位置生成多边形图元。
点标注工具	通过鼠标点击在目标位置留下点图元。
编辑图元	通过鼠标拖动图元顶点来调整标注结果的形状和位置
撤回、恢复	通过撤回和恢复按钮来撤销和恢复上一步标注操作

§ 4.2 软件概要设计

开发一款病理图像分析与标注软件需要确定软件的功能和布局。在本软件中，我们采用层次结构的设计模式并确定系统的功能模块设计和图形化界面模块的设计。

§ 4.2.1 软件层次结构设计

在软件开发领域中，层次架构设计是一种常见的设计模式。该模式将不同的功能块分层，并以高内聚低耦合的方式组织，以提高软件的可移植性和可用性。在集成了深度学习模型的标注软件中，我们采用了层次架构设计，其中包括用户层、逻辑层和视图层。在用户层，用户可以通过点击按钮、控制鼠标等方式发出信号，逻辑层接收并处理这些信号，再将处理结果反馈给视图层，由视图层做出响应并重新渲染软件界面，其层次结构如图 4-1 所示。

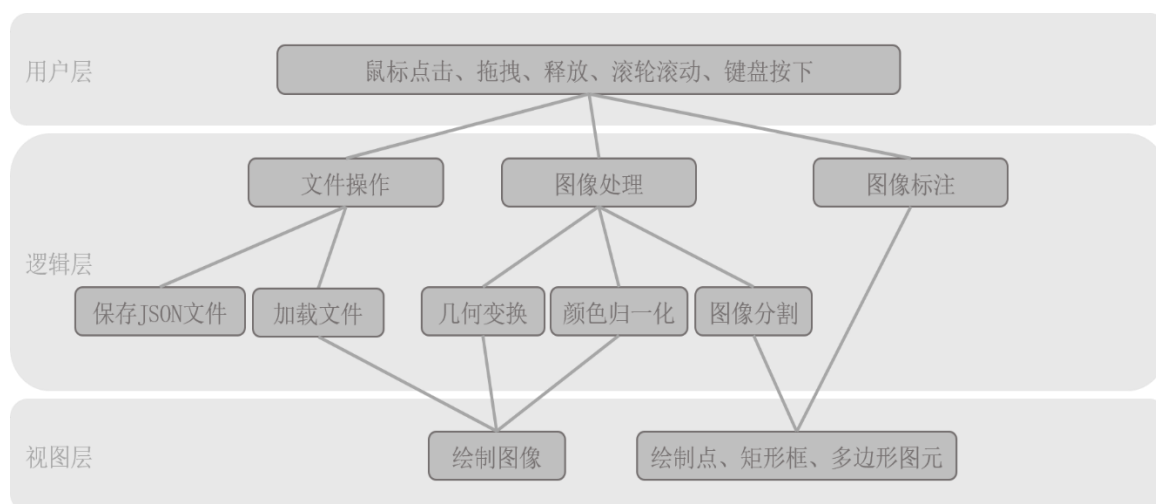


图 4-1 软件的层级结构图

§ 4.2.2 软件开发技术介绍

在开发软件初期，需要确定软件开发技术方案。在集成深度学习算法的病理图像标注软件中，我们将采用 PyQt5 作为界面开发框架，PyQt5 是一个基于 Python 的界面 UI 开发框架，通过使用 PyQt5 可以创建丰富的界面应用软件。我们选用 PyQt5 中的 QGraphicsView、QGraphicsScene、QGraphicsItem 等技术来实现图像标注功能。QGraphicsView 是一个基于 PyQt5 框架的控件，可以用于显示二维图形场景，通常与 QGraphicsScene 结合使用来实现图形项的绘制和管理。QGraphicsScene 是一个用于管理和呈现图形项的容器，它将图形项分组或分层来组织场景中的对象。QGraphicsItem 是一个图形项的基类，通过 QGraphicsItem 可以绘制场景中的对象，如矩形、椭圆等

等基本图元。

在软件中还涉及到图像处理与分析功能，因此，我们选用 `Opencv-python` 作为软件的图像处理功能的工具包。`Opencv-python` 是一种基于 `Python` 的开源计算机视觉库，包括了许多可用的计算机视觉算法，通过使用 `Opencv-python` 工具包，我们可以为用户提供更多的图像处理与分析功能。除此之外，标注软件中还集成了深度学习模型，因此需要选择必要的的数据预处理和深度学习框架，在本软件中，我们使用 `Pytorch` 神经网络库为深度学习分割模型做支撑。`Pytorch` 是一个基于 `Python` 的开源深度学习计算框架，它提供了高度灵活性和高速的计算能力，使得用户能够轻松地推理神经网络模型，除此之外，`Pytorch` 框架安装简单，为软件的可移植性提供了可能。

§ 4.3 软件详细设计

§ 4.3.1 软件界面设计

一款标注软件应该有一个简洁明了的用户界面，供用户可以轻松地进行标注工作，同时，应该使用易于理解和操作的图标和按钮，以使用户快速找到需要的工具和选项。此外，软件应该提供多种标注工具，如矩形、多边形、点等，以满足不同标注任务的需求。在本文开发的组织病理学图像分析与标注软件中，将界面分为菜单栏、工具栏、工作区、文件列表、标签列表以及状态栏。菜单栏为软件的大多数功能提供功能入口。工具栏提供各种编辑工具，方便用户快速操作。工作区则显示用于显示源图像以及图元。文件列表用于展示用户当前工作目录所包含的所有文件，通过点击文件列表能够实现切换图像。标签列表显示用户已标注好的标签。状态栏则是显示鼠标当前的坐标。

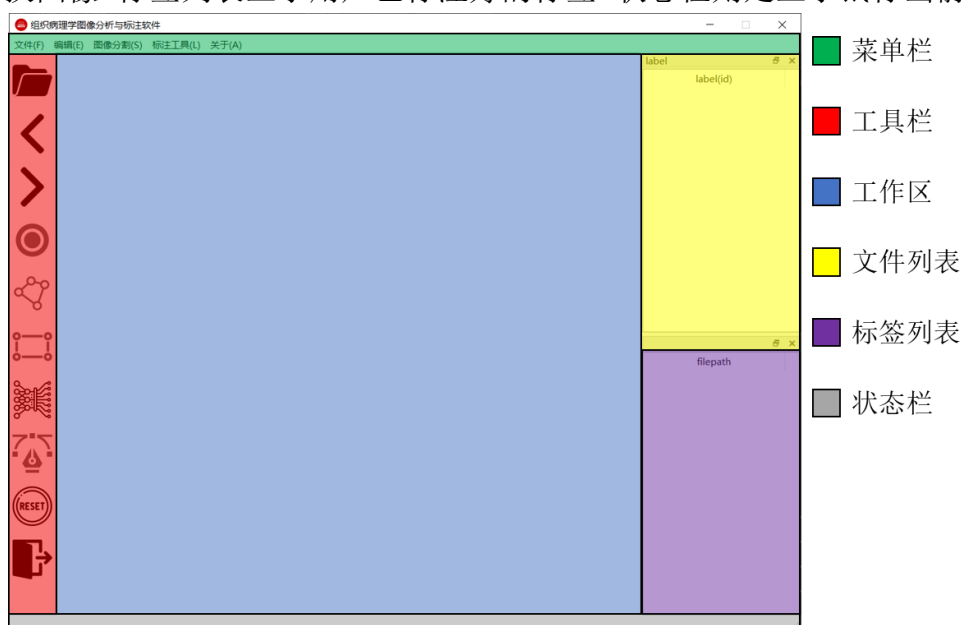


图 4-2 组织病理学病理图分析与标注软件界面

§ 4.3.2 图像分割功能

近年来,深度学习算法在图像分割领域中取得了非常卓越的成绩。在传统的图像分割算法流程中,研究人员需要基于人类对问题的先验知识和经验来手动设计特征,根据这些特征实现图像分割。然而这些手动设计的特征可能不够全面或准确,因此可能会影响分割结果的准确性。相比传统分割算法,深度学习分割算法的主要优势在于其能够使用神经网络自动地从数据中的模式和结构自动学习特征,这使得模型更加适应实际应用场景。此外,深度学习模型的灵活性也更强,可以很容易地应对复杂的场景和任务。在本软件中,我们通过利用已收敛的模型对图像进行分割,生成初步的标注结果。用户可以通过点击菜单栏或者工具栏的分割按钮,可以对目标图像进行推理生成分割掩码。在实际应用中,我们借助 OpenCV 计算机视觉库来找到分割掩码中的最大连通域,并将连通域的近似多边形转化为图元供用户编辑。

§ 4.3.3 图像几何变换功能

图像变换功能包括翻转和旋转两个功能,翻转也称为镜像变换,又可以进一步分为水平镜像和垂直镜像,水平镜像翻转即将图像左半部分和右半部分以图像数值中线为中心轴进行交换,而竖直镜像则是将图像上半部分和下半部分以图像水平中轴线为中心进行交换。在旋转功能中,指的是将图像以某一点为中心旋转一定的角度,形成新的图像过程。当然这个点通常就是图像的中心,旋转前和旋转后的点离中心的位置不变。旋转前后并不会改变图像的 RGB 颜色通道,改变的只是每一个像素的所在位置。以下是几种变换的展示:

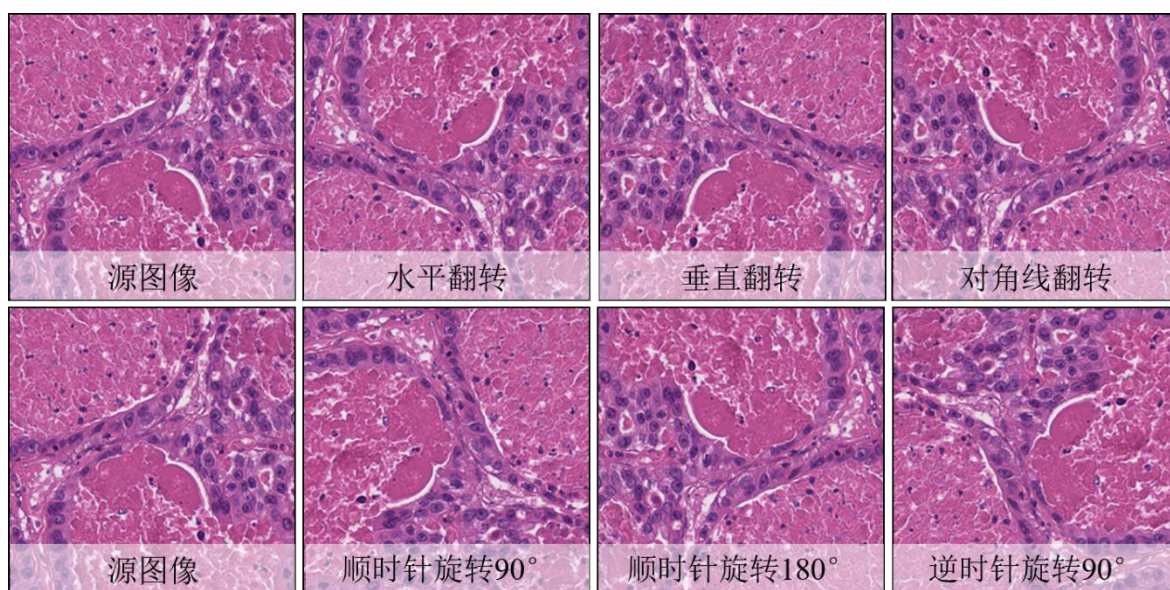


图 4-3 图像几何变换的效果

§ 4.3.4 颜色处理功能

在制作病理切片时,由于染色手法以及设备的影响,可能会导致病理切片成像后的颜色会有偏差,因此,必须根据特定染色方法和设备进行颜色校准。组织病理学图像的颜色校准是一种将不同病理学图像的颜色特征进行统一的方法,以便在不同的图像之间进行比较和分析。在组织病理学中,常用的颜色校准方法包括标准化和归一化。标准化方法将图像的像素值进行线性变换,使得图像的平均值为 0,标准差为 1。这种方法可以消除不同图像之间的亮度差异,但可能会使图像颜色信息丢失。归一化方法是将每个像素的值映射到一个特定的范围内,例如 $[0, 1]$ 或 $[-1, 1]$ 。这种方法可以保留颜色信息,但可能会在图像边缘产生过度增强的效果。这可以通过使用标准化的组织样本来完成,这些样本已经被校准并被广泛使用。总的来说,组织病理学图像的颜色校准是一个重要的步骤,可以提高不同图像之间的比较和分析的准确性和可靠性。为组织病理图像标注提供参考。

针对组织病理学图像,研究人员提出了许多颜色归一化方法^[55-57]。**Macenok** 颜色归一化方法是一种比较常用的方法,它通过估计组织病理学图像中的颜色矩阵来进行颜色归一化,这种方法能够保留组织样本中的颜色信息,同时抑制了在染色过程中产生颜色偏差。与 **Macenok** 方法不同的是, **Reinhard** 归一化方法是将图像中每一个像素的颜色映射到标准 RGB 颜色空间,这种方法比较适用于不同类型的染色方法和设备之间的颜色归一化。针对不同的光照环境, **Vahadane** 能够有效的缓解这一问题,它通过矩阵分解来组织染色和图像亮度进行归一化,可以在处理速度和处理性能中达到一个较好的平衡,常应用于处理大量图像和不同染色方法的情况。图 4-4 展示了不同颜色处理后的效果。

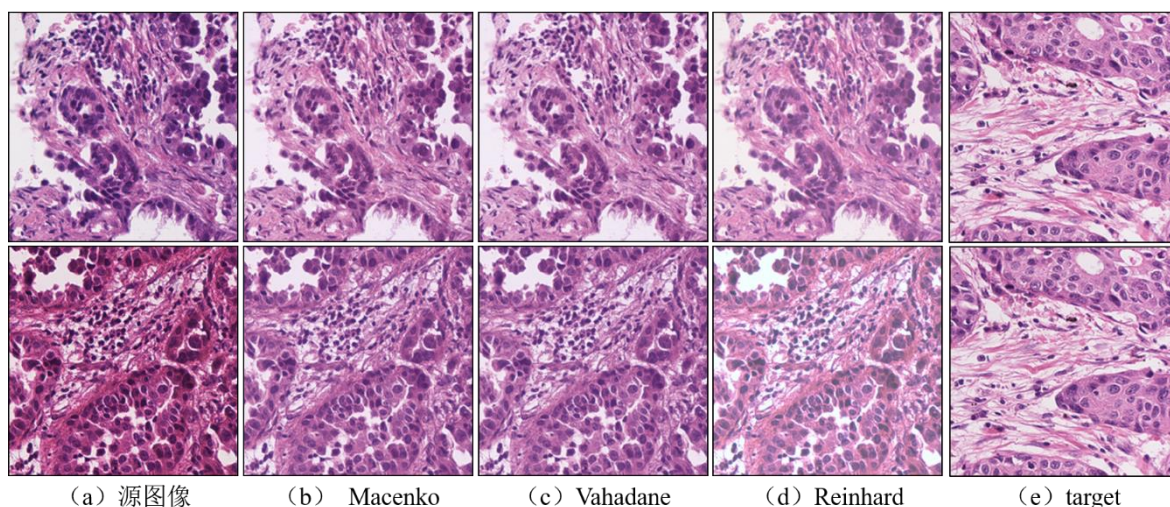


图 4-4 不同颜色归一化的效果

§ 4.3.5 图像标注功能

用户通过选择点、矩形框、多边形三种不同的标注工具对数据进行标注。在操作过程中,用户可以利用鼠标点击、鼠标拖拽、鼠标释放等操作对工作区图像进行标注。其基本流程是通过连接鼠标事件(如鼠标单击事件)来捕捉用户在 QGraphicsView 上单击的位置,并在 QGraphicsScene 中添加 QGraphicsPathItem 或 QGraphicsPolygonItem 对象。在添加矩形框或者多边形图元后,可以通过点击顶点来编辑图元形状,同时也可以将鼠标移动到标签内部高亮区域使其更加突出和易于识别。

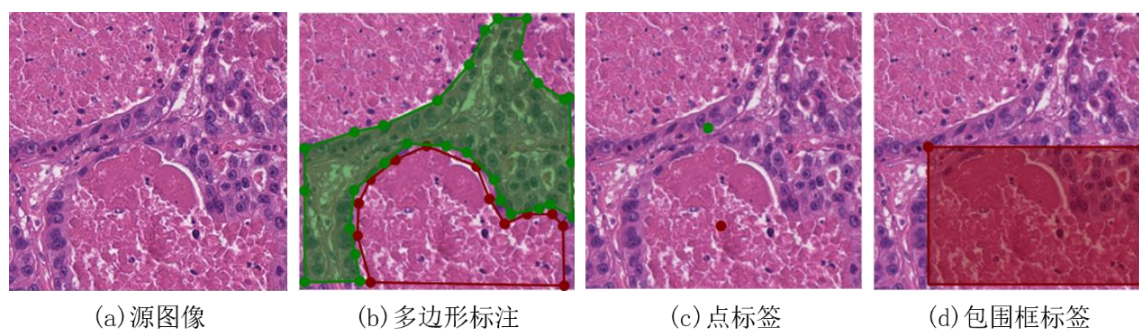


图 4-5 不同标签形式

§ 4.4 软件功能测试

§ 4.4.1 图像分割功能功能测试

用户点击工具栏或者菜单栏的深度学习预测按钮,向逻辑层发送信号,逻辑层收到信号后调用相应的 `inference` 函数,调用分割模型生成分割掩码,根据分割掩码得到近似的多边形形状并记录顶点坐标,最后调用 `createPoly` 函数生成多边形图元。图像分割功能测试所图所示:

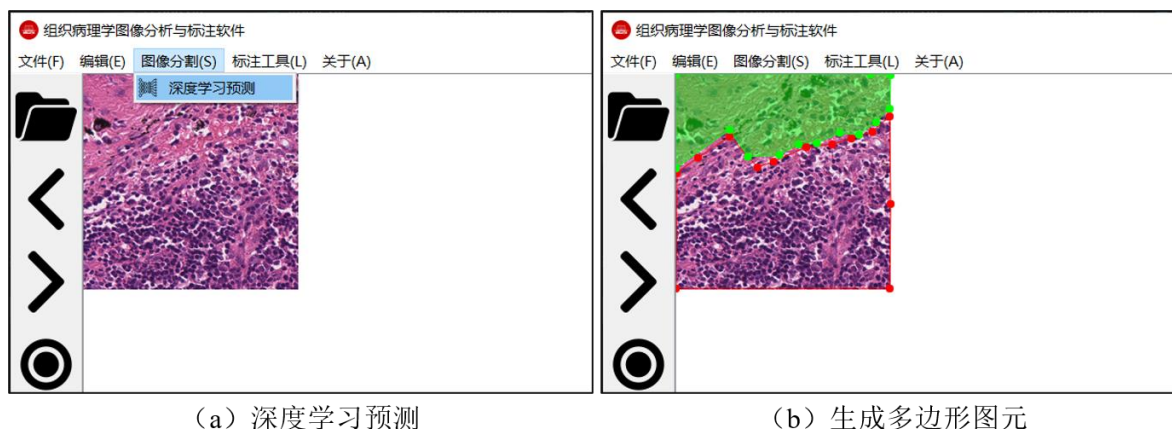


图 4-6 利用深度学习模型生成多边形图元

§ 4.4.2 图像标注功能测试

用户可以通过点击工具栏或者菜单栏的标注工具对图像进行标注，不同标注工具的选择应该是互斥的，所以用户在选择一种标注工具之后，其他标注工具应该变为不可用状态。标注功能测试截图如图所示：



图 4-7 标注功能测试（以多边形标注工具为例）

§ 4.4.3 图像几何变换功能测试

用户通过点击菜单栏的图像几何变换选择相应的变换函数对图像进行几何操作，软件能够支持顺时针旋转、逆时针旋转、垂直翻转、水平翻转以及对角线翻转等功能，用户可以通过组合这几种基本变换来达到更多的变换效果。几何变换功能如图所示：

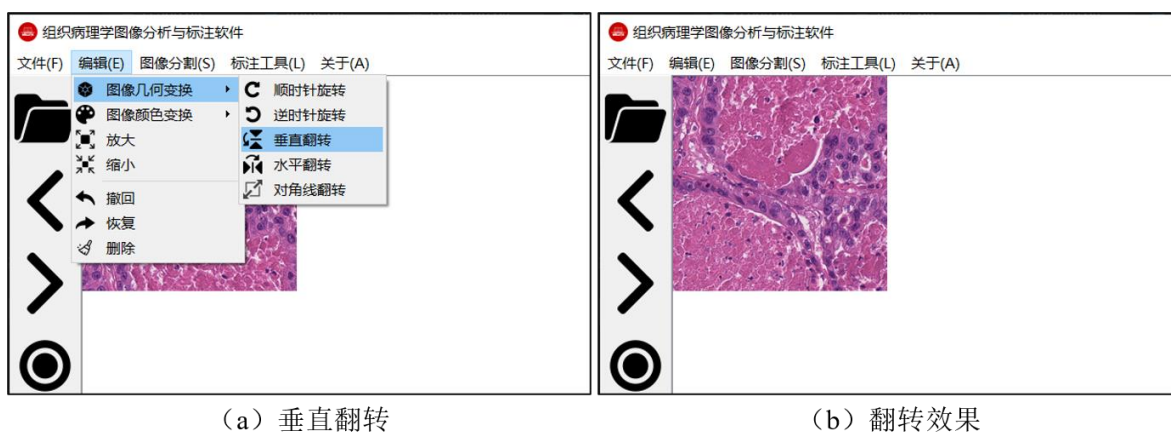


图 4-8 图像几何功能测试（以垂直翻转为例）

§ 4.4.4 颜色处理功能测试

用户通过点击菜单栏的颜色处理功能选择不同的颜色归一化方法对图像颜色进

行归一化处理，颜色处理功能如图所示：

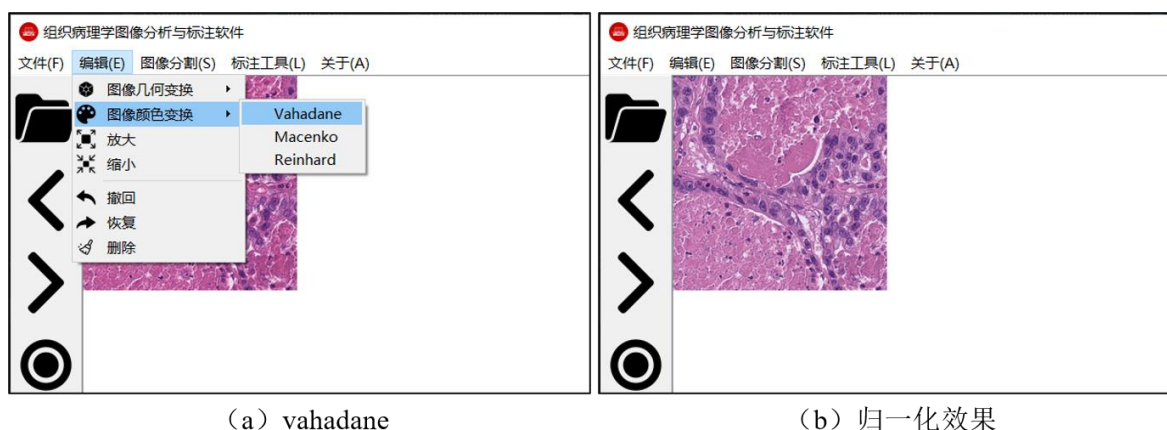


图 4-9 图像颜色归一化功能测试（以 vahdane 归一化）为例

§ 4.5 本章小结

在本章中，我们开发了一款集成了深度学习模型的组织病理学图像标注软件，通过将图像输入已收敛的深度模型进行前向传播，最终得到分割结果。深度学习模型的结果能够为标注人员生成初步的标注结果，帮助标注人员识别病理图像中不同组织的边界信息，标注人员通过微调算法结果来提高标注效率。但开发的软件也有一些不足，对于一些复杂的组织区域，还需要利用涂抹的标注方式能够获取更加精确的组织边界。

第五章 总结与展望

§ 5.1 本文总结

本文的研究内容是如何在仅有图像级标签的条件下更好地实现肺腺癌中四种组织的分割，这是一件极具挑战性的任务。在以往的研究中，研究人员通常只区分图像是否存在癌症组织。然而，除了区分癌症组织和正常组织以外，区分其他组织对于疾病的发展以及治疗也非常重要。本文利用图像级标签训练一个多标签分类网络，待网络收敛后，通过类激活映射技术初步得到不同组织的类激活图，最后在经过一系列后处理步骤细化类激活图得到不同组织的分割结果。这种方法存在两个问题，第一个问题是类激活图只能覆盖组织中最具有判别性的区域，而无法覆盖全部；第二个问题是在训练多标签分类网络时不得不使用二元交叉熵损失函数，该函数并不能对类激活图进行有效的监督。针对这两个问题，在前人研究的基础上，我们提出了一个全新的弱监督语义分割算法并生成更加准确的类激活图，最后通过实验结果和实验分析能够证明方法的有效性。

除此之外，我们认为图像级标签与像素级标签存在天然的监督信息差，为了训练具有卓越分割性能的网络，大量的标注数据是必需的。因此，我们开发了一款组织病理图像分析与标注软件，在软件中集成了深度学习模型，通过已收敛模型对图像进行分割并生成初步的标注结果，最后由标注人员微调生成最终的结果。集成深度学习模型的标注软件能够提高标注人员的标注效率和准确性。此外，软件中还支持了针对组织病理图像的图像处理功能，供病理医生分析使用。

§ 5.2 研究展望

随着病理切片成像技术和深度学习技术的发展，利用深度学习算法对组织病理学图像自动化分析，提高病理学诊断效率是当前的趋势。目前，深度学习技术在病理学图像上的研究比较有限，其主要原因是因为公开数据集的缺乏。针对这一问题，本文认为打通病理学与计算机科学的壁垒，把专家的从业经验和成功的计算机技术利用起来，推动智能医疗的发展。在本文中，我们提出的弱监督语义分割算法性能还不足以应用在临床中。本文认为未来还能够从以下几个方向进行研究：

(1) 可以应用基于 Transformer 的方法实现弱监督语义分割算法，相比较于卷积神经网络而言，基于 Transformer 的网络架构能够捕获更多的上下文信息，这是非常适合组织病理学图像的。当然，由于 Transformer 提取低维特征的能力有限，考虑结

合卷积神经网络与 Transfomer 来研究弱监督语义分割算法。

(2) 组织病理学图像的标注是一个标记成本较为昂贵, 需要病理学专家参与的任务。在这一背景下, 除了在标注软件中集成深度学习算法模型帮助病理学专家进行标注工作以外, 在软件中集成主动学习算法, 通过病理学专家与机器的协作能够得到性能更加优异的模型。主动学习算法的主要思想是基于当前模型的预测向病理学专家提供待标注的样本, 这些样本被认为对于模型的学习具有最高的潜在收益, 通过将标注好的样本加入模型训练后, 可以大幅提高网络参数。主动学习可以显著减少标记数据的需求, 从而提高学习效率和模型性能。

(3) 组织病理学图像分析是为了更好地了解患者情况, 从而确定治疗方案, 但不同患者, 其癌症组织的形态表现不同。因此, 在征得患者的同意后, 收集患者的既有病史、当前癌症情况、癌症发展情况、预后情况等信息, 建立完整的信息数据库, 推动病理学研究。与此同时, 癌症可以分为先天与后天发病, 可以通过收集患者的基因序列以及组织病理学图像, 建立从微观到宏观的自动化分析流程, 推动智能化医疗的发展。

参考文献

- [1] Zheng R, Zhang S, Zeng H, et al. Cancer incidence and mortality in China, 2016[J]. Journal of the National Cancer Center, 2022, 2(1): 1-9.
- [2] Siegel R L, Miller K D, Wagle N S, et al. Cancer statistics, 2023[J]. CA: A Cancer Journal for Clinicians, 2023, 73(1): 17-48.
- [3] Runggay H, Shield K, Charvat H, et al. Global burden of cancer in 2020 attributable to alcohol consumption: a population-based study[J]. The Lancet Oncology, 2021, 22(8): 1071-1080.
- [4] Wilson M L, Fleming K A, Kuti M A, et al. Access to pathology and laboratory medicine services: a crucial gap[J]. The Lancet, 2018, 391(10133): 1927-1938.
- [5] Horton S, Sullivan R, Flanigan J, et al. Delivering modern, high-quality, affordable pathology and laboratory medicine to low-income and middle-income countries: a call to action[J]. The Lancet, 2018, 391(10133): 1953-1964.
- [6] Sayed S, Cherniak W, Lawler M, et al. Improving pathology and laboratory medicine in low-income and middle-income countries: roadmap to solutions[J]. The Lancet, 2018, 391(10133): 1939-1952.
- [7] Long J, Shelhamer E, Darrell T. Fully Convolutional Networks for Semantic Segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015: 3431-3440.
- [8] Badrinarayanan V, Kendall A, Cipolla R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481-2495.
- [9] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation[C]//Navab N, Hornegger J, Wells W M, et al. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Cham: Springer International Publishing, 2015: 234-241.
- [10] Zhou Z, Rahman Siddiquee M M, Tajbakhsh N, et al. UNet++: A Nested U-Net Architecture for Medical Image Segmentation[C]//Stoyanov D, Taylor Z, Carneiro G, et al. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Cham: Springer International Publishing, 2018: 3-11.
- [11] Abdollahi A, Pradhan B, Alamri A. VNet: An End-to-End Fully Convolutional Neural Network for Road Extraction From High-Resolution Remote Sensing Data[J]. IEEE Access, 2020, 8: 179424-179436.
- [12] Chen L C, Papandreou G, Kokkinos I, et al. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834-848.
- [13] Chen L C, Papandreou G, Schroff F, et al. Rethinking Atrous Convolution for Semantic Image Segmentation[A]. arXiv, 2017.
- [14] Chen L C, Zhu Y, Papandreou G, et al. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation[C]//Proceedings of the European Conference on Computer Vision (ECCV). 2018: 801-818.
- [15] Dai J, He K, Sun J. BoxSup: Exploiting Bounding Boxes to Supervise Convolutional Networks for Semantic Segmentation[C]//Proceedings of the IEEE International Conference on Computer Vision. 2015: 1635-1643.
- [16] Zhang K, Zhuang X. CycleMix: A Holistic Strategy for Medical Image Segmentation From Scribble Supervision[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022: 11656-11665.
- [17] Shen W, Peng Z, Wang X, et al. A Survey on Label-Efficient Deep Image Segmentation: Bridging the Gap Between Weak Supervision and Dense Prediction[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023: 1-20.
- [18] Zhang F, Gu C, Zhang C, et al. Complementary Patch for Weakly Supervised Semantic Segmentation[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 7242-7251.
- [19] Shmatko A, Ghaffari Laleh N, Gerstung M, et al. Artificial intelligence in histopathology: enhancing cancer research and clinical oncology[J]. Nature Cancer, 2022, 3(9): 1026-1038.

- [20] 宋杰, 肖亮, 练智超, 等. 基于深度学习的数字病理图像分割综述与展望[J]. 软件学报, 2021, 32(5): 1427-1460.
- [21] 宋国利, 陈杰. 病理图像分析的深度学习研究方法研究综述[J]. 中国科学基金, 2022, 36(2): 225-234.
- [22] Chen H, Qi X, Yu L, et al. DCAN: Deep Contour-Aware Networks for Accurate Gland Segmentation[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016: 2487-2496.
- [23] Qu H, Yan Z, Riedlinger G M, et al. Improving Nuclei/Gland Instance Segmentation in Histopathology Images by Full Resolution Neural Network and Spatial Constrained Loss[C]//Shen D, Liu T, Peters T M, et al. Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. Cham: Springer International Publishing, 2019: 378-386.
- [24] Tokunaga H, Teramoto Y, Yoshizawa A, et al. Adaptive Weighting Multi-Field-Of-View CNN for Semantic Segmentation in Pathology[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 12597-12606.
- [25] Xu Y, Zhu J Y, Chang E I C, et al. Weakly supervised histopathology cancer image segmentation and classification[J]. Medical Image Analysis, 2014, 18(3): 591-604.
- [26] Jia Z, Huang X, Chang E I C, et al. Constrained Deep Weak Supervision for Histopathology Image Segmentation[J]. IEEE Transactions on Medical Imaging, 2017, 36(11): 2376-2388.
- [27] Lin H, Chen H, Dou Q, et al. ScanNet: A Fast and Dense Scanning Framework for Metastatic Breast Cancer Detection from Whole-Slide Image[C]//2018 IEEE Winter Conference on Applications of Computer Vision (WACV). 2018: 539-546.
- [28] Lee J, Choi J, Mok J, et al. Reducing Information Bottleneck for Weakly Supervised Semantic Segmentation[C]//Advances in Neural Information Processing Systems: 卷 34. Curran Associates, Inc., 2021: 27408-27421.
- [29] Jiang P T, Hou Q, Cao Y, et al. Integral Object Mining via Online Attention Accumulation[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). 2019: 2070-2079.
- [30] Sun K, Shi H, Zhang Z, et al. ECS-Net: Improving Weakly Supervised Semantic Segmentation by Using Connections Between Class Activation Maps[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 7283-7292.
- [31] Wei Y, Xiao H, Shi H, et al. Revisiting Dilated Convolution: A Simple Approach for Weakly- and Semi-Supervised Semantic Segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 7268-7277.
- [32] Zhang X, Wei Y, Feng J, et al. Adversarial Complementary Learning for Weakly Supervised Object Localization[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 1325-1334.
- [33] Lee J, Kim E, Yoon S. Anti-Adversarially Manipulated Attributions for Weakly and Semi-Supervised Semantic Segmentation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 4071-4080.
- [34] Chan L, Hosseini M S, Rowsell C, et al. HistoSegNet: Semantic Segmentation of Histological Tissue Type in Whole Slide Images[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019: 10662-10671.
- [35] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition[A]. arXiv, 2015.
- [36] Zhang J, Hua Z, Yan K, et al. Joint fully convolutional and graph convolutional networks for weakly-supervised segmentation of pathology images[J]. Medical Image Analysis, 2021, 73: 102183.
- [37] Bera K, Schalper K A, Rimm D L, et al. Artificial intelligence in digital pathology — new tools for diagnosis and precision oncology[J]. Nature Reviews Clinical Oncology, 2019, 16(11): 703-715.
- [38] Sirinukunwattana K, Pluim J P W, Chen H, et al. Gland segmentation in colon histology images: The glas challenge contest[J]. Medical Image Analysis, 2017, 35: 489-502.
- [39] Xu Y, Jia Z, Wang L B, et al. Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features[J]. BMC Bioinformatics, 2017, 18(1): 281.
- [40] Han C, Lin J, Mai J, et al. Multi-layer pseudo-supervision for histopathology tissue semantic segmentation using patch-level classification labels[J]. Medical Image Analysis, 2022, 80: 102487.
- [41] Srivastava N, Hinton G, Krizhevsky A, et al. Dropout: A Simple Way to Prevent Neural Networks from Overfitting[J]. Journal of Machine Learning Research, 2014, 15(56): 1929-1958.
- [42] Wu Z, Shen C, van den Hengel A. Wider or Deeper: Revisiting the ResNet Model for Visual

- Recognition[J]. Pattern Recognition, 2019, 90: 119-133.
- [43] Zhou B, Khosla A, Lapedriza A, et al. Learning Deep Features for Discriminative Localization[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 2921-2929.
- [44] He K, Zhang X, Ren S, et al. Identity Mappings in Deep Residual Networks[C]//Leibe B, Matas J, Sebe N, et al. Computer Vision – ECCV 2016. Cham: Springer International Publishing, 2016: 630-645.
- [45] Xie S, Girshick R, Dollar P, et al. Aggregated Residual Transformations for Deep Neural Networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 1492-1500.
- [46] Guo M H, Xu T X, Liu J J, et al. Attention mechanisms in computer vision: A survey[J]. Computational Visual Media, 2022, 8(3): 331-368.
- [47] Hu J, Shen L, Sun G. Squeeze-and-Excitation Networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 7132-7141.
- [48] Woo S, Park J, Lee J Y, et al. CBAM: Convolutional Block Attention Module[C]//Proceedings of the European Conference on Computer Vision (ECCV). 2018: 3-19.
- [49] AbdulJabbar K, Raza S E A, Rosenthal R, et al. Geospatial immune variability illuminates differential evolution of lung adenocarcinoma[J]. Nature Medicine, 2020, 26(7): 1054-1062.
- [50] Chen Z, Wang T, Wu X, et al. Class Re-Activation Maps for Weakly-Supervised Semantic Segmentation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022: 969-978.
- [51] Chang Y T, Wang Q, Hung W C, et al. Weakly-Supervised Semantic Segmentation via Sub-Category Exploration[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 8991-9000.
- [52] Chattopadhyay A, Sarkar A, Howlader P, et al. Grad-CAM++: Generalized Gradient-Based Visual Explanations for Deep Convolutional Networks[C]//2018 IEEE Winter Conference on Applications of Computer Vision (WACV). 2018: 839-847.
- [53] Jiang P T, Hou Q, Cao Y, et al. Integral Object Mining via Online Attention Accumulation[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019: 2070-2079.
- [54] Kweon H, Yoon S H, Kim H, et al. Unlocking the Potential of Ordinary Classifier: Class-Specific Adversarial Erasing Framework for Weakly Supervised Semantic Segmentation[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 6994-7003.
- [55] Reinhard E, Adhikhmin M, Gooch B, et al. Color transfer between images[J]. IEEE Computer Graphics and Applications, 2001, 21(5): 34-41.
- [56] Vahadane A, Peng T, Sethi A, et al. Structure-Preserving Color Normalization and Sparse Stain Separation for Histological Images[J]. IEEE Transactions on Medical Imaging, 2016, 35(8): 1962-1971.
- [57] Macenko M, Niethammer M, Marron J S, et al. A method for normalizing histology slides for quantitative analysis[C]//2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. 2009: 1107-1110.