

硕士学位论文

胶质瘤分子分型的非侵入式预测方法
Non-invasive Prediction Method for Molecular
Subtype of Glioma

学科专业： 工程硕士

专业领域： 控制工程

作者姓名： 徐勇

指导教师： 谢斌 副教授

中南大学
2021年5月

中图分类号 TP391.4

学校代码 10533

UDC 004.8

学位类别 专业学位

硕士学位论文

胶质瘤分子分型的非侵入式预测方法

**Non-invasive Prediction Method for Molecular
Subtype of Glioma**

作者姓名：徐勇

学科专业：工程硕士

专业领域：控制工程

研究方向：模式识别与智能系统

二级培养单位：自动化学院

指导教师：谢斌

论文答辩日期 2021年5月27日 答辩委员会主席 

中南大学
2021年5月

学位论文原创性声明

本人郑重声明，所提交的学位论文是本人在指导教师指导下进行的研究工作及取得的研究成果。尽我所知，除了论文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得中南大学或其他教育机构的学位或证书而使用过的材料。与我共同工作的同志对本研究所作的贡献均已在论文中作了明确的说明。

申请学位论文与资料若有不实之处，本人承担一切相关责任。

作者签名： 徐勇 日期： 2021 年 6 月 3 日

学位论文授权使用授权书

本学位论文作者和指导教师完全了解中南大学有关保留、使用学位论文的规定：即学校有权保留并向国家有关部门或机构送交学位论文的复印件和电子版；本人允许本学位论文被查阅和借阅；学校可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用复印、缩印或其它手段保存和汇编本学位论文。

保密论文待解密后适应本声明。

作者签名： 徐勇

日期： 2021 年 6 月 3 日

导师签名： 谢红

日期： 2021 年 6 月 3 日

胶质瘤分子分型的非侵入式预测方法

摘要: 脑胶质瘤是脑肿瘤中发病率最高的原发性恶性肿瘤。通过对肿瘤切片进行病理学检测可以确定肿瘤切片的分子标记物,这可以在分子级别确定胶质瘤的分型(分子分型)。通过分子分型可以更准确地判断胶质瘤临床特性和预后,选择个性化的治疗方案。然而一方面获取富含肿瘤的组织样本以准确测定分子标记物状态是不容易的,另一方面在手术前更早确定分子分型将具有更大的临床指导意义。因此本文针对**基于核磁共振图像(MRI)和病人基本信息等非侵入信息预测胶质瘤分子分型**任务展开了深入的研究,具体的工作内容如下:

(1)对基于非侵入信息预测胶质瘤分子分型的研究现状进行了分析,考虑分子标记物预测方法通常需要人工勾画感兴趣区域(ROI)的问题,将预测分子标记物转化为**ROI分割算法和基于多模态信息的分类算法**,为基于非侵入信息预测胶质瘤分子分型提供了可行的标准流程。

(2)探究了模型的输入图像方式(是否需要勾画ROI)、分类网络骨架和多模态信息融合位置对异柠檬酸脱氢酶(IDH)突变状态预测准确率的影响,证明了**ROI作为附加输入对于预测分子分型的重要性**。设计了多种多模态融合网络用于预测IDH突变状态,最佳模型能取得受试者工作曲线下面积(AUC)为0.87的结果,这比主流的融合方式性能更好。

(3)为了解决胶质瘤分子标记物预测方法通常需要人工勾画ROI的问题,基于临床胶质瘤分割数据集提出了一种名为NASBraNet的语义分割网络架构搜索框架,目的是**实现快速准确的自动ROI分割**。本章节针对临床胶质瘤MRI影像分割需要兼顾准确率和速度的需求,通过多分支选择机制能搜索到较高的分割准确率的教师网络。通过引入扩展系数搜索和延迟正则化,实现了高效率学生网络的架构搜索。并通过教师学生网络共同搜索和知识蒸馏将教师网络的知识有效地迁移到了学生网络上,最终得到了应用于临床胶质瘤MRI影像任务的高效分割网络,并在综合性能上优于主流算法。

(4)为了解决胶质瘤分子标记物预测方法通常需要人工勾画ROI的问题,一个新的多模态深度学习网络(KiNet)被提出用于基于多模态MRI和异构元数据(年龄、性别)等信息非侵入式地预测胶质瘤的Ki67指数。网络使用两个附加分支对两个模态MRI图像进行特征提取,在主分支中

融合多层次附加分支特征和异构元数据,并引入了多尺度损失监督网络训练过程。网络首次实现了无需手工标注 **ROI** 的非侵入式预测胶质瘤的 **Ki67 指数**,取得了 AUC 0.78 的结果。通过该方法,医生有希望在手术前获得胶质瘤的 Ki67 指数,从而制定更有效的治疗方案。

包含图 23 幅,表 14 个,参考文献 99 篇

关键词: 胶质瘤; 分子分型; 影像基因组学; 病灶分割; 多模态学习
分类号: TP391.4

Non-invasive Prediction Method for Molecular Subtype of Glioma

Abstract: Glioma is the primary malignant tumor with the highest incidence among brain tumors. By detecting the molecular markers of tumor slices, the molecular subtype of glioma can be determined, so that the characteristics and clinical prognosis of glioma can be judged more accurately, and personalized plans can be selected according to different molecular classifications. However, it is not easy to obtain tumor-rich tissue samples to accurately determine the status of molecular markers, and it is of greater clinical significance to determine the molecular classification earlier before surgery. Therefore, this article has carried out an in-depth study on predicting the molecular typing task of glioma based on heterogeneous non-invasive information such as MRI and basic patient information. The specific work content is as follows:

1. In this paper, the research status of predicting glioma molecular typing based on non-invasive information is analyzed. Considering that the prediction method of molecular markers usually needs to manually delineate the region of interest (ROI), the prediction molecular markers are transformed into ROI segmentation algorithm and classification algorithm based on multimodal information, which provides a feasible standard process for predicting glioma molecular typing based on non-invasive information .

2. The influence of the input image mode (whether it is necessary to outline ROI), classification network skeleton and multimodal information fusion position of the model on the prediction accuracy of mutation state of isocitrate dehydrogenase (IDH) was explored, and the importance of ROI as additional input for predicting molecular typing is proved. A variety of multimodal fusion networks are designed to predict the mutation state of IDH. The best model can achieve the result that the area under the receiver operating curve (AUC) is 0.87, which is better than the mainstream fusion methods.

3. In order to solve the problem that the prediction method of glioma molecular markers usually needs to manually delineate ROI, a semantic segmentation network architecture search framework named nasbranet is proposed based on clinical glioma segmentation data set. The purpose is to

achieve fast and accurate automatic ROI segmentation. In this chapter, the accuracy and speed of clinical glioma MRI image segmentation need to be considered. Through the multi branch selection mechanism, the teacher network with higher segmentation accuracy can be searched. By introducing extended coefficient search and delay regularization, the efficient architecture search of student network is realized. The knowledge of the teacher network is effectively transferred to the student network through the joint search and knowledge distillation of the teacher student network. Finally, an efficient segmentation network is obtained, which is better than the mainstream algorithm in comprehensive performance.

4. In order to solve the problem that the prediction methods of glioma molecular markers usually need to draw the ROI manually, a new multimodal deep learning network (KiNet) is proposed to predict the Ki67 index of glioma noninvasively based on multimodal MRI and heterogeneous metadata (age, gender). The network uses two additional branches to extract the features of two modal MRI images. In the main branch, multi-level additional branch features and heterogeneous metadata are fused, and multi-scale loss monitoring network training process is introduced. For the first time, the network realized the non-invasive prediction of Ki67 index of glioma without ROI, and achieved the result of AUC 0.78. Through this method, doctors hope to obtain the Ki67 index of glioma before surgery, so as to develop a more effective treatment plan.

Figure 23, Table 14, Reference 99.

Keywords: Glioma; Molecular subtype; Radiogenomics; Lesion segmentation; Multimodal heterogeneous learning

Classification: TP391.4

目 录

第 1 章 绪论.....	1
1.1 研究背景及意义.....	1
1.2 国内外研究现状.....	3
1.2.1 脑肿瘤影像分割方法研究进展.....	4
1.2.2 胶质瘤分子分型预测方法研究进展.....	5
1.3 本文主要研究内容和组织结构.....	7
第 2 章 胶质瘤分子分型预测算法理论基础.....	9
2.1 卷积神经网络.....	9
2.2 主流卷积神经网络模型.....	9
2.2.1 AlexNet 模型.....	9
2.2.2 ResNet 模型.....	10
2.2.3 全卷积网络.....	11
2.2.4 U-Net 模型.....	11
2.3 神经网络架构搜索.....	12
2.4 知识蒸馏.....	13
2.5 多模态深度神经网络模型.....	14
2.6 本章小结.....	15
第 3 章 胶质瘤 IDH 突变状态预测.....	17
3.1 引言.....	17
3.2 数据集预处理.....	17
3.3 基于多模态图像的 IDH 突变状态预测算法.....	18
3.3.1 网络骨架和输入模式选择.....	19
3.3.2 多模态分类模型.....	19
3.4 实验结果与讨论.....	20
3.4.1 算法评估准则.....	20
3.4.2 不同网络骨架和输入模式实验结果对比.....	21
3.4.3 不同多模态融合方式实验结果对比.....	22
3.4.4 分析与讨论.....	22

3.5 本章小结	24
第4章 脑部MRI影像中分子分型ROI分割.....	25
4.1 引言	25
4.2 可微分架构搜索	26
4.3 ROI分割网络架构搜索.....	26
4.3.1 搜索空间.....	26
4.3.2 搜索空间松弛连续化.....	29
4.3.3 细粒度的延迟正则化优化.....	30
4.3.4 教师/学生网络共同搜索.....	31
4.3.5 网络结构参数更新过程.....	31
4.3.6 离散架构导出.....	32
4.4 实验结果与讨论	33
4.4.1 数据集介绍.....	33
4.4.2 网络架构搜索结果.....	33
4.4.3 算法性能评估.....	34
4.4.4 分析与讨论.....	35
4.5 本章小结	36
第5章 胶质瘤KI67指数预测.....	37
5.1 引言	37
5.2 数据集预处理	37
5.3 多任务多模态深度神经网络模型	38
5.3.1 多任务多模态深度神经网络模型结构.....	39
5.3.2 损失函数设计.....	40
5.4 实验结果与讨论	41
5.4.1 算法评估准则.....	41
5.4.2 与主流算法的实验结果对比.....	41
5.4.3 消融实验.....	42
5.4.4 分析与讨论.....	42
5.5 本章小结	43
第6章 总结与展望.....	45
6.1 论文总结	45
6.2 研究展望	46

参考文献.....	47
攻读学位期间主要的研究成果.....	55
致谢.....	56

第 1 章 绪论

1.1 研究背景及意义

本文研究的脑胶质瘤是脑肿瘤中发病率最高的原发性恶性肿瘤。它起源于神经胶质细胞，是典型恶性肿瘤，具有弥散性、浸润性、边界模糊等特征^[1]。

世界卫生组织（WHO）根据病理特征将胶质瘤分为 4 级，也是目前临床病理诊断的标准^[2]。根据中国胶质瘤联合组织（Chinese Glioma Cooperation Group, CGCG）的统计资料，级别最高的 WHO IV 级胶质瘤中位生存期只有 14 个月，而 WHO II 级胶质瘤中位生存期为 7 年左右。

但是，经典的组织病理学主观性较强，不能客观、系统、准确反映肿瘤组织的基因学背景和生物学特征。随着人类基因组计划的完成及高通量技术的发展，肿瘤形态学结合基因特征的分子病理、分子分型基础上的胶质瘤个体化治疗等概念，已逐渐被神经外科学者们认同并应用。通过检测分子标志物可以确定胶质瘤的分子分型，能够更准确地判断胶质瘤患者临床预后，并且对组织学难以明确诊断和分级的肿瘤提供鉴别依据^[2]。此外，这些新近发现的分子变异有可能成为未来治疗的新靶点（表 1-1）。

表 1-1 胶质瘤中分子标记物的生物学功能及诊疗价值

分子标志物	生物学功能	诊断价值	预后价值
IDH1/2 突变	与基因损伤修复、表观遗传学调控及细胞免疫的行为密切相关	是区分弥漫性胶质瘤和胶质细胞增生的标志物	存在 IDH1/2 突变的肿瘤患者预后好
1p/19q 联合缺失	不明。候选基因 CIC 和 FUBP1 正在检测中	是少突胶质细胞瘤诊断性标志物	有联合性缺失的肿瘤患者预后较好
TERT 启动子区突变	增加 TERT 的表达，与端粒酶的活性相关	与原发胶质母细胞瘤和少突胶质细胞瘤相关	有争议的，与其他分子标志物有关
ATRX 突变	与端粒替代延长机制有关	是星形细胞瘤诊断性标志物	存在 ATRX 突变的肿瘤患者预后较好
MGMT 启动子甲基化	干扰 DNA 修复，与 IDH1/2 突变肿瘤中的 G-CIMP 基因相关	无	对于高级别星形细胞瘤胶质瘤患者放/化疗有好的疗效
EGFR 扩增	对细胞的生长、增殖和分化等生理过程有重要作用	与胶母细胞瘤密切相关	大于 60 岁的胶母细胞瘤病人预后差
Ki-67 指数	细胞增殖相关的核抗原	判断恶性程度、确定分级	对于低级别弥漫性胶质瘤可用于预测预后
miR-21	直接调控下游多个靶基因	能够通过脑脊液等检测	与患者不良预后相关

其中 Ki67 指数在临床上可用于衡量肿瘤的增殖程度、预测病人生存期和复发情况^[3]。一般来说, 高的 Ki67 指数对应着更短的生存期和更高的复发率^[4, 5], 因此, 不同的 Ki67 指数往往对应着不同的治疗方案, 若能在手术前得到 Ki67 指数将具有很大的指导意义^[6]。

临床上常使用 T2 Flair、T1、T1c、T2 四个模态的磁共振成像 (MRI) 影像作为脑肿瘤诊断的依据, 图 1-1 展示了某患者四种模态的临床图像。脑脊液在 T2 Flair 模态图像中是低信号组织, 在图像中的表现是黑色。T2 Flair 模态对病灶更敏感, 因脑脊液信号被抑制所以能够检查出被脑脊液掩盖的病灶, 比如脑出血。相较于其他模态图像, 该模态图像除脑肿瘤区域比较明亮而其他区域都比较模糊, 如图 1-1(a), 其中脑脊液在图像中呈低信号为黑色。T1 模态图像主要用来显示组织的解剖结构, 如图 1-1(b)。T1c 模态图像突出显示了病灶的边缘信息, 与周围组织形成鲜明对比, 如图 1-1(c), 这样对脑组织结构的显示也有一个比较明显的对比度。T2 中病变信号突出, 适用于观察病变, 即肿瘤整体的灰度都与正常组织差异明显, 如图 1-1(d)。

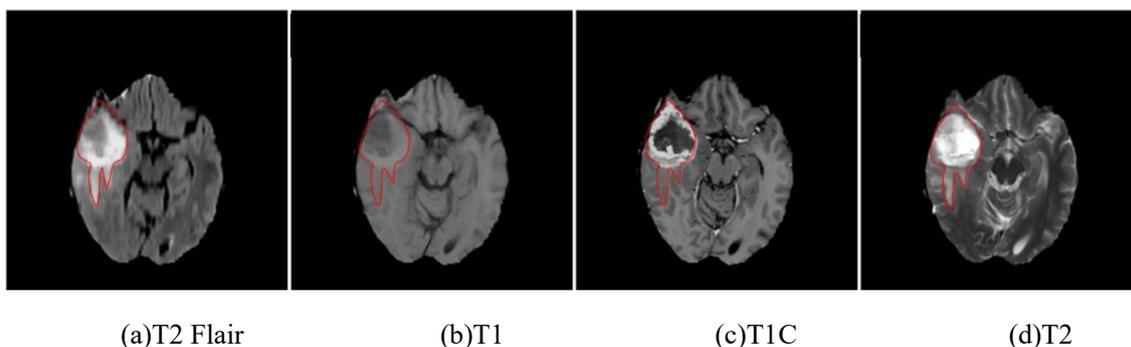


图 1-1 四种模态的脑肿瘤图像

图 1-2 展示了收集于中南大学湘雅医院神经外科的不同 Ki67 指数的患者的多模态 MRI 图像, 专业医师肉眼对比 T1c 和 T2 Flair 两个模态的 MRI 图像可以基本确定肿瘤大小、位置等, 基于深度神经网络的肿瘤分割算法可能更加高效, 但是仅基于肿瘤区域难以分辨不同 Ki67 指数的肿瘤在 MRI 图像上的差异。Ki67 指数的确定需要通过肿瘤组织取样, 使用 Ki67 抗体对其进行染色, 并计算被染色细胞的比例^[7]。然而, 获取富含肿瘤的组织样本以计算 Ki67 指数可能是一项挑战。癌症基因组图谱(TCGA)的一份报告表明, 只有 35%的活检样本含有足够的肿瘤成分, 可以进行准确的分子特征鉴定^[8]。开发一个非侵入式的 Ki67 指数鉴定方法将对患者的护理更加有利。

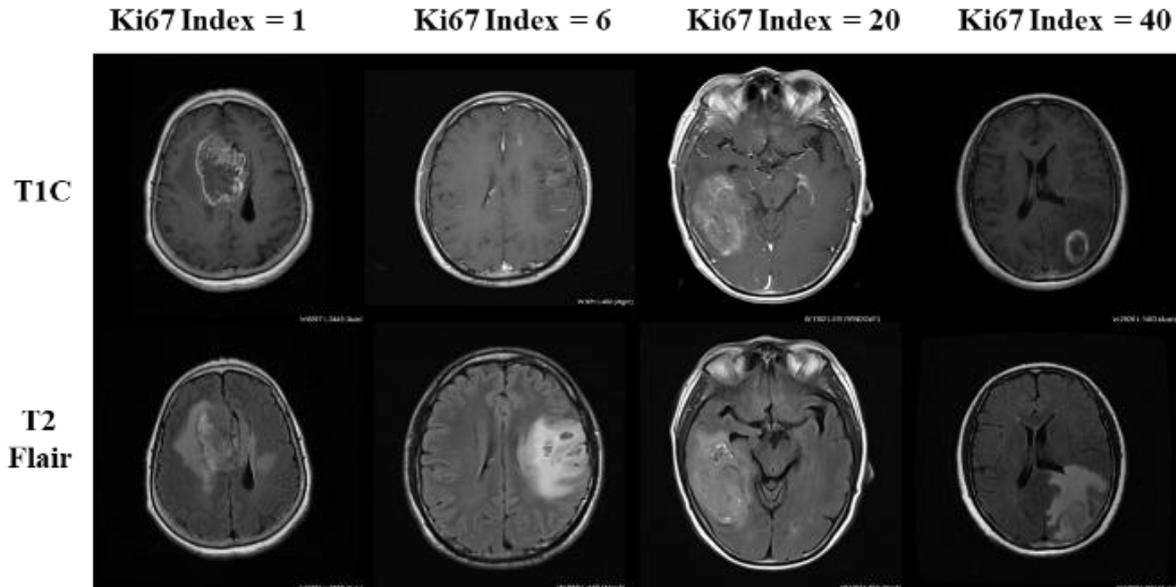


图 1-2 不同的 Ki67 指数的 T1C 和 T2 Flair 模态的临床 MRI 影像。

本项目提出一项基于真实临床数据的研究计划，通过异构机器学习技术对多模态脑部 MRI 图像以及病人年龄，性别元数据等多类数据进行融合，实现非侵入式地对神经胶质瘤的病灶的自动定位，并得出胶质瘤的分子标记物情况，从而可以术前预测胶质瘤的分子分型、制定最合理的诊疗方案。

1.2 国内外研究现状

在过去的数十年中，随着模式识别技术的发展和数据集大小的增加，医学图像分析领域的研究呈现出指数级增长。

2012 年，Lambin 等人^[9]提出了影像组学的概念。他们将影像组学解释为从医学图像中提取高通量的影像特征确定基因水平的特性，并提出了影像组学研究的标准流程。首先获取高质量和标准化的成像，肿瘤区域可以由有经验的放射科医师或放射肿瘤科医生来分割，但为了在临床实际中便于应用，通过算法自动分割得到肿瘤区域是一个更优的选择。随后从先前定义的肿瘤区域提取高通量的定量成像特征。通过就是分析所选特征数据来获取关于肿瘤患者的病理级别信息。最终目标是通过将成像特征获得肿瘤等疾病分子级别的特征，从而更好的帮助诊断治疗。

考虑到现有关于分子特征预测算法的研究，大多都需要勾画肿瘤区域作为先验知识。为了在无专家知识的情况下非侵入式地预测胶质瘤分子分型，本文将基于多模态数据的胶质瘤分子分型预测方法拆分为基于网络架构搜索的脑部 MRI 影像分割算法和基于多模态神经网络的分类算法。因此，本节将会从 MRI 脑肿瘤分割方法和胶质瘤分子分型预测进展两个方面介绍国内外相关的研究进展。

1.2.1 脑肿瘤影像分割方法研究进展

磁共振成像(MRI)处理,特别是区分肿瘤组织和脑中其他部分的语义分割在胶质瘤的诊断和治疗中起着重要的作用^[10]。对于神经外科医生来说,在手术前尽可能精确地对肿瘤区域进行标注是一项重要的任务,因为一个理想的脑肿瘤分割不仅可以保留足够的健康组织,而且还可以防止肿瘤的复发^[11]。在过去的几年里,来自学术界和工业界的科学家一直在不遗余力地探索计算机辅助解决方案,让医生从费力和耗时的标注工作中解脱出来。

深度学习的盛行推动了语义分割方法的发展。全卷积网络(FCN)开始了利用卷积神经网络(CNN)进行密集图像预测的尝试^[12]。受到 FCN 类似思想的启发,许多前沿的 CNN 在分割任务中找到了自己的方法,例如完全卷积密集网(FCDN)^[13]和 DeepLab^[14]。2015 年,在 IEEE 国际生物医学成像研讨会(ISBI)挑战赛上,针对果蝇细胞跟踪任务首次提出了 U-Net^[15]。在那之后不久便因在二维(2D)和三维(3D)医学图像数据集上展示的有效性和效率而闻名^[16]。

对于胶质瘤的分割,U-Net 也是最常用的架构之一。例如,Kamnitsas 等人制作了 FCN、U-Net 和 DeepMedic^[17]的组合模型,其中 DeepMedic 是 DeepLab 的变种,用于脑部病变分割,并在 2017 年多模态脑瘤分割挑战(BraTS)中获得一等奖^[18]。2018 年 BraTS 的获胜者 Myronenko 受益于变分自动编码器(VAE)和 U-Net 结构的结合^[19]。Isensee 等人。提出了这样的论点,即一个训练好的 UNet 就足以完成分割任务,而不需要额外的附件^[20]。然而,即使是这样,考虑到现成操作模块的多样性可能会导致大量的候选结构,为 U-Net 中的下采样和上采样路径挑选正确的构建模块可能仍然是一项艰巨的工作。

作为自动机器学习(AutoML)的一个分支,神经网络架构搜索(NAS)专门致力于从众多的候选体系结构中自动发现最优的网络结构。NAS 的蓬勃发展归功于 Zoph 和 Le,他们首先提出了以强化学习的方式训练递归神经网络(RNN)的想法,以确定卷积模块的自变量和滤波器数量等超参数^[21]。在 NASNet 中,Zoph 等人。提出了基于小尺度数据集的两步搜索策略,即首先寻找两种称为基本单元的“单元”,然后将这些“单元”堆积在一起,形成较大数据集的最终解^[22]。NASNet 的一个主要问题是巨大的成本,这通常相当于数百个 GPU 在一起工作几天。如何在降低时间和存储空间消耗的同时保持高效率一直是 NAS 领域的研究热点。渐进式 NAS(PNAS)通过更小的搜索空间、启发式搜索策略和经验代理评估器获得比 NASNet 快 7 倍的速度^[23]。高效的 NAS(ENAS)通过搜索空间中基本模块之间的权重共享,极大地提高了 NASNet 的效率^[24]。为了进一步削减内存支出,研究人员将目光投向了超网络学习控制器。One-shot NAS^[25]和无代理 NAS^[26]去掉了强化或进化学习控制器,它们为每个操作单元安排一个概率,并直接在目标数据集上对它们进行一次性训练,而不是在较小的数据集上进

行训练。可微分结构搜索(DARTS)引入了从离散搜索到连续搜索的数学松弛,并用梯度下降更新过程代替了控制器学习^[27, 28]。

NAS 在分类任务上的成功也刺激了语义分割场景的努力。陈等人通过递归搜索编码器和解码器块^[29],证明了该方法的可行性。刘等人设计了 Auto-DeepLab,并提出了分层搜索空间的概念,它覆盖了“单元”级和骨干网络级搜索空间^[30]。在医学图像分割方面,NAS-UNet 分别寻找最优的下采样和上采样单元,构建 UNet 结构,并在包括前列腺 MRI、肝脏计算机断层扫描(CT)和神经超声图像在内的二维医学图像数据集上进行了测试^[31]。朱等人开发了一种基于 DARTS 的可微分 NAS UNet,用于在 2D 和 3D CT 数据集上分割肺和胰腺^[32]。

尽管取得了一些初步的成功,但在设计细分模型方面成功的人员领域专业知识似乎尚未完全集成到医学图像分割 NAS 框架中。例如,人为设计的实时分割架构^[33,34]通常利用具有适当深度,宽度,运算符和下采样率的多分辨率分支,并发现它们对网络的高准确率做出重要贡献。此外,现有的医学图像 NAS 分割网络往往对于高精度追求较高,而忽略了临床应用中的低延时和低资源消耗要求。

本文根据临床胶质瘤分割数据集提出了一种名为 NASBraNet 的 NAS 框架,目的是实现快速的推理速度和高的准确率。本文设计了一个特殊的搜索空间,能够支持对不同分辨率的多个分支进行优化。这些搜索的分支被自动整合以用于最终预测。为了进一步在准确性与延迟之间取得平衡,并避免朝任一指标崩溃(例如,良好的延迟但准确性较差),本文使用了一种精细的延迟正则化,以更好的权衡延迟和准确性。而且,为了更完整地兼顾准确率和性能,本文的 NAS 框架同时搜索复杂的教师网络和轻量级的学生网络结构,通过知识蒸馏得到高推理速度高准确率的分割网络。

1.2.2 胶质瘤分子分型预测方法研究进展

放射基因组学有两种含义。这个术语最早的用法是描述放射治疗的敏感性或有效性,以及正在治疗的肿瘤的基因组特性的影响。另一个定义,也是本文关注的,是使用放射学影像来预测组织的基因组特性。本文将使用“分子”而不是“基因组”的说法,因为许多重要的发现从技术上讲不是基因组,而可能是表现基因组(例如,基因组一部分的甲基化,这通常会使该基因失活)、染色体或分子水平,但实际上并不代表基因构成的变化。已经有研究认识到成像结果和分子特性之间的趋势,例如少突胶质瘤的 1p/19q 染色体共缺失常常导致肿瘤边缘模糊和 T2 信号不均。虽然这些研究可以产生统计学上显著的 P 值^[35],但对特定病例的基因组特性的前瞻性预测仍然不够准确,不能用于临床应用。本文还注意到,大多数用于放射基因组预测的技术可以应用于其他成像任务,包括疾病的分类(良性、恶性、炎症性、传染性等)或治疗反应的评估。共同点是,这些算法能够学习图像中与类别相关的模式。

表 1-2 胶质瘤放射基因组学部分相关研究

作者	标记物	数据集	肿瘤类型	结果
Jakola ^[36]	IDH	25, FLAIR	2~3 级弥漫性胶质瘤	AUC 0.905
Lj ^[37]	MGMT	193, T1 FLAIR	弥漫性胶质瘤	AUC 0.88
Park ^[38]	IDH, 1p/19q	93, T2, ADC, FA	二级胶质瘤	AUC 0.807
Chang ^[39]	IDH	496, FLAIR, T2, T1	弥漫性胶质瘤	AUC 0.95
Han ^[40]	MGMT	26, T1, T2, FLAIR	胶质母细胞瘤	AUC 0.61
Korfiatis ^[41]	MGMT	155, T2	胶质母细胞瘤	Accuracy 0.95 ± 0.04
Korfiatis ^[42]	ATRX	135, FLAIR	胶质母细胞瘤	ResNet f1 score 0.91 SVM f1 score 0.73
Chang ^[43]	IDH, 1p19q, MGMT	259, T1, FLAIR	弥漫性胶质瘤	Accuracy 0.94 IDH, 0.92 1p19q, 0.83 MGMT
Eichinger ^[44]	IDH	79	2~3 级弥漫性胶质瘤	AUC 0.952
Bisdas ^[45]	IDH	37	2~3 级弥漫性胶质瘤	AUC 0.88
Zhang ^[46]	IDH, TP53	103	2~3 级弥漫性胶质瘤	IDH AUC 0.792, TP53 AUC 0.869
Ren ^[47]	IDH, ATRX	57	弥漫性胶质瘤	IDH AUC 0.93
Arita ^[48]	IDH/TERT	169	2~3 级弥漫性胶质瘤	Accuracy 0.82

表 1-2 中的缩写含义如下，IDH，异柠檬酸脱氢酶突变；FLAIR，液体衰减反转恢复；AUC，受试者工作特征曲线下面积；Accuracy，准确度；f1 score：用来衡量模型分类性能的指标；MGMT，甲基转移酶启动子甲基化；ADC，表观扩散系数；FA，分数各向异性；CE，增强对比度；DTI，扩散张量成像；WHO，世界卫生组织；MRI，磁共振成像；TP53，肿瘤蛋白 p53 基因；ATRX， α 地中海贫血/智力低下综合征 X 连锁；EADC，指数表观扩散系数；TERT，端粒酶逆转录酶

表 1-2 总结了放射基因组学最近的研究，更具体地说，都是使用机器学习来预测利用磁共振成像(MRI)的分子标记物。大多数研究主要集中在弥漫性胶质瘤及其异柠檬酸脱氢酶(IDH)、甲基鸟嘌呤 DNA 甲基转移酶(MGMT)、端粒酶逆转录酶(TERT)和染色体臂 1p 和 19q(1p19q)状态，也有 EGFR 和肿瘤蛋白 p53 基因 (TP53) 的报道。对其他类型脑瘤的研究很少。由于世界卫生组织(WHO)改变了对弥漫性胶质瘤的报告方式的建议，激发了人们对分子标志物的兴趣^[50]。具体地说，WHO 建议弥漫性胶质瘤具有 IDH 突变状态，1p19q 染色体缺失，应报告 TERT 扩增，组织学分级是次要的。文章[51]展示了使用这些分子生物标记物对临床行为、治疗反应和结果相似的胶质瘤进行分组的价值。识别这些生物标记物可能会使患者的治疗更具针对性，从而改善结

果。

将现代机器学习方法应用于医学图像可能在很大程度上有助于提高医学诊断，这一点令人非常兴奋。在检测不准确(1p19q 荧光原位杂交(FISH)检测中有 5~10%是假阳性)或由于样本不足而无法提供结果的情况下，通过常规成像确定脑瘤基因组学的能力将是有帮助的。肿瘤在基因上是异质性的，成像也可能提供比组织活检更好的“全局性评估”，并且可能有助于描述宿主对肿瘤的反应。成像生物标记物可能有助于进一步对受试者进行分类，有可能成为临床试验的有用分层工具。影像学可以补充或确认组织病理学的初步诊断和治疗，还可以提供对治疗对肿瘤影响的洞察力。一些人认为，简单的体积测量可能表明低级别胶质瘤(LGG)何时可能转变为多形性胶质母细胞瘤(GBM)^[52]，但这是有争议的^[53]，更深入的定量/机器学习方法可能会提高本文对这一挑战的理解。总而言之，在常规的临床影像中可能有大量的信息可以使用机器学习方法来提取用于预测分子级别的特性。

有很多关于研究弥漫性胶质瘤的成像表型和肿瘤分子标志物之间的联系的研究^[54, 55]。MRI 特征被发现可以预测低级别胶质瘤的生存和分子数据^[56]。经典机器学习已经利用成像特征从多模式 MRI 特征中以中等精度预测 IDH 和 TP53 突变状态^[57, 58]。最近，利用针对分类和分割的深度学习的研究表明，在乳房摄影肿瘤分类、间质性肺疾病分类、脑癌分类、皮肤癌分类和视网膜分类中取得了成功。深度学习已经成为许多基于图像的竞争的首选技术，无论是在医学领域，还是在更广泛的计算机科学领域，在这些不同的挑战中，深度学习都实现了性能的大幅跃升。

类似的方法已被用于 MRI 预测肿瘤分级以及鉴别原发性中枢神经系统淋巴瘤和胶质母细胞瘤^[59]。Zacharachi 等人^[60]利用纹理特征和监督分类器对不同类型的脑瘤和胶质瘤进行分级。Citak-er 等人^[61]实现了多模态方法，支持向量机分类器达到了 0.93 的准确率。

已报道的大多数方法需要手动勾画肿瘤。并且目前关于 Ki67 的预测研究较少，仅注意到邢福勇等人基于染色后的病理图像设计了一个弱监督的语义分割网络用于预测肿瘤的 Ki67 指数^[62]，尚未发现利用非侵入式信息预测 Ki67 指数的研究。

因此在本研究中，一个新的多模态深度学习网络(KiNet)被提出用于非侵入式地预测 Ki67 指数。KiNet 可以有效融合多模态 MRI 和异构元数据(年龄、性别)等信息，首次尝试实现 Ki67 指数非侵入式预测。

1.3 本文主要研究内容和组织结构

为了在无专家知识的情况下非侵入式地预测胶质瘤分子分型，如图 1-3 所示，本算法一共分成了三个基本模块，分别是数据预处理、肿瘤分割、分子标记物状态预测。原始图像经过配准、颅骨剥离、偏置场校正等预处理操作得到用于肿瘤分割的图像，随后通过网络架构搜索得到的分割网络进行肿瘤分割，标记肿瘤的位置。肿瘤位置信

息会作为辅助信息输入到多模态分子标记物预测网络得到分子标记物状态，医生通过综合分析多种分子标记物状态可知胶质瘤的分子分型。

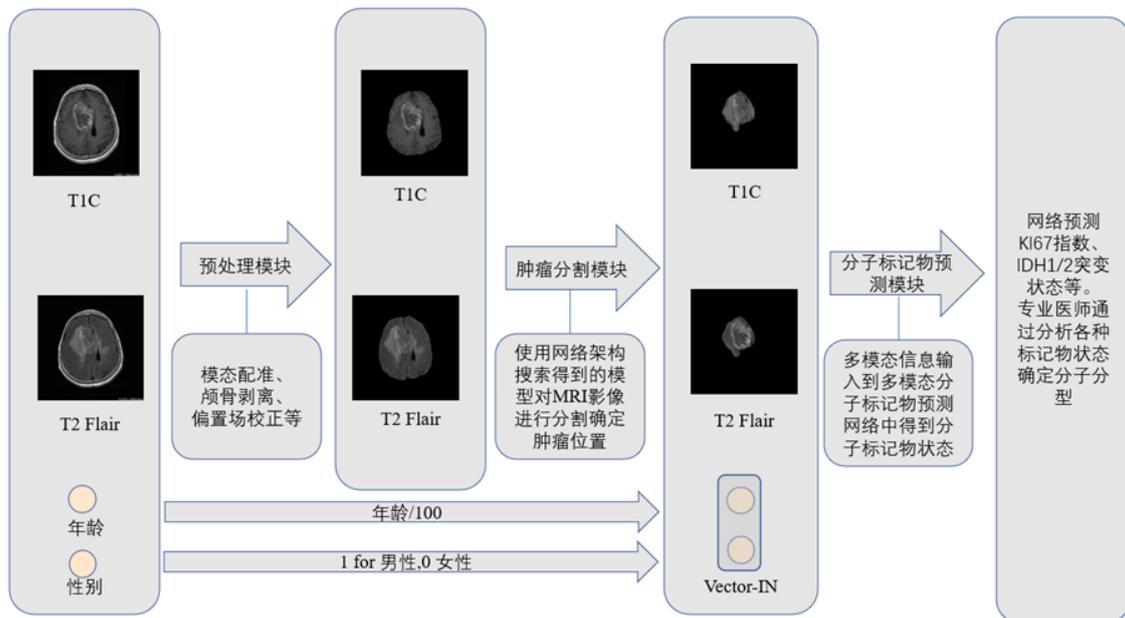


图 1-3 整体算法框架

本文将基于多模态数据的胶质瘤分子分型预测方法拆分为基于网络架构搜索的脑部 MRI 影像分割算法和基于多模态神经网络的分类算法，第一章中描述了胶质瘤分子分型预测的研究背景和意义，由于现有研究大多需要医生在原始 MRI 影像勾画相关感兴趣区域（ROI）才可以进行后续分子分型预测，提出了将该任务其划分为分割和多模态分类两个任务，介绍了两种任务的研究现状，从而引出本文的研究内容。第二章阐明了本文算法所需要掌握的一些基础知识。第三章以预测 IDH 突变为例，验证了在 MRI 影像中勾画 ROI 对于预测 IDH 突变的重要作用，并探究了不同模态融合方式对于分类性能的影响。第四章介绍了本文探索了基于网络架构搜索技术的 MRI 影像 ROI（脑肿瘤区域）分割算法，并与主流算法对比验证了本文算法的有效性。第五章介绍了本文在第四章 MRI 影像 ROI（脑肿瘤区域）分割算法基础上，提出了融合多模态 MRI 影像的多任务多模态深度神经网络，实现了无专家知识的情况下非侵入式地预测胶质瘤 Ki67 指数。在第六章中，本文回顾了本文的工作，同时分析了工作的意义和不足，针对工作中不完善的部分，提出了一些在深入研究中值得关注的方向。

第 2 章 胶质瘤分子分型预测算法理论基础

2.1 卷积神经网络

使用神经网络处理图像数据时，把本层神经元与之前层输出所有神经元（像素）完全连接是不切实际的，一方面计算量过大，另一方面忽略了数据的空间结构。卷积神经网络（CNN）中的卷积层结构是权值共享的，即让一组神经元使用相同的权重，这就使得参数数量减少，整体模型的复杂度也随之降低了很多^[72]。CNN 复合多个“卷积层”和“下采样层”对输入信号进行加工，然后在全连接层实现与输出目标之间的映射。每个卷积层都包含多个特征图，每个特征图是一个卷积滤波器提取输入的一种特征^[73]。卷积操作在卷积层内进行，通过滑窗的方式将线性卷积核和输入信号的每一处局部感受野区域进行内积计算，然后通过一个非线性的激活函数，产生一个对映射输入信号位置的激活值，最终获得相应的特征图。下采样层其作用是在减少数据量的同时保留有用的信息。在卷积神经网络中，深层的卷积核通常对应着高层特征，浅层的卷积核对应着低层特征^[74]。换言之，通过多层处理，低层的特征逐渐被转化成高层特征，用“简单模型”即可完成复杂的分类等学习任务。卷积神经网络的核心思想就是通过权值共享、下采样等方法最终使得获得的特征在某种程度上具有尺度、位移和形状的不变性。

2.2 主流卷积神经网络模型

2.2.1 AlexNet 模型

AlexNet 解决的是图像分类问题^[75]，其中输入是 1000 种不同类别（例如猫，狗等）之一的图像，而输出是 1000 维的向量。向量的第 i 个元素表示输入图像属于第 i 类的概率。输出向量的所有元素的总和被归一化为 1。

图 2-1 为 AlexNet 的网络结构。AlexNet 比以前用于计算机视觉任务的 CNN 更大。它具有 6000 万个参数和 65 万个神经元。当然现在可以在非常大的数据集上运行更为复杂的 CNN，也可以非常高效地在更快的 GPU 上运行。但是早在 2012 年，AlexNet 是巨大的。

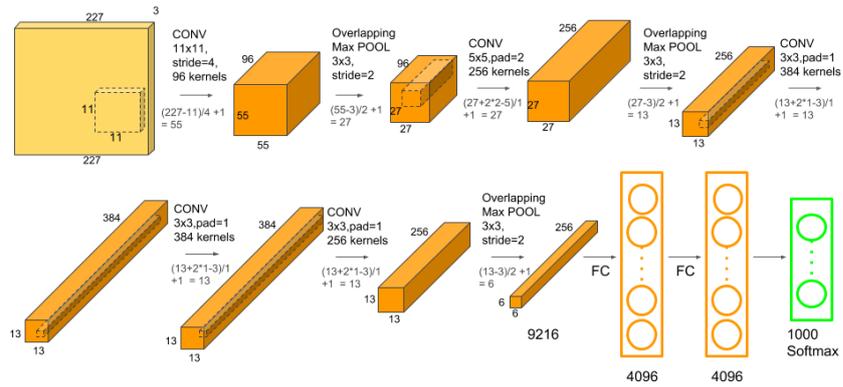
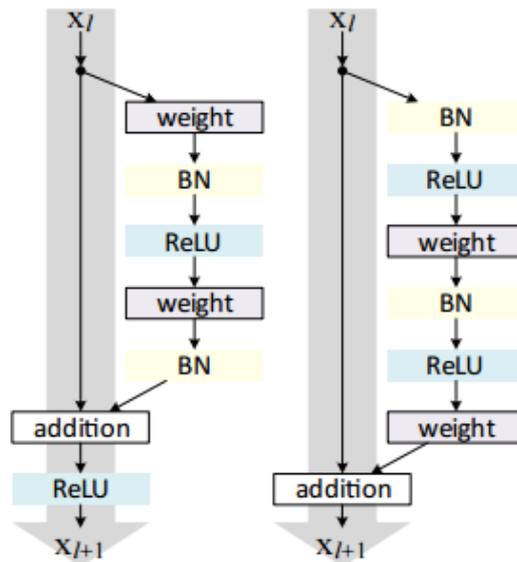


图 2-1: AlexNet 网络架构

2.2.2 ResNet 模型

与传统的顺序网络体系结构（例如 AlexNet）不同，ResNet 而是一种依靠微结构模块（也称为“网络中网络结构”）的“新颖结构”形式^[76]。术语微结构模块是指用于构建网络的一组“构建块”。一组微结构构建块（以及标准卷积，池化等层）构成了宏观结构（即最终网络）。

最早由何恺明等人提出。在其 2015 年论文《用于图像识别的深度残差学习》中，ResNet 架构已成为一项开创性工作，表明可以通过使用残差模块使用标准 SGD（以及合理的初始化）来训练极深的网络。他们在 2016 年展示可以通过更新残差模块以使用身份映射来获得更高的准确性^[77]。



(a) 原始残差模块 (b) 使用预激活更新的残差模块

图 2-2 : ResNet 中的残差模块

即使是 ResNet 比 VGG16 和 VGG19 深得多，模型的大小实际上是相当小的，对于 50 层的 ResNet，使用全局平均池化代替全连接层使得模型的大小降低到 102MB。

2.2.3 全卷积网络

语义图像分割任务是图像的每个像素都赋为一个类别。在医学图像分析领域，图像分割可用于病灶分割和手术引导等。

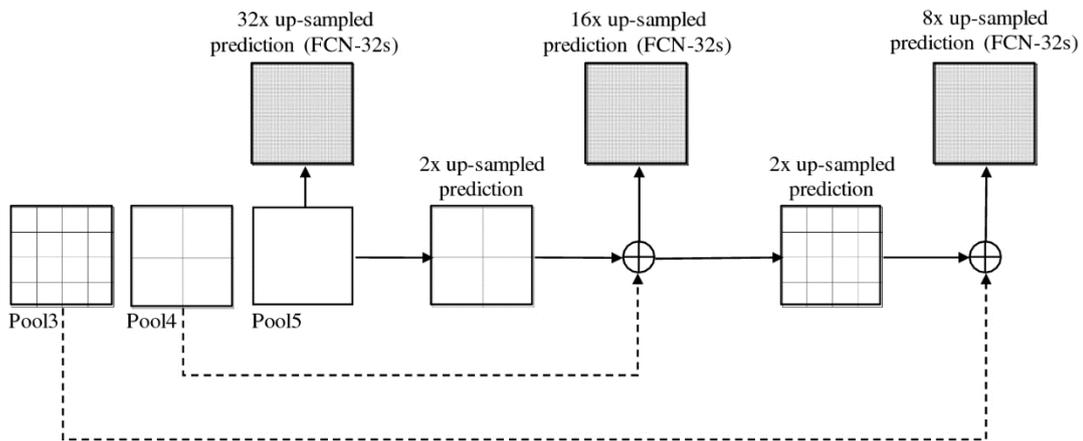


图 2-3 FCN 的上采样过程

作为最早的基于 CNN 的高影响力语义分割模型之一，Long 等人提出了全卷积网络（FCN）的像素级标记^[78]。他们提出对输出激活图进行上采样（逆卷积），从而可以得到和原始图像大小相同的密集预测。为了在图像中保留上下文空间信息，作者提出将输出与较浅层的输出融合。融合步骤如图 2-3 所示。

卷积神经网络的完全连接层需要固定大小的输入。在 FCN 中采用的解决方案是用 1x1 卷积层替换全连接层。因此，FCN 可以对任何输入大小的图像执行语义分割。

2.2.4 U-Net 模型

U-Net 建立在 FCN 的概念上^[79]。与上面的编码器-解码器体系结构类似，它的体系结构可以分为三个部分：

下采样路径包括 4 个块，其中每个块两个 3x3 卷积层（每个卷积层后接批归一化层），随后通过 2x2 最大池化。在每个池化层（在每个块之后），特征图的数量将增加一倍，64 -> 128 -> 256 依此类推。水平瓶颈模块包括两个 3x3 卷积，然后是 2x2 逆卷积。与下采样路径的扩展或上采样路径也包含 4 个块，其中每个块包含两个 3x3conv，然后进行 2x2 上采样（逆卷积）。每个块之后，此处的特征图数量减半。

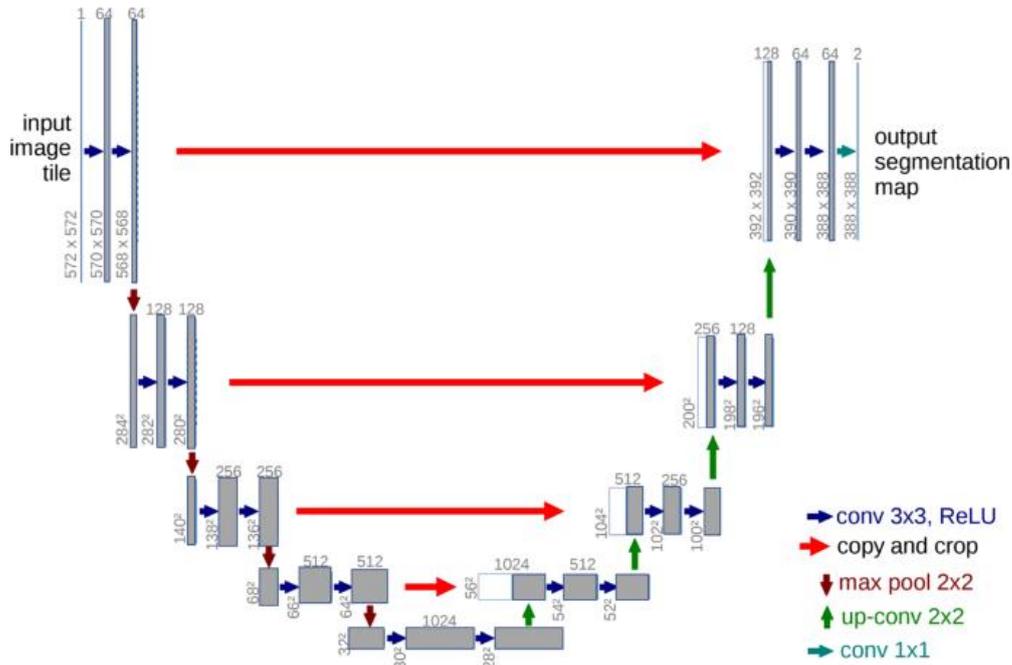


图 2-4: UNet 网络架构

图 2-4 是 UNet 的网络架构，图中每个灰色的方块表示一个多通道的特征图。通道数量标注在方块上方。特征图大小在方块边缘的左下方标明。白色方块表示复制的特征图。不同箭头表示不同的计算操作。

U-Net 还具有跳跃连接以便进行局部细节还原，如图 2-4 白色方块所示。将上采样的输出与来自相应层级的下采样路径裁剪后的（由于在每个卷积中丢失边界像素而被裁剪）特征图进行级联。最后，结果输出通过 3×3 卷积层以提供分割输出，其中特征图的数量等于所需的分割类别数。

2.3 神经网络架构搜索

深度学习通过以端到端的方式与统计模型一起学习表示形式，提供了绕过手动特征工程过程的希望。但是，神经网络体系结构本身通常是由专家以困难的特别的方式设计的。神经网络架构搜索（NAS）被定义通过自动搜索优于手工设计的模型结构。

实际上，NAS 是超参数优化的子集。而且，专用的 NAS 方法实际上并不是完全自动化的，因为它们依赖于人为设计的体系结构作为起点。2016 年，MIT 和 Google 的学者们将强化学习引入到深度神经网络结构的搜索中^[64,65]，取得了不俗的成果。在几个小型的经典数据集上击败同时代同级别的手工设计网络。但这个方法有个缺点是消耗计算资源巨大。但是，在过去的几年中，专用的 NAS 方法在提高准确性，降低计算成本和减小结构大小方面表现出了显著的进步，并且最终可以在神经网络结构设计上超过人类的性能。

Elsken 等人^[63]总结 NAS 三个主要组成部分分别为：

(1) 搜索空间。该组件描述了要考虑的一组可能的神经网络体系结构。这些搜索空间是为特定于应用程序而设计的，例如，用于计算机视觉任务的卷积网络空间或用于语言建模任务的循环网络空间。因此，NAS 方法不是完全自动化的，因为这些搜索空间的设计从根本上依赖于人类设计的架构作为起点。即使这样，仍然存在许多架构决策。实际上，在这些搜索空间中考虑的可能架构的数量通常超过 10^{10} 。

(2) 优化方法。该组件确定如何探索搜索空间以找到良好的体系结构。这里最基本的方法是随机搜索，同时还引入了各种自适应方法，例如，强化学习，进化搜索，基于梯度的优化和贝叶斯优化等。尽管这些自适应方法在如何确定要评估的体系结构方面有所不同，但它们都试图将搜索偏向于性能更佳的体系结构。所有这些方法都具有在传统超参数优化任务中引入的对应方法。

(3) 评估方法。该组件衡量通过优化方法考虑的每种体系结构的性能。最简单但最昂贵的计算选择是完全训练体系结构。人们可以选择利用局部训练，还引入了特定于 NAS 的评估方法（例如权重共享和超网络）来利用神经网络的结构来提供更节能的启发式性能评估。部分训练方法通常比完全训练消耗少一个数量级，而 NAS 特定的评估方法比完全训练消耗少 2-3 个数量级^[67]。

2.4 知识蒸馏

在工程中，一般希望部署到应用中的模型除了具有较好的准确率，还需要使用尽量少的计算资源(存储空间、计算单元等)，产生较低的时延。

在深度学习的背景下，过参数化的深度神经网络的学习能力很强，一般可以得到较好的准确率，但由于其需要大量的计算，对部署非常不利。但是从零训练一个小模型，从经验上看是很难达到大模型的准确率，于是从小模型（学生模型）学习一个已经训练好的大模型（教师模型）的预测结果的方案被提出来。

Rich Caruana 等人指出可以让学生模型近似教师模型^[68]。在深度学习理论中，通常假定输入到输出有一个未知潜在的函数关系，重新训练一个学生模型就是从已有的数据集中近似这个未知的函数。如果让学生模型近似教师模型，考虑到教师模型是已知的，由于可以使用很多非训练集内的伪数据来训练学生模型，所以训练变得更简单。

传统深度学习网络训练的目标是让学生模型的 softmax 分布与标签生成的独热编码相同，现在只需要让学生模型与教师模型在给定输入下的 softmax 输出分布匹配。显然在同样的输入下 softmax 输出相比独热编码会提供更丰富的信息量。

知识蒸馏的目标是让学生模型与教师模型的 softmax 输出差异尽量小。但直接使用 softmax 输出作为监督的问题在于在一般的 softmax 函数中输出是一个接近独热编码的向量，其中一个值很大，其他的都很小。此时，输出所蕴含的信息量仍然较少。相较类似独热编码这样的硬性输出，输出更“软”一些会有利于网络训练。

一种方法是直接比较模型 softmax 的输入 (logits) 来避免这个问题。记教师模型

产生的 logits 是 v_i ，学生模型产生的 logits 是 z_i ，本文需要最小化： $\frac{1}{2}(z_i - v_i)^2$

Hinton 等人提出了知识蒸馏的概念[69]，将 softmax 函数扩展为：

$$q_i = \frac{e^{\frac{z_i}{T}}}{\sum_j e^{\frac{z_j}{T}}} \quad (2-1)$$

其中 T 是温度，这是从统计力学中的玻尔兹曼分布中借用的概念。当 $T=1$ 时，即为常用的 softmax 函数，当温度 T 趋向于 0 时，softmax 输出将收敛为一个独热向量；温度 T 趋向于无穷时，softmax 的输出则更“软”。因此，在训练学生模型的时候，可以使用较高的 T 使得 softmax 产生的分布足够软，这时让学生模型的 softmax 输出近似教师模型；在训练结束以后再使用正常的温度 T 来预测。具体地，在训练时本文需要最小化两个分布的交叉熵(Cross-entropy)，记学生模型利用公式 2-1 产生的分布是 q ，教师模型产生的分布是 p ，则本文需要最小化：

$$C = -p^T \log q \quad (2-2)$$

在前面提到的这个过程中，本文先让温度 T 升高，然后在测试阶段降低温度 T ，从而将教师模型中的知识提取出来，因此将其称为是蒸馏。

当然，如果转移时使用的是有标签的数据，那么也可以将标签与学生模型 softmax 分布的交叉熵加入到损失函数中去，实验表明这将大大改善学生模型的表现。

2.5 多模态深度神经网络模型

最基础的多模态融合方法就是通道拼接，顾名思义，通道拼接的方式就是将每种模态抽取出来的表达连接到一起后（后面连接全连接层）学习整体的特征，这种方式简单，并且能够提供一个不错的基准。

在已有的有关多模态融合的文献中常常用模型在哪一层融合做出定义，通常包括早期融合、晚期融合以及杂交融合等。但是对于在模型的哪一层进行融合能带来最好的效果，学者们没有能达成一致。在深度学习神经网络模型中，不同模态融合以及模态特征表征是难以区分的，往往采用通道拼接，逐元素相加等方法来使得多模态信息特征映射到同一个多模态模型空间维度，这些方式能被视为学习一种联合表征^[70]。而与之不同的是，联合表示有更多的补充性，能最大化发掘出不同多模态模型表征的相关性。

通道拼接的方法相当于在各自模态的特征分别独立抽取之后做融合，但是不干预特征抽取的过程。这显然会漏掉一些不同模态之间的相关性的信息，比如，一个人尖叫的视频片段，里面有人的嘴巴长得很大的这种视觉特征，也会有尖叫声这种声音特征，这两种特征出现的时间都是相关连的。如果在抽取特征的过程中，相互有一个监督或者借鉴的关系，那样应该会更好，这是 CentralNet 提出的动机^[71]。文献[71]提出了 CentralNet 同时借鉴多模态表征，先构建多个用于独立处理单一模态的模型，并提

出一个中心网络连接各个模态的网络，把不同模态的表征映射到相同的公共空间。

图 2-5(a)是基本的多模态融合方法，首先独立的模型 M_1, M_2 提取出来各自模态的特征，然后一起输入到融合模块（Fusion）进行融合，融合得到的特征送入分类器去分类。

图 2-5(b)是 CentralNet 的示意图，它各个模态模型 M_1, M_2 中间分支的特征也拿到公式 2-3 加权送入下一层。如此一层一层地做，传到最后一层得到一个融合特征 h_{C_x} ，利用这个融合特征送入分类器分类。

$$h_{C_{i+1}} = \alpha_{C_i} h_{C_i} + \sum_{k=1}^n \alpha_{M_i^k} h_{M_i^k} \quad (2-3)$$

CentralNet 模型借助多任务对每个模态的表达进行约束，以期融合后的表达能够获取更好的泛化能力。训练的损失函数是各个模态损失和融合后特征送入融合分类器后得到的损失的一个加权和：

$$loss = loss_C + \sum_k \beta_k loss_{M^k} \quad (2-4)$$

CentralNet 把两个可以独立在各自模态做出判断的网络的中间层联系起来，利用联合多任务训练来优化，最后也得到了不错的结果。

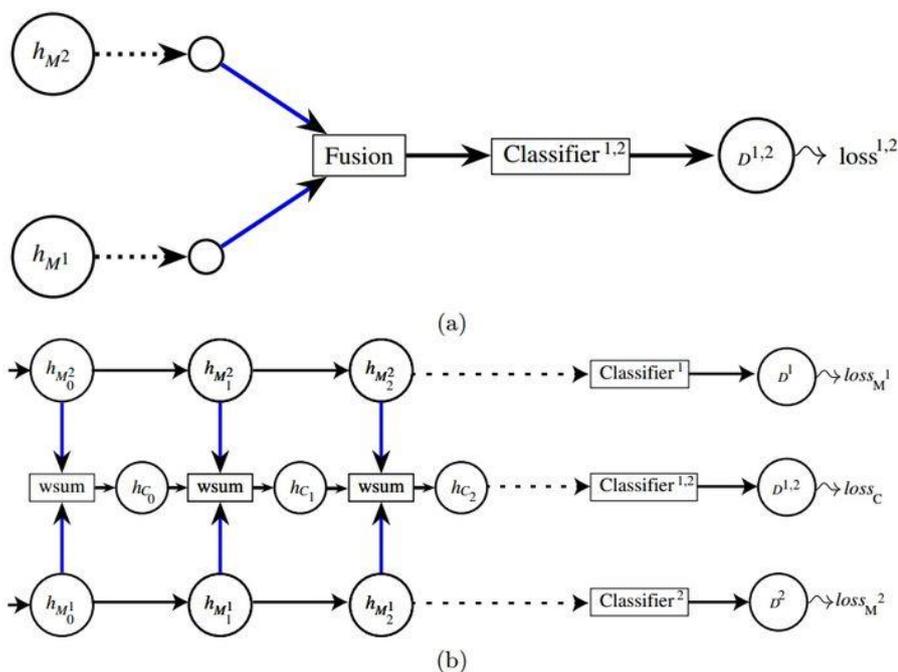


图 2-5 CentralNet 模型多模态融合方法

2.6 本章小结

在本章中，简要介绍了需要理解本文提出的分子分型预测算法所需的一些基本知识，包含第一节中的卷积神经网络相关介绍，第二节是用于本文模型对比的相关神经

网络模型, 第三节和第四节介绍了第 4 章 ROI 分割算法中所需的神经网络架构搜索和知识蒸馏相关知识, 第五节介绍了第 5 章中多模态神经网络的一些知识和主流多模态模型。以上的基础知识是理解后文的技术方法所必须的, 但由于篇幅问题, 所以在本章中统一介绍。

第 3 章 胶质瘤 IDH 突变状态预测

3.1 引言

异柠檬酸脱氢酶(IDH)突变状态是胶质瘤诊断和治疗的重要指标之一。具有这种突变酶的胶质瘤比具有野生型 IDH 的相同级别的肿瘤预后更好,治疗方法也有所不同^[89]。目前,确定 IDH 突变胶质瘤的唯一方法是对通过活检或手术切除获得的组织标本进行免疫组织化学(IHC)或基因测序。由于 IDH 突变和 IDH 野生型胶质瘤之间的差异可能具有重要的治疗意义,所以在手术前尝试区分这两种肿瘤类型是非常有意义的。IDH 突变状态的早期确定直接影响治疗决策。对于那些看起来是低度恶性胶质瘤但实际上是 IDH 野生型的肿瘤,通常采用早期干预而不是观察治疗。对于 IDH 突变的胶质瘤,特定的化疗干预措施(如替莫唑胺)更有效^[90,91]。此外,已证明在 III-IV 级 IDH 突变的肿瘤中,手术切除未增强的肿瘤(比全部切除增强肿瘤)更具有生存益处^[92]。然而,IDH 突变状态的确定需要使用组织取样。获取富含肿瘤的组织样本以确定 IDH 状态可能是一项挑战。癌症基因组图谱(TCGA)的一份报告表明,只有 35%的活检样本含有足够的肿瘤成分,可以进行准确的分子特征鉴定。开发一种强有力的非侵入性方法将对这些患者的护理更加有利。

在基于 MRI 的肿瘤遗传和分子生物学预测方面,深度学习方法的发展超过了传统的机器学习方法^[93]。最近,Matsui 等人同时利用了多模态 MRI, PET, and CT data 预测 IDH 分型,达到了 68%的准确率,添加病人年龄性别等信息后,能取得更高的准确率^[94]。Chang 等人最近的研究使用深度学习技术来基于 MRI 非侵入性地确定 IDH 突变状态,使用癌症成像档案馆(TCIA)的数据库,准确率达到 94%^[95]。这些方法均使用 2D(切片)分类方法,设计和开发切片分类模型的一个已知限制是数据泄漏问题^[96,97]。Chandan 等人提出以上方法在临床应用都受到了较大的限制,由于需要多对比度采集,扫描时间长而经常受到患者运动的影响。所以他们基于 TCGA 数据集 T2 加权(T2)图像建立一个高精度的全自动深度学习 IDH 分类 3D 网络,交叉验证准确率达到 97%。然而在临床中获取密集的扫描序列成本较大,这限制了三维网络的应用。

本章节的目的是在排除使用切片数据可能导致的数据泄露问题后,验证手工勾画的肿瘤区域作为 ROI 作为预测的原始输入的作用,探究一种更鲁棒的多模态图像切片融合方式,缓解患者运动对多参数成像的模态不匹配,从而获得更好的 IDH 预测结果。

3.2 数据集预处理

这项回顾性研究的数据来源于近三年(2017-2020)就诊于中南大学湘雅医院神经

外科，所有纳入研究对象患都经由术后病理诊断确诊神经胶质瘤，共包含 263 例患者数据，数据内容包含患者基本信息（性别年龄）、MRI 图像、术后病理 IDH 突变状态化验结果（野生型标注为 0，突变型标注为 1）。其中患者有 113 例男性，年龄分布范围是 13 岁到 75 岁，平均年龄是 45 岁。MRI 数据包含四个模态(T1,T1+C,T2,T2flair)。在研究中，利用五折交叉验证来评估模型的预测和泛化能力。

医院整理得到的图像数据 NG-263 为 MRI 影像，这得益于 MRI 影像在脑部神经胶质瘤方向的广泛应用。根据专业医师提供的先验知识，在 MRI 各个序列中，T1、T1+C、T2 及 T2Flair 四个序列图像会较大程度地反映脑部的真实结构，对神经胶质瘤的突变识别有着重要的指导意义。每例患者数据集中都包含了以上四个序列的约 20 张图像，单个患者共约 80 张图像。在图像的预处理阶段，

1) 首先分析了单一序列中不同层的图像的成像特点，两端的层图像中不包含有肿瘤的实体，这将会对神经胶质瘤的 IDH 突变诊断产生很大的阻碍，所以我们请专业的医师团队对数据集中的图像就其是否具有明显的肿瘤实体进行了严密挑选标注，并在另一个专业团队的再次核对下确定用于下一步研究的数据集 ENG-263。

2) 在分析 IDH 突变胶质瘤通常情况下位于一侧的某一部位，而脑部 MRI 图像的结构成分又十分复杂（如脑脊液白质等），卷积神经网络的特征学习会包含过多驳杂的无用特征，这样对于研究 IDH 突变辅助诊断十分不利。考虑到这一点，我们打算就分子分型所需关键感兴趣区域（ROI，一般为肿瘤区域）进行标注处理，该任务由经验丰富的专业医师团队执行，具体的标注软件是 LabelImg。

3) 数据集中 MRI 影像包含四个扫描序列，由于扫描的对象都是同一病例，那么不同序列之间是否会存在相关特征，我们的研究会探索同一病例的四个不同序列图像在时序排列上的关联性。但是在寻找研究可行性解释的过程中，所设想的研究是建立在同一病例的不同序列图像达到像素级对齐的基础上，而区别于其他三个对齐的序列，T1+C 序列在注射增强后，不可避免地会引入偏移，但还是保留着脑部与、其他三个序列图像中的相同的结构。在实验的预处理阶段，我们进行了不同序列的配准，整理得到同病例中序列间对齐的图像数据集并对图像进行平移，旋转及翻转来数据增强，得出满足研究泛化需求的图像数据集 SNG-263。

处理后所得的数据集在应用于卷积神经网络模型前，研究针对所有病例中的每个序列图像，都计算得出该序列图像数据的均值和方差，在模型处理阶段执行均值方差归一化操作，在带来较快收敛速度的同时能够保留尽可能多图像信息。

3.3 基于多模态图像的 IDH 突变状态预测算法

多模态数据的融合能够为模型决策提供更多的相关信息，其目的是处理和关联来自多个模态的信息特征。如本文脑部胶质瘤的多参数 MRI 生成的多模态图像，能够反映同一对象不同层次的特征信息。目前多模态数据融合的方式主要有基于抽象特征层

面的早期融合、晚期融合以及中期融合等，不同的融合方式会带来不同的参数量以及模型诊断效果。

3.3.1 网络骨架和输入模式选择

以卷积神经网络为主流的深度学习模型已经被应用在各类研究任务中，卷积神经网络的模型改进方向也受到研究学者们的重点关注，其中迈出重要步伐的改进是 ResNet 模型^[76]。

ResNet 本着复制浅层网络的特征来使得深层网络至少具有浅层网络的性能的假设，有效改善了卷积神经网络模型在网络层数加深后出现的梯度消失或爆炸的情况。

最近，有研究^[88]提出更全面的 ResNeSt 模型。该模型首先在 ResNet 网络残差块的结构中，利用小尺寸卷积的组合在缩小模型参数量的同时引入更多的感受野尺寸，还在块结构中加入注意力思想来增强相关特征表达。

本研究设计五折交叉验证实验，来确立本研究任务的网络骨架以及验证感兴趣区域 (ROI) 掩膜带来的性能提升，即输入模式对网络性能的影响。实验基于 SNG-263 数据集，权重采用在 ImageNet 数据集上预训练得到的参数初始化，以 IMG (4 序列图像通道)，ROI (4 掩膜图像通道) 和 MIX (4 序列图像+4 掩膜图像) 作为特征向量输入，分别在 ResNet50，ResNeSt50 模型上设置 6 对照实验来预测 IDH 突变，并根据相关的评价指标进行评估结果。

3.3.2 多模态分类模型

在前文中我们处理多模态图像的方式是将四个序列图像按灰度图读入，并通过通道叠加，构成多通道的输入图像，这就是目前最主流的多模态数据早期融合方式。早期融合能够将各个模态图像的原始特征结合在一起，但是原始特征比起深层特征，存在大量的冗余信息，且以全局特征为主，这对于困难样本的特征识别并无明显的提升。

后期融合^[98]一般是将不同模态的图像数据分别训练并进行输出融合的方式，区别于早期融合，后期融合会得到不同模态各自独立的深层特征，更偏重于控制总损失来实现预测决策优化，这样可以避免单独一个分类器中的错误累加。在本研究中，考虑到序列图像与掩膜图像之间的对应关系，我们将同一序列与其掩膜图像作早期通道融合，在后期融合模型通过连接 4 个序列训练模型的输出特征图，并输入到一个全连接层的方式实现。

至于中期融合，则可以选择融合的位置，这对于具体的研究任务能带来非常大的灵活性，它是将图像转化为高维特征向量表达，并在模型中间层以特异的方式结合，这样的训练模型容易获取到不同模态间的共性，所谓共性，正是具体研究所需要的识别特征表达。在本研究中，我们在不同的模型位置尝试了两种中间融合模型构造，首先是如上一个实验一般，得到 4 个同一序列及其掩膜的输入特征向量。第一种融合模

型是将块结构的输入以通道融合的方式连接，旨在关联探寻其高级特征之间的特征共性。第二种融合模型同样是在块结构之前，将 4 组特征向量通过通道降维，与 4 组特征向量的独立输出连接在一起，在寻找高级特征的同时，也关注 4 组序列图像的包含位置特征信息在内低级信息上的关联特征。

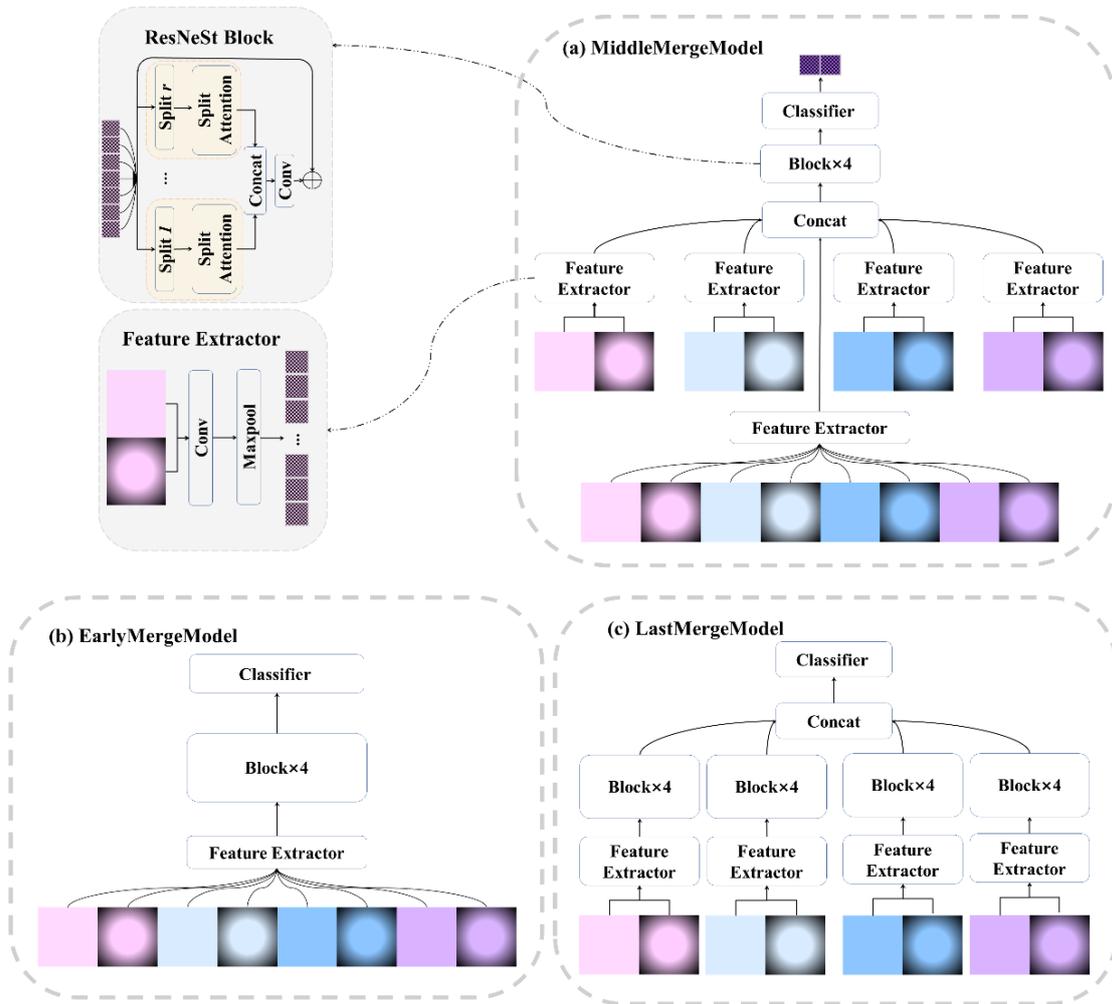


图 3-1 多模态融合网络结构

我们的研究在探索最优网络骨架和输入模式后，基于 SNG-263 数据集，分别在三种融合方式的创新模型上进行了五折交叉验证实验，融合模型的结构如图 3-1 所示。

3.4 实验结果与讨论

3.4.1 算法评估准则

一般来讲，对于二分类问题而言，得到的结果可以分成以下四种情况：实际为正例，且判断为正例（真阳性,TP）、实际为正例，但判断为负例（假阴性, FN）、实际为负例，且判断为负例（真阴性, TN）及实际为负例但判断为正例（假阳性, FP），如表 3-1 所示。

表 3-1 分类算法结果分类

预测结果\真实结果	正例	负例
正例	TP	FP
负例	FN	TN

准确率(accuracy)、敏感度(sensitivity)和特异度(specificity)是评价医学鉴定方法性能的常用指标。准确率表示分类的正确率,敏感度是对阳性样品的敏感度,特异度可以理解成敏感度的在阴性样本上的说法,通过对所有出现的结果进行统计,可以得到 TP, TN, FP 和 FN 的数量,便可以以此计算算法的准确率、敏感度和特异度,计算公式如式(3-1)、式(3-2)和式(3-3)所示:

$$\text{accuracy} = \frac{TP+TN}{TP+FP+FN+TN} \quad (3-1)$$

$$\text{sensitivity} = \frac{TP}{TP+FN} \quad (3-2)$$

$$\text{specificity} = \frac{TN}{TN+FP} \quad (3-3)$$

通过设定不同的阈值 τ_r ,可以得到一组敏感度及特异度的值,以敏感度为纵坐标,并以 $(1-\text{specificity})$ 为横坐标,可以绘制出受试者工作 (ROC) 曲线。一般而言,在比较不同的分类模型时,可以将每个模型的 ROC 曲线都画出来,比较曲线下面积 (AUC) 做为模型优劣的指标, AUC 面积越大说明算法性能越好。

3.4.2 不同网络骨架和输入模式实验结果对比

在研究网络骨架及输入模式对网络性能的影响的实验中,6 组对照实验均采用五折交叉验证,在 SNG-263 数据集的模型预测结果如表 3-2 所示。

表 3-2 网络骨架及输入模式对网络性能的影响

模型	准确率	敏感性	特异性	AUC
ResNet+IMG Model	0.646	0.578	0.435	0.654
ResNeSt+IMG Model	0.687	0.616	0.454	0.677
ResNet+ROI Model	0.697	0.617	0.476	0.702
ResNeSt+ROI Model	0.734	0.624	0.563	0.754
ResNet+MIX Model	0.728	0.618	0.577	0.742
ResNeSt+MIX Model	0.766	0.734	0.715	0.842

由对照实验的结果可知,使用 ResNeSt 模型作为网络骨架的效果明显要好过 ResNet 模型,且在连接上掩膜通道后,模型性能大约能提升 7-8%,对于医学研究任务的预测结果而言,会产生重要的意义。对表 3-2 分析可以得出,模型对 IDH 野生型的预测都会比较准确,而 ResNeSt 模型相比于 ResNet 模型,在对 IDH 野生型的预测上会更有优势。

3.4.3 不同多模态融合方式实验结果对比

在选定最优的网络骨架和输入模式后，基于最优的网络骨架和输入模式，分别使用三种融合方式的模型在 SNG-263 数据集上进行了五折交叉验证实验。表 3-3 中给出 3 种融合模型在 SNG-263 数据集上的预测评估指标，图 3-2 是 3 个融合模型的 ROC 曲线，从整体来看，不同的模型融合方式能够给模型预测的性能带来不同的提升效果。我们可以从表 3-3 中看到，早期融合模型预测性能会优于后期融合模型性能，结合融合方式的差异，我们可以得出结论，针对于预测 IDH 突变这一研究任务，MRI 图像中的不同序列之间的确存在某种关联特征，这类特征信息在早期融合模型中能够被网络学习到，而侧重于独立分类器决策优化的后期融合模型则几乎不能学到这个层次的信息。另外，中期融合模型效果又优于早期融合模型效果，说明不同的序列图像之间的深层特征关联性会大于浅层特征的关联性，对我们研究的任务有更大的影响。

表 3-3 不同多模态融合方式实验结果对比

模型	准确率	敏感性	特异性	ROC
EarlyMergeModel	0.766	0.734	0.715	0.842
LastMergeModel	0.736	0.725	0.691	0.824
MiddleMergeModel	0.779	0.784	0.644	0.871

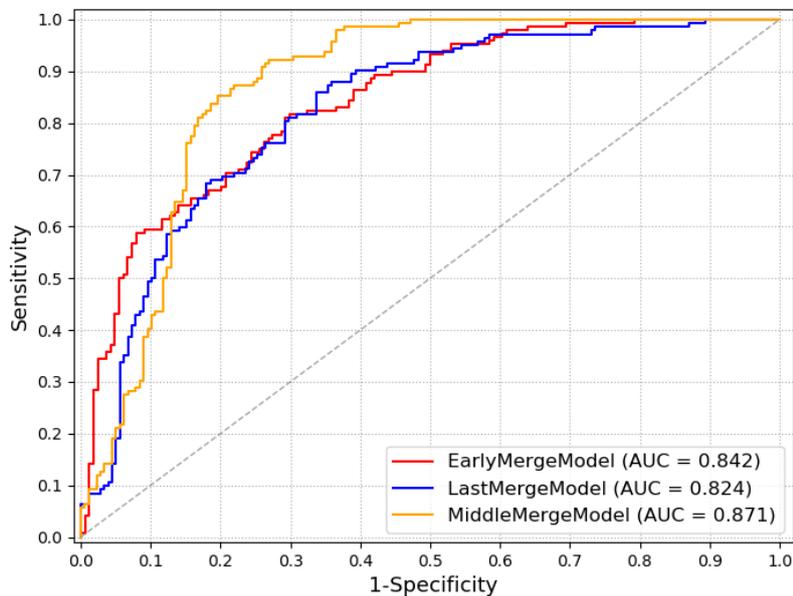


图 3-2 融合模型的 ROC 曲线

3.4.4 分析与讨论

本章探究了基于多模态 MRI 影像预测胶质瘤突变状态的深度学习方法，基于 T1, T1C, T2 和 T2Flair 四个模态的 MRI 图像构建出的最好预测模型表现出了可靠的性能，

这表明基于 MRI 影像等侵入式特征预测分子分型的可行性，值得提出的是我们的数据集中包含所有 WHO 分级的胶质瘤，在使用前不需要人为确定分级，这意味着更强的临床实用性。

为了探究为什么使用原始图像+ROI 模型有更好的效果，我们对 ResNeSt+IMG Model、ResNeSt+ROI Model、ResNeSt+MIX Model 三个模型做了梯度权重类激活图 (Grad-CAM)分析，来查看网络对原始图像的注意力分布。Grad-CAM,即使用流入最后卷积层的任何类别(如在一个分类网络的“狗”)的梯度产生一个粗定位热力图，用于强调图像中用于预测指定概念的重要区域^[99]。图 3-3 为生成的 Grad-CAM 图。

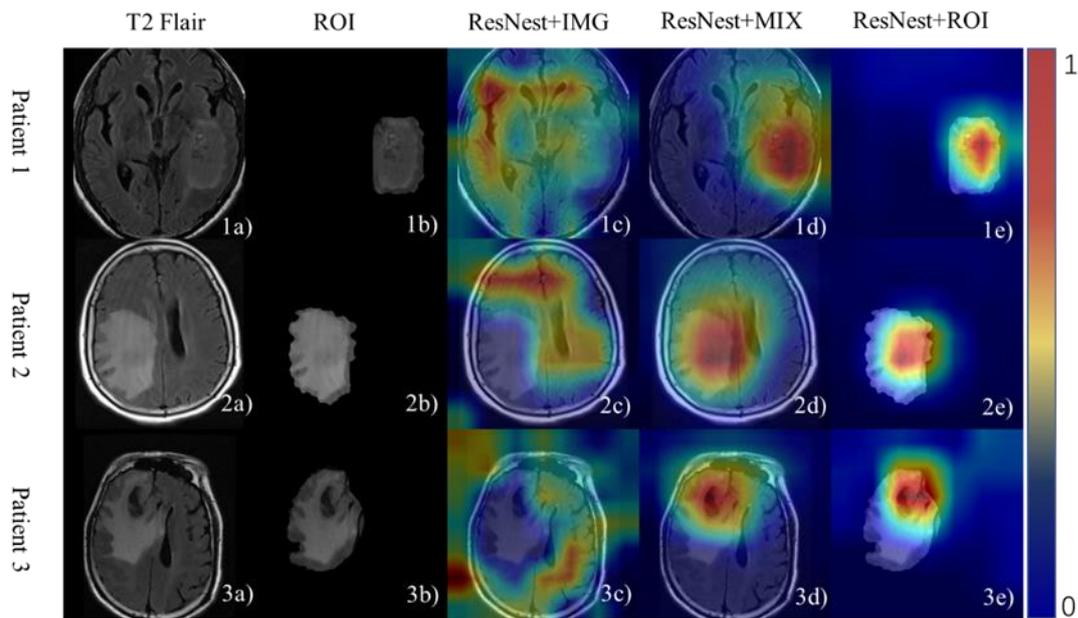


图 3-3 Grad-CAM 分析结果图。a 为 T2 Flair 模态的影像，b 为手工标记的肿瘤区域，c、d、e 分别为三个模型的 Grad-CAM 可视化结果，从蓝色到红色表示对网络结果的重要程度递增。

从图 a 可以看到，对于 ResNeSt+IMG 模型，Grad-CAM 中重要区域分布非常混乱，有时甚至没有覆盖到肿瘤区域，这也跟较差的准确率对应，这可能是由于较少的数据量导致网络无法学习到关键的特征。而 ResNeSt+MIX 模型的 Grad-CAM 表明，模型对肿瘤及肿瘤周围区域有更高的注意力，在更精细的 ResNeSt+ROI CAM 图中可以发现模型的注意力更集中在肿瘤的核心区域而不是水肿区。然而从结果上分析基于 ROI 的模型性能不如 ResNeSt+MIX 模型，根据 ResNeSt+IMG 模型的 Grad-CAM 图推测，可能是由于 IDH 突变状态也会影响肿瘤周围脑组织的分布，所以在预测 IDH 突变状态时，我们仍需要将肿瘤周边的区域考虑进去。

值得注意的是，使用随机选择切片来构建机器学习预测模型可能会有数据泄露的风险，因为它们在所有病例随机挑选切片以生成训练、验证和测试切片。结果，可以在训练、验证或测试数据子集中找到来自同一主题的相邻切片。由于相邻切片通常共享大量信息，这种方法可能会通过在测试阶段引入偏差人为地提高准确性。这可能是

之前已经发表的研究中使用类似的方法但是却得到差异较大的结果的根本原因，实际上在没有特别处理这个问题之前，我们方法的准确率超过了 90%。因此，在划分数据集的时候，我们特别注意到先按照病人编号划分训练、验证数据子集，并进行了 5 折交叉验证，最终得到了可信赖的结果。

基于机器学习的医学预测任务，通常都会使用多参数成像的 MRI 生成的多模态 MRI 图像，但由于各模态 MRI 不一定会在同一时间生成，容易受到患者运动的影响，造成各模态图像空间上不匹配。模态配准就是为了改善这个问题，但是仍无法彻底解决。基于此我们研究了几种多模态 MRI 融合的 CNN 分类网络。早期融合的就是在输入层直接将各个模态图像组合在一起成为多通道图像，这也是现在的主流做法，这种融合方法要求各模态完全配准。另一种方式是后期融合网络，每个模态的图像经过设计的分类器成为特征向量，在最后将各个模态的特征向量融合经过全连接层得到最终结果，这种方法对于模态匹配没有要求，甚至可以融合异构信息，比如图像和文本的融合。但通过结果来看，后期融合网络的准确率要比主流的早期融合网络低，这是由于对高层的特征向量进行融合无法利用到不同模态图像浅层特征的相关性。最后一种即为我们的中期融合方法，每个模态通过浅层的特征编码器生成特征图，将各个特征图所有通道拼接成为一组特征图，再输入后续的分类器，经过浅层的特征编码器生成的特征图仍保留着浅层的纹理、位置等浅层特征，在这个阶段进行融合既能消除可能的模态不完全匹配所造成的干扰，又能较充分的利用多模态的对照信息，所以取得了最好的性能。

3.5 本章小结

在本章中，建立了一个包含 263 例样本的多模态 MRI 数据集，探究了模型的输入图像方式（是否需要勾画感兴趣区域）、分类网络骨架和多模态信息融合位置对 IDH 预测准确率的影响并通过 Grad-CAM 对其结果进行注意力分析，结果证明了 ROI（脑肿瘤区域）作为附加输入对于预测分子分型的重要性。设计了多种多模态融合网络用于预测 IDH 突变状态，最佳模型能取得 AUC 为 0.87 的结果，这比主流的融合方式性能更好，同时能容忍一定程度的模态不匹配。

第 4 章 脑部 MRI 影像中分子分型 ROI 分割

4.1 引言

类似于本文第二章中的方法，报道过的胶质瘤分子标记物预测方法通常需要人工勾画用于预测算法所需的感兴趣区域（ROI，一般是肿瘤分布区域），这限制了临床上胶质瘤分子标记物预测算法的应用，为了尽量少的使用预处理步骤和医生的专业知识实现临床胶质瘤分子标记物预测，基于医院的 MRI 临床数据开发一个高效准确的 ROI 分割算法是一个必要工作。

使用神经网络对医学图像进行分割源自 Ronneberger 等^[79]提出使用 UNet 对 2D 医学图像进行分割。即使是现在 UNet 仍然被作为基本框架添加各种先进的构建块被广泛应用在各种医学图像分割任务当中。考虑到现成构建块的多样性可能会导致大量的候选结构，为 UNet 中的下采样和上采样路径挑选正确的构建模块可能仍然是一项艰巨的工作。神经体系结构搜索（NAS）算法的最新成功为设计更好的语义分割模型（特别是在其他资源约束的延迟下）开辟了新的视野。

Auto-DeepLab^[80]首次引入了网络级搜索空间，以优化分割任务的分辨率（除单元格结构外）。在医学图像分割方面，NAS-UNet 分别寻找最优的下采样和上采样单元，构建 UNet 结构，并在包括前列腺 MRI、肝脏计算机断层扫描(CT)和神经超声图像在内的二维医学图像数据集上进行了测试^[81]。

尽管取得了一些初步的成功，但在设计细分模型方面成功的人员领域专业知识似乎尚未完全集成到医学图像分割 NAS 框架中。例如，人为设计的实时分割架构通常利用具有适当深度，宽度，算子和下采样率的多分辨率分支，并发现它们对网络的高准确率做出重要贡献。此外，现有的医学图像 NAS 分割网络往往对于高精度追求较高，而忽略了临床应用中的低延时和低资源消耗要求。

在本章中，为了解决胶质瘤分子标记物预测方法通常需要人工勾画 ROI 的问题，基于临床胶质瘤分割数据集提出了一种名为 NASBraNet 的 NAS 框架，目的是实现快速的推理速度和高的准确率。本文使用了一个特殊的搜索空间，能够支持对不同分辨率的多个分支进行优化。为了进一步在准确性与延迟之间取得平衡，并避免朝其中一个指标崩溃（例如，良好的延迟但准确性较差），本文使用了一种精细的延迟正则化，以有效的权衡延迟和准确性。而且，为了更完整地兼顾准确率和性能，本文的 NAS 框架同时搜索复杂的教师网络和轻量级的学生网络结构，通过知识蒸馏得到高推理速度高准确率的分割网络。

4.2 可微分架构搜索

本章用到的 NAS 框架是基于可微分架构搜索 (DARTS) 方法, 它是一种基于梯度的方法 (Gradient-based method) [66]。基于强化学习和进化算法的方法本质上都是在离散空间中搜索。一个要搜索最优结构的块, 可以看作是包含 N 个有序结点的有向无环图。结点代表隐式表征 (例如特征图), 连接结点的有向边代表算子操作。DARTS 方法中最关键的技巧是将候选算子使用 softmax 函数进行混合。这样就将搜索空间变成了连续空间, 目标函数成为了可微函数。这样就可以用基于梯度的优化方法找寻最优结构了。搜索结束后, 这些混合的操作会被权重最大的操作替代, 形成最终的结果网络。搜索效率高是这种方法最大的优点。

4.3 ROI 分割网络架构搜索

本章针对分割脑部 MRI 中 ROI 的任务使用了一种多分辨率的网络架构搜索方法 NASBraNet, 它的搜索空间受启发于先前手动设计成功的网络并且具有多分支的特点, 优化方法基于梯度的可微分架构搜索方式, 评估方法使用分割损失和一种细粒度的延迟正则化, 以提高网络的推理速度。并且本文将 NASBraNet 扩展成一个师生共同搜索框架, 可以获得更轻量化更准确的学生网络

4.3.1 搜索空间

搜索空间核心目标是搜索具有整体低延迟的多分辨率分支, 多分辨率分支在先前的手动设计作品中显示出优秀的性能[33,34]。图 4-1 展示了本文的 NAS 框架搜索空间, 框架自动选择和合并不同分辨率的分支, 每个分支构成的基本单位是由可搜索超核连接的单元 (Cell)。各个分支对应着不同分辨率的输出, 这些输出在头模块中逐渐聚合在一起。每个单元可以单独搜索, 并且可以具有两个不同下采样率的输入和输出。对于单元间的单个超核, 还可以搜索其宽度扩展系数。

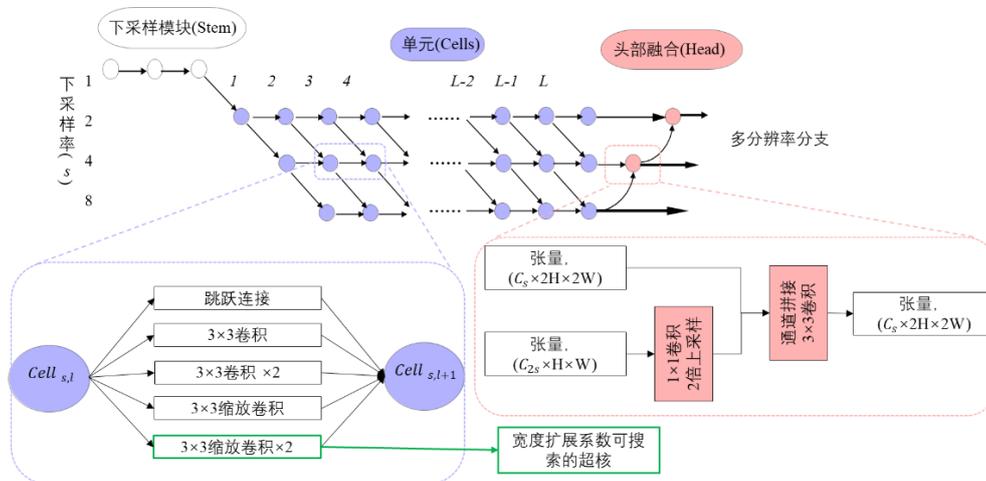


图 4-1 多分辨率分支搜索空间

(1) 可搜索的多分辨率分支

受到 Auto Deeplab 的启发, 本文在 L 层单元内搜索空间分辨率, 其中每个单元从两个连接的前驱单元获得输入并输出不同分辨率的两组特征图输出。由于手工设计的网络表明具有不同分辨率输出的多分支能改善网络性能, 因此本文的目标是在该 L 层框架中选择不同输出分辨率的 b ($b > 1$) 个分支。具体来说, 本文可以为最后一层单元选择 b 个不同的最终输出分辨率, 并通过回溯解码每个分支 (第 4.3.6 节)。这可以使本文的 NAS 框架能够探索具有不同分辨率的 b 个单独的分支, 这些分支由头部整合模块逐步合并。

网络设计遵循每次下采样时增加通道数量的惯例。为了扩大模型容量而不会产生太大的延迟, 首先使用 Stem 模块将输入图像缩小到原始尺度的 1/2, 然后设置可搜索的下采样率 $s \in \{2, 4, 8\}$ 。图 4-2 显示了本文的多分辨率搜索空间能够涵盖优秀的手工设计的网络以进行实时分割。

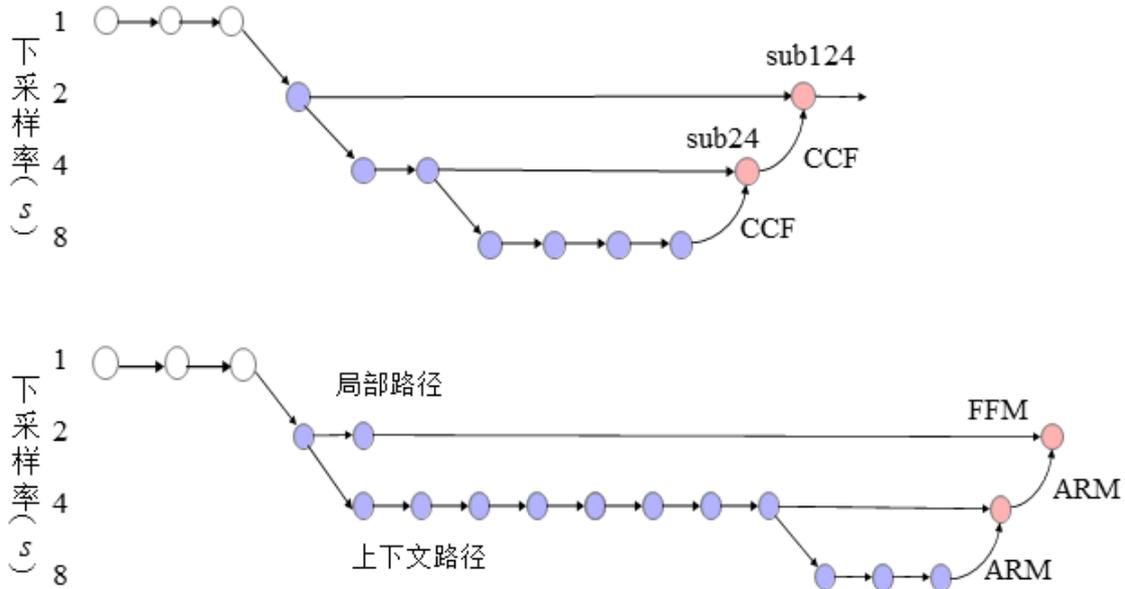


图 4-2 顶部: ICNET^[33] 网络结构 底部: Bisenet^[34] 网络结构

由于本文可搜索到的下采样率 $S \in \{2, 4, 8\}$ 和所选分支的数量 $b = 2$, 本文的超网络需要选择三种可能的分辨率组合: $\{2, 4\}$, $\{2, 8\}$ 和 $\{4, 8\}$ 。对于每个组合, 本文的头部模块将合并两种分辨率的分支。本文的超网络根据 Tan 等人提出的方法选择最佳分支组合^[82], 如式 4-1:

$$\text{Target}(m) = \text{ACC}(m) \times \left[\frac{\text{LAT}(m)}{T} \right]^w \quad (4-1)$$

其中 m 是整合 b 个分支搜索到的模型, $\text{ACC}(m)$ 是模型准确率, $\text{LAT}(m)$ 是延迟。 w 是权重因子, 定义如下:

$$w = \begin{cases} \alpha, & \text{if } \text{LAT}(m) \leq T \\ \beta, & \text{otherwise.} \end{cases} \quad (4-2)$$

本文设置 $\alpha = \beta = -0.07$ ，将目标延时 T 设定为 8.3 ms。

(2)选择高效且有更大感受野的算子

当本文的目标是提高推理速度时，选择超核的算子 O 的直接指标是执行算子的速度，而不是间接度量如浮点计算量等（FLOPs）。确保在空间上下文中的有足够大感受野也很重要。本文在表 4-1 中分析了主流的算子，包括它们的常见间接延迟度量（FLOPs，参数量等）以及与在安装 TensorRT 库的 NVIDIA 1080Ti GPU 上实际的测量延迟。

分组卷积通常用于减少 FLOPs 和参数数量^[84]。与标准卷积相比，分两组的卷积具有相同的感受野，但速度快 13%，同时参数量减少一半（由于它降低了模型学习能力而不是优选）。空洞卷积具有更大的感受野，并且在密集的预测任务中被广泛应用^[33,34]。然而，如表 4-1 所示，空洞卷积（空洞率 2）有着明显更高的延迟，尽管这不直接反映在 FLOPs 和参数量中。鉴于此，文献^[83]设计了一个名为“缩放卷积（zoomed convolution）”的新变种，其中输入顺序地进行双线性下采样，标准卷积和双线性上采样操作。与标准卷积相比，这种特殊设计使延迟降低了 40%的同时将感受野变为原来的 2 倍。因此，本文的搜索空间包括以下算子：

- 跳跃连接
- 3×3 卷积
- 3×3 卷积 $\times 2$
- 缩放卷积：双线性下采样 + 3×3 卷积 + 双线性上采样
- 缩放卷积 $\times 2$ ：双线性下采样 + 3×3 卷积 $\times 2$ + 双线性上采样

表 4-1：不同卷积算子的性能分析。延迟是使用带有 TensorRT 库的 1080Ti 尺寸 $1 \times 256 \times 32 \times 64$ 的输入测量。每个算子只包含一个卷积层。相对感受野(RF)是比较于标准卷积（第一行）得到。

算子类型	延迟 (ms)	FLOPs(G)	参数量(M)	相对感受野
卷积	0.15	1.21	0.59	1
分组卷积 (2 组)	0.13(-13%)	0.60(-50%)	0.29(-51%)	1
空洞卷积 (空洞率 2)	0.25(+67%)	1.10(-9%)	0.59	2
缩放卷积	0.09(-40%)	0.30(-75%)	0.15(-75%)	2

如 Ma 等人所述，网络碎片会显著妨碍并行度，降低实际效率^[85,86]。因此，本文选择顺序搜索空间，即卷积层在本文的网络中顺序堆叠。在图 4-1 中，每个单元都是可微分的并且如果导出离散架构后将仅包含一个算子^[30]。值得注意的是，每个单元在整个搜索空间上都可以单独搜索。

(3)宽度可搜索的超核

为了得到运行效率更高的模型，每个单元间的算子还可以选择不同通道数（扩展系数）的灵活性。在本文的工作中，本文搜索连续单元之间的连接宽度。然而，由于算子和宽度的可能存在指数级数量的组合，这使得宽度搜索变得难以进行。为了解决这个问题，文献^[83]提出了可以在一个支持一组扩展系数 $X \subseteq N^+$ 的卷积核中直接搜索

扩展系数 χ , 本文中 $X = \{16, 24, 32, 40, 48\}$ 。该方法将算子的通道数根据比率 χ 由少变多的变化, 转化为逐渐从一个通道数最多的超核中获取更大的子集。在架构搜索期间, 对于每个超核, 通过“Gumbel-Softmax”技巧, 在随机梯度下降的每个步骤中, 只采样、激活和反向传播一个扩展系数。这种设计可以得到一个简化且节省内存的超网络。

“Gumbel-Softmax”技巧详细来讲, 就是可以从一组扩展系数 $X \subseteq N^+$, 实现从 X 的随机采样得到一个扩展系数 χ , 并且这个采样过程是近似可微分的。对于每个扩展系数 χ^i 都会有一个相对应的概率 γ^i , 其中 $\sum_{i=1}^{|\chi|} \gamma^i = 1$ 。首先采样一个 Gumbel 噪声 $o^i =$

$$-\log(-\log(u)), u \sim \text{Unif}[0,1]。然后选择 j = \operatorname{argmax}_i \left(\frac{\exp\left(\frac{\log(\gamma^i) + o^i}{\tau}\right)}{\sum_{m=1}^{|\chi|} \exp\left(\frac{\log(\gamma^m) + o^m}{\tau}\right)} \right), 而不是 j =$$

$\operatorname{argmax}_i(\gamma^i)$ 。在本文实验中温度参数 $\tau = 1$ 。

为了遵循分辨率下采样增加通道数量的惯例, 在本文的搜索空间中, 宽度为 $\chi \times s$, 其中 $s \in \{2, 4, 8\}$ 。搜索方法允许每对连续单元之间的连接灵活地选择自己的扩展系数, 而不是在整个搜索空间中使用统一扩展系数。

4.3.2 搜索空间松弛连续化

本文综合权衡速度与精度, 选用的搜索策略为可微分架构搜索。于是需要对搜索空间进行松弛连续化。将下采样率记为 s , 层标号记为 l 。为了方便搜索空间结构, 将每个单元与前面两个不同下采样率的单元输出相连接:

$$\overline{I}_{s,l} = \beta_{s,l}^0 \overline{O}_{\frac{s}{2} \rightarrow s, l-1} + \beta_{s,l}^1 \overline{O}_{s \rightarrow s, l-1} \quad (4-3)$$

每个单元最多可以具有两个不同的下采样率的输出, 如式 3-4 和式 3-5, 进入其后继单元:

$$\overline{O}_{s \rightarrow s, l} = \sum_{k=1}^{|\mathcal{O}|} \alpha_{s,l}^k O_{s \rightarrow s, l}^k(\overline{I}_{s,l}, \chi_{s,l}^j, \text{stride} = 1) \quad (4-4)$$

$$\overline{O}_{s \rightarrow 2s, l} = \sum_{k=1}^{|\mathcal{O}|} \alpha_{s,l}^k O_{s \rightarrow 2s, l}^k(\overline{I}_{s,l}, \chi_{s,l}^j, \text{stride} = 2) \quad (4-5)$$

扩展系数 $\chi_{s,l}^j$ 是通过“Gumbel-Softmax”技巧根据 $p(\chi = \chi_{s,l}^j) = \gamma_{s,l}^j$ 采样得到。这里, α , β 和 γ 是归一化的标量, 分别对应着每个算子 $O^k \in \mathcal{O}$, 每个前驱单元输出 \overline{O}_{l-1} 和每一个扩展系数 $\chi \in X$, 通过这几个参数可以对架构进行优化和导出。

α , β 和 γ 是通过 softmax 进行归一化的方式如公式 4-6、4-7 和 4-8:

$$\sum_{k=1}^{|\mathcal{O}|} \alpha_{s,l}^k = 1, \forall s, l \text{ and } \alpha_{s,l}^k \geq 0, \forall k, s, l \quad (4-6)$$

$$\beta_{s,l}^0 + \beta_{s,l}^1 = 1, \forall s, l \quad \text{and} \quad \beta_{s,l}^0, \beta_{s,l}^1 \geq 0, \forall s, l \quad (4-7)$$

$$\sum_{j=1}^{|\chi|} \gamma_{s,l}^j = 1, \forall s, l \quad \text{and} \quad \gamma_{s,l}^j \geq 0, \forall j, s, l \quad (4-8)$$

其中 s 是下采样率， l 是超网络的层索引。

4.3.3 细粒度的延迟正则化优化

低延迟是对于实用网络是很重要的，但很具有挑战性。在搜索过程中，超空网或搜索策略经常陷入错误的“局部最优”，其生成的架构具有极低的延迟但精度差，特别是在搜索的早期阶段。此外，搜索的网络倾向于使用更多的跳过连接而不是选择低扩展系数^[87]。潜在的原因是，寻找具有极低延迟的架构（例如，简单的选择最轻量化的算子）显然比发现高精度的精巧结构容易。

为了解决这个问题，采用了更细粒度延迟正则化。本文首先通过使用相应的延迟替换公式 4-3 到公式 4-5 中的算子 O ，实现与单元中的算子相似的延迟的连续松弛。通过构建一个包含所有算子的延迟查找表，支持估计连续松弛后的延迟。通过累加在网络中所有延迟来估计网络延迟。

超网络的延迟对于算子 O ，下采样率 s 和扩展系数 χ 的有着不同敏感性。算子采用“ 3×3 卷积 $\times 2$ ”和“缩放卷积”时有一个巨大的延迟差。在窄和宽的膨胀比之间存在类似的延迟间隙（虽然更温和）。然而，像“2”和“8”这样的下采样率并不同，因为分辨率下采样也会带来输入和输出通道的数量加倍。若对这三者不加区分的采用同一个权重作为延迟正则项，很容易导致超网络搜寻到的网络结构崩溃，因此需要在训练时仔细的调参。

表 4-2 超网络对于不同变量类型造成的延迟变化的敏感度

变量	$\Delta\text{Latency (ms)}$
O (算子)	10.42
s (下采样率)	0.01
χ (扩展系数)	5.54

本文通过调整三者中的一个来定量地比较 O ， s 和 χ 对超网络延迟的影响。以 O 作为示例，本文首先均匀初始化 β 和 γ ，并计算选用最慢的运算符和最快运算符的超网络的时间差 $\Delta\text{latency}(O)$ ，输入尺寸为 $(3, 224, 224)$ 。对 s 和 χ 进行了类似的计算，得到表 4-2，表中的 $\Delta\text{latency}$ 的值表示超网络对算子而不是下采样率有着更高的灵敏度。

基于搜索空间的不同敏感度使用了一种正则化的延迟优化。本文将超网络的延迟解耦成搜索空间的三个方面 (O, s, χ)，并将每个方面乘以不同的系数进行标准化：

$$\begin{aligned} \text{Latency}(O, s, \chi) = & w_1 \text{Latency}(O | s, \chi) + w_2 \text{Latency}(s | O, \chi) \\ & + w_3 \text{Latency}(\chi | O, s) \end{aligned} \quad (4-9)$$

其中 w 表示各个延时约束的权重,根据表 2 中的不同敏感度,实验取 $w_1 = 0.001$,
 $w_2 = 0.997$, $w_3 = 0.002$ 。

4.3.4 教师/学生网络共同搜索

知识蒸馏是将一个大而复杂的网络(教师网络 T)学到的知识转移到一个小得多的网络(学生网络 S)中的有效方法。在本文的框架中,本文可以在一次运行中同时搜索两个网络。

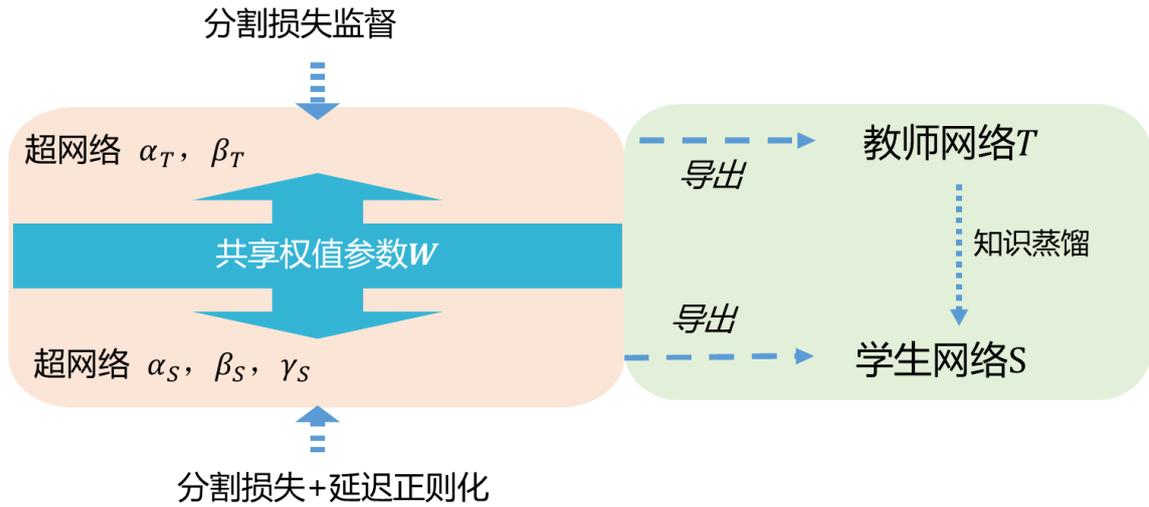


图 4-3 师生网络共同搜索框架

具体来说,如图 4-3 所示,本文通过在一个超网络中采用两组架构参数 (α_T, β_T) 和 $(\alpha_S, \beta_S, \gamma_S)$ 来同时搜索复杂的教师网络和轻量级学生网络。请注意,教师网络总是选择最宽的扩展系数,不会对其进行搜索。这种共同搜索方式不会带来任何额外内存使用或超网大小的开销,因为教师和学生搜索过程中共享相同的超网络权重 W 。两组架构在搜索过程中被反复优化,训练过程中只对学生应用延迟正则化。因此,本文搜索的教师网络是基于与学生网络使用的相同搜索空间和权重 W 的复杂网络。在学生网络从头开始的训练过程中,本文将教师网络 T 的一个蒸馏损失应用到学生网络 S 身上。

$$L_{\text{distillation}} = E_{i \in \mathbb{R}} \text{KL}(q_i^s | q_i^t) \quad (4-10)$$

KL 表示 KL 散度。 q_i^s 和 q_i^t 分别使用 S 和 T 预测像素 i 的特征。分割损失和蒸馏损失的权重相等。

4.3.5 网络结构参数更新过程

超网络定义为 \mathcal{M} ,架构搜索期间总的优化目标函数为:

$$L = L_{\text{seg}}(\mathcal{M}) + \lambda \cdot \text{Lat}(\mathcal{M}) \quad (4-11)$$

分割损失 L_{seg} 采用 FocalLoss。Lat(\mathcal{M})是超网络的估计延迟时间，而 λ 是平衡系数，在实际搜索中将 λ 设置为 0.01。

由于架构参数 α 、 β 和 γ 全部都在微分计算图里面，可以通过梯度下降法优化。与文献[30]中训练方法类似，将训练数据集按照 1:1 的比例随机分为两个不相交的集合训练集 A 和训练集 B，然后按照以下步骤进行优化。

在架构搜索实验中，首先对超网络进行预训练 20 轮，只更新权重参数 W 而不更新网络结构参数 α 、 β 、 γ ，然后开始搜索 30 轮的网络结构参数。

在训练集 A 求 $\nabla_{\alpha,\beta,\gamma}L_{\text{seg}}(\mathcal{M} | W, \alpha, \beta, \gamma)$ 更新超网络权重 W

在训练集 B 上求 $\nabla_{\alpha,\beta,\gamma}L_{\text{seg}}(\mathcal{M} | W, \alpha, \beta, \gamma) + \lambda \cdot \nabla_{\alpha,\beta,\gamma}LAT(\mathcal{M} | W, \alpha, \beta, \gamma)$ 更新网络结构参数 α 、 β 、 γ

当把超网络扩展为教师学生网络共搜索结构后，有教师和学生两组网络结构参数 (α_T, β_T) 和 $(\alpha_S, \beta_S, \gamma_S)$ ，网络的训练步骤便如下：

在训练集 A 上求 $\nabla_W L_{\text{seg}}(\mathcal{M} | W, \alpha_T, \beta_T)$ 更新网络权重 W

在训练集 A 上求 $\nabla_W L_{\text{seg}}(\mathcal{M} | W, \alpha_S, \beta_S, \gamma_S)$ 更新网络权重 W

在训练集 B 上求 $\nabla_{\alpha_T,\beta_T,\gamma_T} L_{\text{seg}}(\mathcal{M} | W, \alpha_T, \beta_T)$ 更新教师网络结构参数 α_T, β_T

在训练集 B 上求 $\nabla_{\alpha_S,\beta_S,\gamma_S} L_{\text{seg}}(\mathcal{M} | W, \alpha_S, \beta_S, \gamma_S) + \lambda \cdot \nabla_{\alpha_S,\beta_S,\gamma_S} LAT(\mathcal{M} | W, \alpha_S, \beta_S, \gamma_S)$ 更新学生网络结构参数 $\alpha_S, \beta_S, \gamma_S$

优化多个宽度：要在一个超核中训练多个宽度，在超网络中计算所有不同宽度选项的损失是不现实的。因此，本文通过训练最大和最小扩展系数来近似多宽度优化。在对超网络进行预训练期间，以最小宽度，最大宽度和 2 个随机宽度训练每个算子。在搜索网络结构参数时，本文训练每个运算符的最小宽度，最大宽度以及 Gumbel Softmax 当前采样宽度 γ 。

4.3.6 离散架构导出

搜索完成后，本文根据得到的网络结构参数 α 、 β 和 γ 导出离散架构：

α 、 γ : 本文通过取 α 和 γ 的 argmax 来选择最佳算子和扩展系数。

β : 对于每个单元 $cell_{s,l}$ ，将 β^0 和 β^1 分别视为两个前驱单元 $cell_{\frac{s}{2},l-1}$ 和 $cell_{s,l-1}$ 的输出进入 $cell_{s,l}$ 的概率。因此，通过取 $l^* = \text{argmax}_l (\beta_{s,l}^0)$ ，本文找到了对当前分辨率($\frac{s}{2} \rightarrow s$)进行下采样的最佳位置 $cell_{s,l}$ 。对于有着两个可搜索的下采样位置的分支，就应考虑的是联合概率的 argmax : $(l_1^*, l_2^*) = \text{argmax}_{l_1, l_2} (\beta_{s,l_1}^0 \cdot \beta_{2s,l_2}^0)$

值得注意的是，如果多分辨率分支的单元算子的类型、分辨率和扩展系数相同，那么它们会共享单元的权值和特征图，这种设计有助于加快网络推理速度。一旦分支中的单元分叉，分支之间的共享将被停止，它们变成单独的分支。

4.4 实验结果与讨论

为了验证本文提出算法的性能表现,先基于医院临床数据建立了分割数据集,将本文算法先从分割精度及运行时间两个层面与目前主流的分割算法展开对比,然后分析了本算法关键部分的作用,最后展示了算法在典型案例下的可视化结果。实验的运行软件环境是 Ubuntu16.04 和 Pytorch1.7,所用的电脑硬件配置主频为 4.0GHz 的 Intel Core i7 处理器, GPU 为 NVIDIA 的 GeForce GTX 1080Ti 显卡,在网络推理速度测试阶段使用了 TensorRT 加速。

4.4.1 数据集介绍

本研究来源于近三年(2017-2020)就诊于中南大学湘雅医院神经外科的患者,所有纳入研究对象的患者都经由术后病理诊断确诊神经胶质瘤,共包含 263 例患者数据,数据内容包含患者水平位两个模态(T1C, T2 Flair)MRI 图像。其中患者有 113 例男性,年龄分布范围是 13 岁到 75 岁,平均年龄是 45 岁。每例患者数据均包含两个模态各约 20 张图像。

为了排除颅骨等非脑部组织对算法的干扰,首先对头部 MRI 图像进行颅骨剥离得到脑部 MRI 图像,对肿瘤位置进行标注处理任务由经验丰富的专业医师团队执行,具体的标注软件是 LabelImg,将两个模态脑部 MRI 图像合并为形状为(224,224,2)的两通道图像。为了增强算法的泛化性能,对图像进行平移,旋转及翻转来进行数据增强,在输入网络前执行均值方差归一化操作。

4.4.2 网络架构搜索结果

超网络共有 $L=16$ 层,下采样率为 $\in \{2, 4, 8\}$ 。默认使用分支数量 $b=2$,因为更多的分支将会造成较高延迟。对于任意的下采样率 s 和层 l 的扩展系数为 $\chi_{s,l} \in X\{16,24,32,40,48\}$ 。多分辨率分支有千余种不同的路径。对于单元和扩展系数,本文有 $(1+4 \times 5)^{(15+14+13)} \approx 3.4 \times 10^{55}$ 不同的组合。于是总的搜索空间约为 10^{58} 数量级。

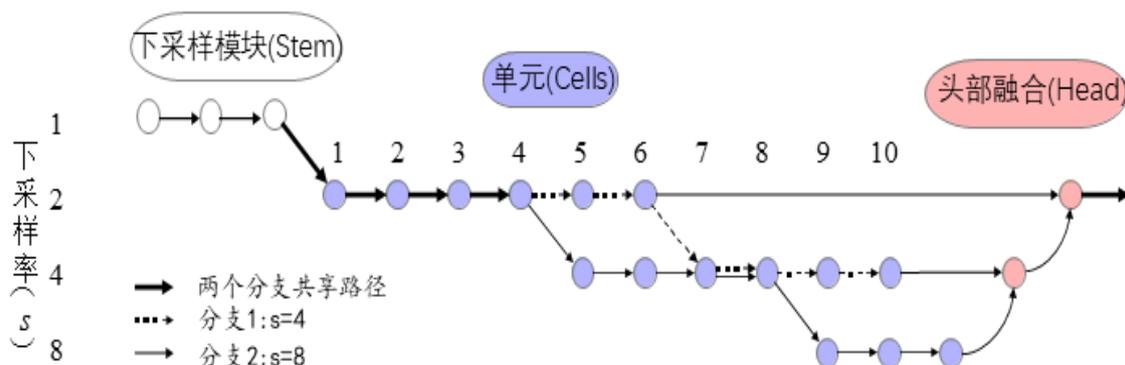


图 4-4 通过架构搜索得到的 NASBraNet 架构

图 4-4 显示了发现的最佳空间分辨率(NASBraNet)。本文的 NASBraNet 以适当的深度实现了多分辨率。这两个分支共享前三个运算符，然后分离，并选择整合下采样率为 4 和 8 的输出。表 4-3 和表 4-4 列出了 NASBraNet 两个分支的所有单元的参数，其中大量使用缩放卷积，这表明了低延迟和大感受野的重要性。

表 4-3 NASBraNet 中最终下采样率为 4 的分支使用的单元列表

单元序号	算子	下采样率	扩展系数	输出通道数
1	卷积×2	2	16	32
2	卷积×2	2	16	32
3	缩放卷积×2	2	16	32
4	缩放卷积×2	2	16	32
5	卷积×2	2	16	32
6	卷积×2	2	16	32
7	缩放卷积×2	4	16	64
8	缩放卷积×2	4	48	192
9	缩放卷积×2	4	32	128

表 4-4 NASBraNet 中最终下采样率为 8 的分支使用的单元列表

单元序号	算子	下采样率	扩展系数	输出通道数
1	卷积×2	2	16	32
2	卷积×2	2	16	32
3	缩放卷积×2	2	16	32
4	缩放卷积	2	32	64
5	缩放卷积×2	4	16	64
6	缩放卷积×2	4	16	64
7	缩放卷积×2	4	16	64
8	缩放卷积×2	4	16	64
9	缩放卷积×2	8	16	128
10	缩放卷积×2	8	32	256

4.4.3 算法性能评估

对于肿瘤区域，常用 DSC 分割准确率，定义如下：

$$DSC(P, T) = \frac{|P_1 \cap T_1|}{(|P_1| + |T_1|)/2} \quad (4-12)$$

其中 P 代表模型预测，T 代表真值标签。T1 和 T0 是 T 的子集，表示所讨论的肿瘤区域的阳性和阴性。对于 P1 和 P0 也是如此。不难发现，如果得到的 DSC 值越大，则说明目标分割越准确，最理想的值是 1，此时预测范围与真实范围完全重叠。

考虑到本研究收集的数据量并不算丰富，为了更完整的验证网络性能采用了五折交叉验证，表 4-5 展示了其验证的平均分割 DSC 和推理速度。

表 4-5 与主流分割方法测试 DSC 和推理速度对比

方法	DSC (%)	推理速度 (Hz)	分辨率
UNet ^[15]	75.6	91.7	224×224
ICNet ^[33]	88.8	56.4	224×224
NAS-UNet ^[81]	86.5	82.9	224×224
NASBraNet (本文方法)	89.4	178.3	224×224

从表中明显可以看到通过本文的网络架构搜索方法得到的 NASBraNet 模型，在准确率和推理速度上都要高于主流的 UNet 等网络，特别是在推理速度上能达到同样使用网络架构搜索得到的 NAS-UNet 的两倍以上，这证明了该方法在临床数据中的实用性。

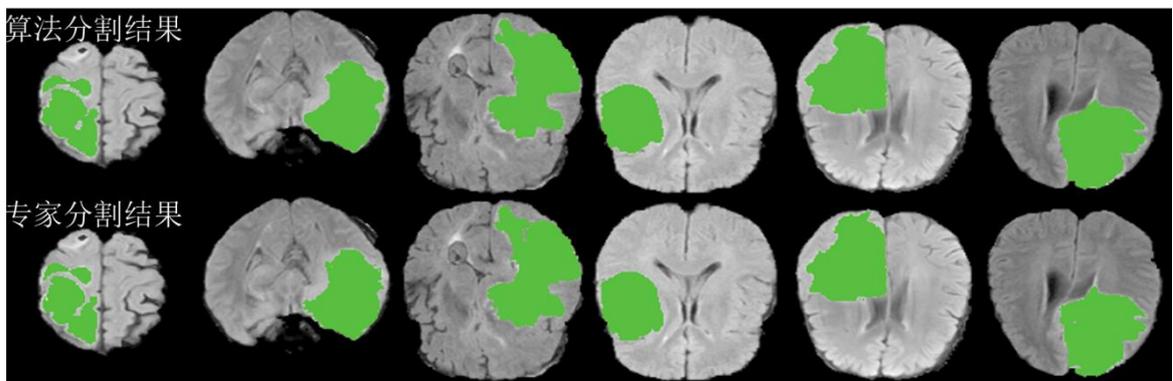


图 4-5 部分患者分割结果与真值对比示例，

对部分患者测试图像进行可视化得到图 4-5，上面一行为算法分割结果，下面一行为多名专业医师确认的分割结果，两者整体重合度较高。由图可知，本文方法能够对不同病人的脑肿瘤在位置、大小、形状等方面存在较大差异的情况下有一个较理想的分割结果，可用于后续分子分型预测等进一步精细化的诊断。

4.4.4 分析与讨论

为了评估所提 NAS 框架的有效性，本文还进行了消融研究。更具体地说，本文考察了算子(O)、下采样率(s)、扩展系数(χ)、分支数(b)以及知识蒸馏对精度和推理速度的影响，实验的输入大小为 224×224 。在表 3-6 中，当本文从单分支($b=1$)扩展到多分支($b=2$)时，本文的推理速度下降了，但本文在分割 DSC 上获得了很大的改善，这表明多分辨率设计对分割任务是有利的。通过启用搜索扩展系数(χ)，本文发现在不牺牲精确度(70.5%)的情况下，可以得到更快速的网络，这证明了可搜索超核在保持高精度的同时从消除冗余通道中获益。

表 4-6 不同搜索和训练策略的消融实验结果。

设置	DSC(%)	推理速度(Hz)	FLOPs(G)	参数量(M)
$O, s \chi = 32, b = 1$	82.2	193.6	13.5	3.1
$O, s \chi = 32, b = 2$	86.4	132.4	21.4	5.4
教师网络 (T) 学生网络 (S) 共同搜索				
$O, s, \chi b = 2$	86.7	178.3	14.0	2.2
$T \rightarrow T$ 剪枝	81.6	158.9	14.7	2.3
$T \rightarrow S$ (NASBraNet)	89.4	178.3	14.0	2.3

注：“ \rightarrow ”：知识蒸馏

本文进一步评估了本文的师生共同搜索框架的有效性。在共同搜索之后，本文得到了教师架构(T)和学生架构(S)。如上所述，使用可搜索扩展系数(χ)搜索S，获得 163.9 的 FPS 和 70.5%的 DSC。相比之下，当本文直接对教师进行压缩(通过选择最窄的扩展系数进行通道剪枝)，并从训练好的教师模型通过知识蒸馏进行训练时，在推理速度仅为 158.9 的情况下，DSC 为 81.6%，这表明本文的共同搜索框架结果优于基于剪枝的压缩。最后，当本文采用从训练好的教师网络将知识蒸馏到搜索到的学生时，将学生的准确率提高到了 89.4%，这就是本章最终的网络 NASBraNet。这表明，通过共同搜索发现的学生架构和通过教师的知识蒸馏进行训练对于获得准确、快速的实时分割模型都是至关重要的。

4.5 本章小结

在本章中，为了解决胶质瘤分子标记物预测方法通常需要人工勾画 ROI 的问题，基于临床胶质瘤分割数据集提出了一种名为 NASBraNet 的语义分割网络架构搜索框架，目的是实现快速准确的自动 ROI 分割。本章节针对临床胶质瘤 MRI 影像分割需要兼顾准确率和速度的需求，使用了一种高性能多分辨率网络架构搜索框架，基于常用的单分支和算子搜索空间添加了多分支选择机制能搜索到较高的分割准确率的教师网络。通过引入扩展系数搜索和延迟正则化，实现了高效率学生网络的架构搜索。并通过教师学生网络共同搜索和知识蒸馏将教师网络的知识有效地迁移到了学生网络上，最终得到了应用于临床胶质瘤 MRI 影像任务的高效分割网络，并在综合性能上优于主流算法。

第 5 章 胶质瘤 Ki67 指数预测

5.1 引言

Ki67 指数在临床上可用于衡量肿瘤的增殖程度、预测病人生存期和复发情况。一般来说, 高的 Ki67 指数对应着更短的生存期和更高的复发率^[4,5], 因此, 不同的 Ki67 指数往往对应着不同的治疗方案, 若能在手术前得到 Ki67 指数将具有很大的指导意义^[6]。Ki67 指数的确定需要通过对肿瘤组织取样, 使用 Ki67 抗体对其进行染色, 并计算被染色细胞的比例^[7]。然而, 获取富含肿瘤的组织样本以计算 Ki67 指数可能是一项挑战。癌症基因组图谱(TCGA)的一份报告表明, 只有 35%的活检样本含有足够的肿瘤成分, 可以进行准确的分子特征鉴定^[8]。开发一个非侵入式的 Ki67 指数鉴定方法将对患者的护理更加有利。

最主流的非侵入式胶质瘤诊断方法是多参数核磁共振成像 (MRI), 专业医师肉眼对比 T1C 和 T2 Flair 两个模态的 MRI 图像可以方便地确定肿瘤大小、位置等, 但是难以分辨不同 Ki67 指数的肿瘤在 MRI 图像上的差异。

近年来, 基于非侵入式机器学习方法确定分子标记物的发展非常迅速。但关于 Ki67 的预测研究较少, 目前仅观察到邢福勇等人基于染色后的病理图像设计了一个弱监督的语义分割网络用于预测肿瘤的 Ki67 指数^[62], 尚未发现利用非侵入式信息预测 Ki67 指数的研究。

报道过的胶质瘤分子标记物预测方法通常需要人工勾画用于预测算法所需的感兴趣区域 (ROI, 一般是肿瘤分布区域), 这限制了临床上胶质瘤分子标记物预测算法的应用。

因此在本研究中, 首先对原始 MRI 影像利用第 4 章中的分割算法提取 ROI 后, 合成用于分子标记物预测的融合图像。随后一个新的多模态深度学习网络 (KiNet) 被提出用于非侵入式地预测 Ki67 指数。结合第 4 章的分割算法, KiNet 可以在无医生勾画肿瘤区域的条件下有效融合多模态 MRI 和异构元数据 (年龄、性别) 等信息, 首次实现 Ki67 指数非侵入式预测。

5.2 数据集预处理

这项回顾性研究的数据来源于近三年 (2017-2020) 就诊于中南大学湘雅医院神经外科的患者, 所有纳入研究对象的患者都经由术后病理诊断确诊神经胶质瘤, 共包含 263 例患者数据, 数据内容包含患者基本信息元数据 (性别、年龄)、水平位两个模态 (T1C, T2 Flair) MRI 图像、术后病理化验 Ki67 指数结果 (根据 WHO 的指南, 本文将 Ki67 指数低于 10%称为低 Ki67 指数, 标记为 0, 其余的称为高 Ki67 指数, 标记为

1^[15])。其中患者有 113 例男性, 年龄分布范围是 13 岁到 75 岁, 平均年龄是 45 岁。高 Ki67 指数和低 Ki67 指数分别有 140 和 123 例, 每例患者数据均包含两个模态各约 20 张图像和元数据。

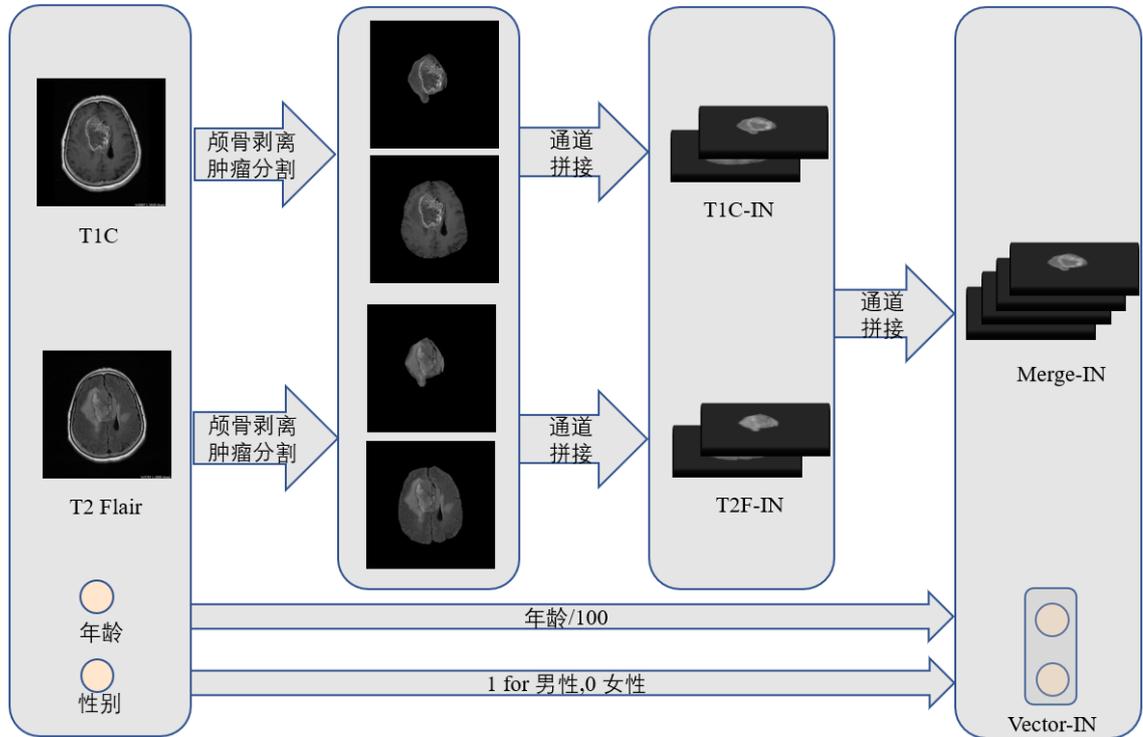


图 5-1 数据预处理流程

数据预处理流程如图 5-1 所示, 为了排除颅骨等非脑部组织对算法的干扰, 首先对头部 MRI 图像进行颅骨剥离得到脑部 MRI 图像, 并通过肿瘤分割算法分割出了肿瘤区域, 将各模态脑部 MRI 图像和分割出的肿瘤图像合并为形状为 (224,224,2) 的两通道图像, 两个模态的两通道图像分别称为 T1C-IN,T2F-IN。除此之外, 将 T1C-IN 和 T2F-IN 的图像合并为形状为 (224,224,4) 的四通道图像, 称为 Merge-IN。为了增强算法的泛化性能, 对图像进行平移, 旋转及翻转来进行数据增强, 在输入 KiNet 前执行均值方差归一化操作。对于患者的基本信息元数据, 将性别和年龄编码为 $2 * 1$ 的向量 (性别男记为 0, 性别女记为 1; 年龄除 100 转化成 0-1 之间的浮点数) 与图像数据构成异构数据作为 KiNet 最终的输入。

5.3 多任务多模态深度神经网络模型

在基于医学图像进行相关研究时, 多模态的方法已经作为一个关键的技术。多模态数据的融合能够为模型决策提供更多的相关信息, 其目的是处理和关联来自多个不同模态的不同层次信息特征 (如本文脑部胶质瘤的多参数 MRI 生成的多模态图像), 带来优于单一模态的模型预测决策。

5.3.1 多任务多模态深度神经网络模型结构

在本文中,受 CentralNet 模型思想的启发, KiNet 被提出用于 Ki67 指数预测,。如图 5-2 所示,模型包含了两个独立的单模态附加分支以及一个主分支,三个分支的目标都是预测 Ki67 指数。每个分支都是一个独立的分类器,主干结构与 ResNet 相同,都包含四个特征提取阶段,按照惯例四个阶段空间分辨率逐步缩小特征图通道逐步增加,第四个阶段后通过全局平均池化和全连接层输出分类结果。特别的,对于主分支而言,在每个阶段输出的特征图后都会添加全连接层获取当前分辨率尺度的分类结果,在最后通过通道拼接融合多尺度分类结果。

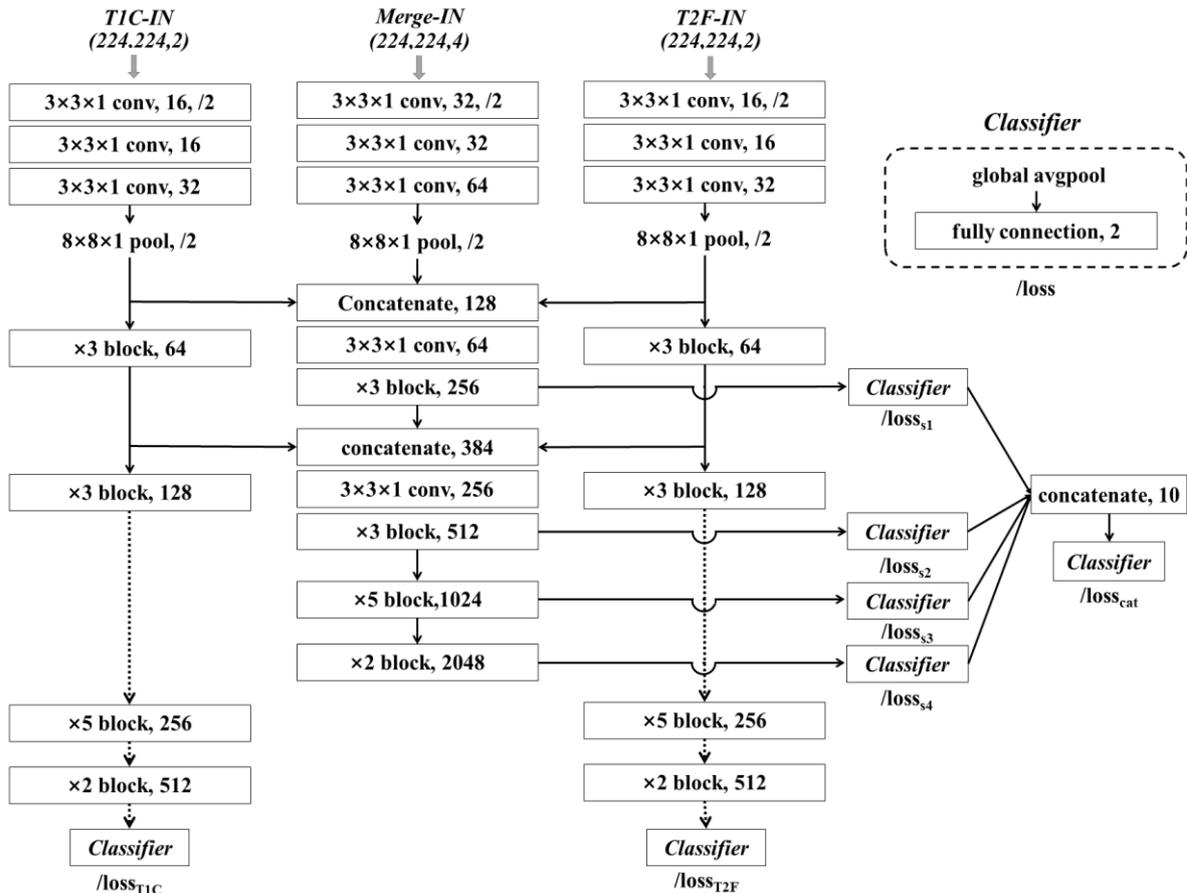


图 5-2 KiNet 整体网络结构

图 5-2 展示了模型的整体框架, TIC-IN, T2F-IN 和 Merge-IN 经浅层特征提取器后的特征图经通道拼接(Concatenate)后作为主分支的输入。主分支第二个阶段的输入是由前一个阶段输出与附加分支模型的第一个阶段的输出按通道拼接得到。将主分支多尺度输出与异构元数据向量通道拼接后经过一个全连接层得到最终的决策。

同样的,各个分支的每个特征阶段分别由若干个构建块组成,如图 5-3 所示,是模型基本构建块的结构,基本构建块 block 采用了与 ResNeSt^[88]相同的分离注意力构建块,与现有的 ResNet 变体相比,不需要增加额外的计算量,却能够获得更好的分

类性能，关于 ResNeSt 和构建块的详细解释可以参照文献[88]。

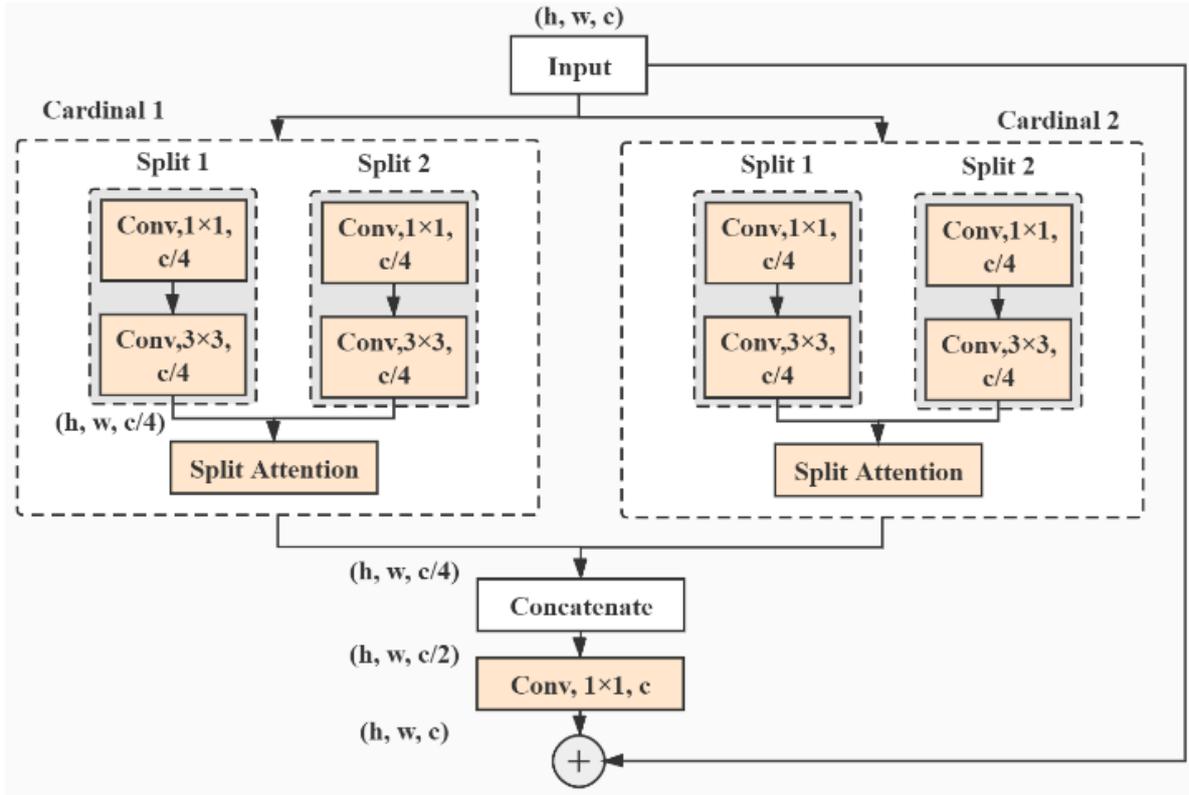


图 5-3 KiNet 中的使用的卷积块结构

5.3.2 损失函数设计

本文构建了一个多任务多模态网络用于 Ki67 指数分类，包括了独立附加分支的预测任务以及主分支的多尺度预测任务。接下来本文介绍网络训练阶段的损失函数。

本文将 Ki67 指数预测任务作为一个二分类任务，选用了交叉熵损失函数，如公式 3-1 所示：

$$loss = -(y \log(p) + (1 - p) \log(1 - p)) \quad (5-1)$$

主分支的损失函数被定义为：

$$loss_{master} = \beta_1 \times loss_{s1} + \beta_2 \times loss_{s2} + \beta_3 \times loss_{s3} + \beta_4 \times loss_{s4} + \beta_5 \times loss_{cat} \quad (5-2)$$

总损失函数 Loss 被定义为：

$$Loss = \alpha_1 \times loss_{T1C} + \alpha_2 \times loss_{T2F} + \alpha_3 \times loss_{master} \quad (5-3)$$

其中， $loss_{T1C}$ 、 $loss_{T2F}$ 是两个独立辅助分支的损失， $loss_{s_i}$ 是主分支中的对应的第 i 个 stage 输出的交叉熵损失， $loss_{cat}$ 是主分支多尺度分类结果与异构元数据拼接操作后的输出交叉熵损失，在本文的实验中， $\alpha_i (i = 1, 2, 3) = 0.35, 0.25, 0.4$ ， $\beta_j (j = 1, 2, 3, 4, 5) = 0.025, 0.025, 0.2, 0.35, 0.4$ 。

5.4 实验结果与讨论

5.4.1 算法评估准则

kappa 系数是用在统计学中评估一致性的一种方法，也可以用来进行分类模型准确度的评估，这个系数的取值范围是 $[-1,1]$ ，实际应用中，一般是 $[0,1]$ 。这个系数的值越高，则代表模型实现的分类准确度越高，当大于 0.4 时一般可以认为模型与现实的一致性较高。

5.4.2 与主流算法的实验结果对比

实验基于整理得到的数据集，采用了五折交叉验证方法，验证了 KiNet 和主流网络架构在 Ki67 指数预测任务上的性能对比，其中 ResNet, ResNeSt 都采用不同模态的图像数据经由通道叠加的早期融合方式，即网络的输入是 Merge-IN

表 5-2 给出了 KiNet 和主流网络架构在 Ki67 指数预测任务上的性能对比，其中 ResNet, ResNeSt 都采用不同模态的图像数据经由通道叠加的早期融合方式，即网络的输入是 Merge-IN。图 5-4 展示了不同神经网络模型的 ROC 曲线。

表 5-2 不同神经网络对 Ki67 指数的预测性能

模型	准确率	敏感度	特异度	AUC	Kappa
ResNet-50	0.691	0.693	0.689	0.724	0.422
ResNeSt-50	0.724	0.734	0.715	0.764	0.451
CentralNet	0.725	0.745	0.705	0.762	0.450
KiNet (本文)	0.752	0.755	0.749	0.788	0.472

由表 5-2 中数据和图 5-4 可以看出，CentralNet 中通过使用中心网络融合两个附加分支多层次特征相比简单的模态通道叠加，对于 Ki67 指数预测任务效果更好。KiNet 提出的网络架构更有利于不同层次特征信息之间的关联性表达，使其在本任务中表现胜过主流网络。

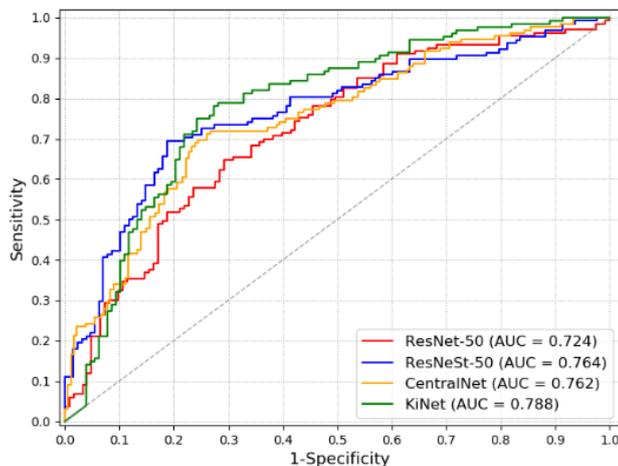


图 5-4 不同网络模型在 Ki67 指数预测任务的 ROC 曲线

5.4.3 消融实验

为了直观展示 KiNet 各个子模块的作用，本文设计了消融实验来验证 KiNet 中的子模块（多尺度损失，Merge-IN，元数据）对模型性能的影响，表 5-3 给出了 KiNet 的消融实验的结果。

表 5-3 KiNet 消融实验的预测性能

序号	模型	多尺度损失	Merge-IN	元数据	准确率	AUC
1	KiNet	0	0	0	73.5	0.766
2	KiNet	0	0	1	73.6	0.768
3	KiNet	0	1	0	74.2	0.773
4	KiNet	0	1	1	74.3	0.774
5	KiNet	1	0	0	74.6	0.778
6	KiNet	1	0	1	74.8	0.781
7	KiNet	1	1	0	75.0	0.785
8	KiNet	1	1	1	75.2	0.788

注：1/0 表示有/无

从表格数据中可以看出，实验 8 具有最好的预测性能，再次验证 KiNet 中网络结构设计的合理性。实验 1 和 5（2 和 6、3 和 7、4 和 8）反映了 KiNet 中设计多尺度损失的有效性，不同尺度的损失会表现出很多隐藏的图像噪声以及特征信息，对于模型鲁棒性的提升以及特征学习都有着很大的帮助。经过多次的调参实验，最终确定了多尺度损失权重 β_1 到 β_5 分别为 0.025,0.025,0.2,0.35,0.4 时网络能取得最好的性能，这在某种程度上说明深层特征对于模型预测结果的影响更大。比较实验 1 和 3（2 和 4、5 和 7、6 和 8）的结果，KiNet 中的添加额外的 Merge-IN 能为 Ki67 指数预测带来帮助。Merge-IN 通过直接组合原始图像，能有效实现不同模态间融合，有利于模型学习到不同模态图像的位置、灰度等浅层特征信息关联性。实验 1 和 2（3 和 4、5 和 6、7 和 8）中能够看出，包含病患的性别以及年龄在内的异构元数据带来的性能提升小于 0.2%，这表明针对于神经胶质瘤患者的 Ki67 指数预测这一具体任务，患者的性别和年龄对本文的模型性能提升较小，这可能是因为神经胶质瘤在人群中的易感人群范围很广，性别因素对神经胶质瘤的增殖能力影响不大导致。

5.4.4 分析与讨论

本文中主分支结构的最终确立经过了很多探索性实验，受到 CentralNet 模型的先入为主的影响，本文首先考虑表征融合方式是带权重的逐元素相加操作，其中权重是可学习的参数。具体的，针对 Ki67 指数预测任务，逐元素相加的权重被初始化为均匀分布，对于训练完的模型作参数可视化的结果如图 5-5，显示的是利用 T1C 和 T2Flair 两个模态训练 CentralNet 实验中，CentralNet 中三个子分支各自的累加权重，

权值可以认为是对中心网络的贡献度。可以看到主分支随着网络加深,贡献越来越大。到第三层和第四层,分支网络贡献很小。

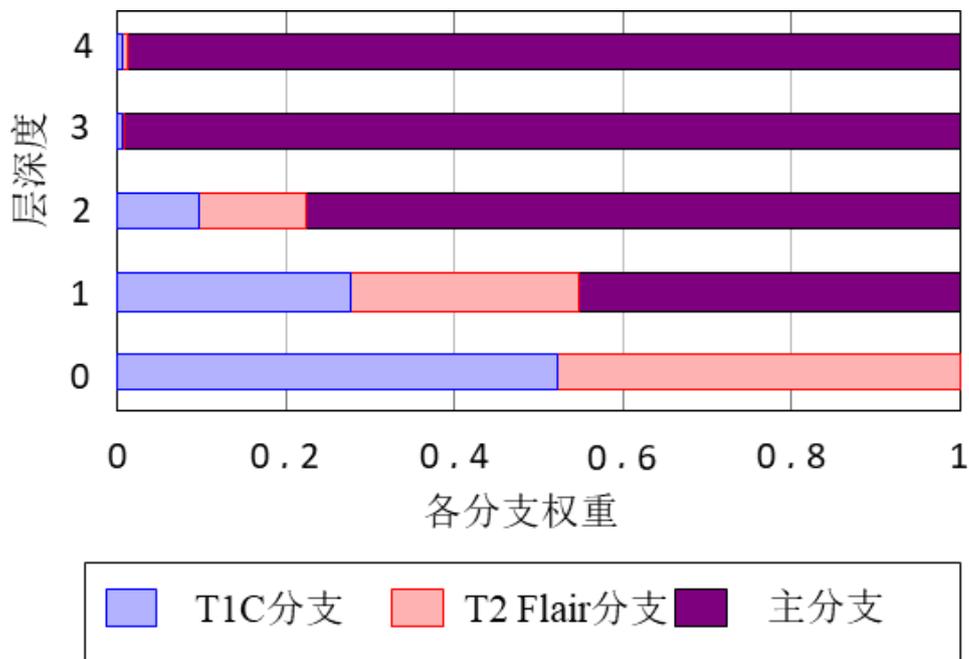


图 5-5 训练完成的 CentralNet 模型的模态融合参数可视化

通过对各分支权重定量分析,主分支结构中的第一个阶段中的权重比接近 1:2:1 (附加分支:主分支:附加分支);第二个阶段中的权重比接近 1:4:1 (附加分支:主分支:附加分支);而对于最后两个阶段,两个附加分支的输入权重小于 0.02。基于此,本文尝试用通道拼接来代替主元素相加操作。对主分支的第一个阶段输入作 1:2:1 的通道拼接操作,第二个阶段的输入作 1:4:1 的通道拼接操作,至于最后两个阶段的输入,本文摒弃了对附加分支的联结来精简本文的网络。而对于每个分支的损失权重,分别在 T1c 和 T2 Flair 两个模态图像的独立附加分支上预测 Ki67 指数,经多次实验确定 T1c 图像效果会高出 T2 Flair 2%,因此对于 T1c 附加分支设定了更大的损失权重,经多次调试参数后,最终将损失权重 α_1 到 α_3 设定为 0.35,0.4,0.25 (T1C 附加分支,主分支,T2 Flair 附加分支)。相比于使用逐元素相加操作进行多模态特征融合的网络,KiNet 的融合方式能带来超过 1.5%的性能提升。

5.5 本章小结

在本章中,为了解决胶质瘤分子标记物预测方法通常需要人工勾画 ROI 的问题,一个新的多模态深度学习网络(KiNet)被提出用于基于多模态 MRI 和异构元数据(年龄、性别)等信息非侵入式地预测胶质瘤的 Ki67 指数。网络使用两个附加分支对两个模态 MRI 图像进行特征提取,在主分支中融合多层次附加分支特征和异构元数据,并引入了多尺度损失监督网络训练过程。网络首次实现了无需标注先验 ROI 的非侵入

式预测胶质瘤的 Ki67 指数，取得了 AUC 0.78， Kappa 系数为 0.47 的结果。通过该方法，医生有希望在手术前获得胶质瘤的 Ki67 指数，从而制定更有效的治疗方案。

第 6 章 总结与展望

6.1 论文总结

分子标记物可以确定胶质瘤的分子分型，基于肿瘤遗传学的分子病理能够更准确地判断胶质瘤患者临床预后，并且对组织学难以明确诊断和分级的肿瘤提供鉴别依据。胶质瘤的分子标记物需要对切除的肿瘤组织做病理学化验才能确定，某些分子突变甚至需要基因测序才能够确定，化验周期可能长达几十天，基因测序的价格动辄上万，这对于早期手术方案的制定极为不利。若能利用非侵入式的信息，在术前得出肿瘤的分级和分子分型，将对治疗方案的有极大的指导意义。

考虑到现有关于分子特征预测算法的研究，大多都需要勾画肿瘤区域作为先验知识。为了在无专家知识的情况下非侵入式地预测胶质瘤分子分型，本文将分子分型预测方法拆分为基于网络架构搜索的脑部 MRI 影像分割算法和基于多模态神经网络的分类算法。

在预测 IDH 突变状态任务中，探究了模型的输入图像方式(是否需要勾画 ROI)、分类网络骨架和多模态信息融合位置对 IDH 预测准确率的影响，结果证明了 ROI (脑肿瘤区域) 作为附加输入对于预测分子分型的重要性。设计了多种多模态融合网络用于预测 IDH 突变状态，最佳模型能取得 AUC 为 0.87 的结果，这比主流的融合方式性能更好，同时能容忍一定程度的模态不匹配。

在临床胶质瘤 MRI 影像分割领域，针对临床胶质瘤 MRI 影像分割需要兼顾准确率和速度的需求，使用了一种高性能多分支网络架构搜索框架，通过多分支和高效的算子搜索空间能搜索到较高的分割准确率的教师网络。通过引入扩展系数搜索和延迟正则化，实现了高效率学生网络的架构搜索。并通过教师学生网络共同搜索和知识蒸馏将教师网络的知识有效地迁移到了学生网络上，最终得到了应用于临床胶质瘤 MRI 影像任务的高效分割网络。

在基于多模态神经网络的分类这一研究领域，一个新的多模态深度学习网络 (KiNet) 被提出用于基于多模态 MRI 和异构元数据 (年龄、性别) 等信息非侵入式地预测胶质瘤的 Ki67 指数。网络使用两个附加分支对两个模态 MRI 图像进行特征提取，在主分支中融合多层次附加分支特征和异构元数据，并引入了多尺度损失监督网络训练过程。本文首次实现了非侵入式预测胶质瘤的 Ki67 指数，取得了 AUC 0.9, Kappa 系数为 0.5 的结果。通过该方法，医生有希望在手术前获得胶质瘤的 Ki67 指数，从而制定更有效的治疗方案。

6.2 研究展望

本文对基于非侵入式信息的胶质瘤分子分型做出了较深入的研究，并取得了一定的进展，但是仍然存在很多问题需要进一步改进和探讨：

(1) 本文的主题是非侵入式的胶质瘤分子分型预测，临床对于胶质瘤的分子分型需要医生综合考虑多种分子标记物才可以得到较准确的分子分型，限于时间和实际临床数据集的限制，并且考虑到其他分子标记物的预测已经有报道，文章焦点仅集中在了某几种分子标记物（IDH 突变状态，Ki67 指数）的预测，通过现有的研究仍无法端到端得到胶质瘤的分子分型，这可通过收集多种分子标记物的数据和参考已报导的方法得到多种分子标记物，并结合医生的临床知识写入系统可做到全自动的胶质瘤分子分型系统。

(2) 对于本文提出的胶质瘤分割网络结构搜索框架而言，尽管在医院的临床数据集上对比一些主流网络取得了较好的效果，但对于算法的泛化性能却未得到验证，后续我们将在开源的数据集合上验证算法性能，并且争取收集到不同医院的数据进行多中心验证才可以确定算法的稳定性

(3) 本文提出的基于多模态神经网络的分子标记物预测方法，虽然相较于主流方法有着更好的性能，但是准确率仍然不算很高，后续可以考虑结合更多的医生知识和定量的临床特征。

参考文献

- [1] Ricard D, Idhahane A, Ducray F, et al. Primary brain tumours in adults[J]. *The Lancet*, 2012, 379(9830): 1984-1996.
- [2] Eckel-Passow J E, Lachance D H, Molinaro A M, et al. Glioma groups based on 1p/19q, IDH, and TERT promoter mutations in tumors[J]. *New England Journal of Medicine*, 2015, 372(26): 2499-2508.
- [3] Xing F, Su H, Neltner J, et al. Automatic ki-67 counting using robust cell detection and online dictionary learning[J]. *IEEE Transactions on Biomedical Engineering*, 2013, 61(3): 859-870.
- [4] Arshad H, Ahmad Z, Hasan S H. Gliomas: correlation of histologic grade, Ki67 and p53 expression with patient survival[J]. *Asian Pac J Cancer Prev*, 2010, 11(6): 1637-1640.
- [5] Torp S H, Johannesen E, Lindboe C F. Comparison of different Ki67 antibodies in human glioblastomas[J]. *Clinical molecular pathology*, 1995, 48(4): M191.
- [6] Li L T, Jiang G, Chen Q, et al. Ki67 is a promising molecular target in the diagnosis of cancer[J]. *Molecular medicine reports*, 2015, 11(3): 1566-1572.
- [7] Reid M D, Bagci P, Ohike N, et al. Calculation of the Ki67 index in pancreatic neuroendocrine tumors: a comparative analysis of four counting methodologies[J]. *Modern pathology: an official journal of the United States and Canadian Academy of Pathology, Inc*, 2016, 29(1): 93.
- [8] Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human glioblastoma genes and core pathways[J]. *Nature*, 2008, 455(7216): 1061.
- [9] Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis[J]. *European journal of cancer*, 2012, 48(4): 441-446.1
- [10] Lenting K, Verhaak R, Ter Laan M, et al. Glioma: experimental models and reality[J]. *Acta neuropathologica*, 2017, 133(2): 263-282.
- [11] Kwon D, Niethammer M, Akbari H, et al. PORTR: Pre-operative and post-recurrence brain tumor registration[J]. *IEEE transactions on medical imaging*, 2013, 33(3): 651-667.
- [12] 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.
- [13] Jégou S, Drozdal M, Vazquez D, et al. The one hundred layers tiramisu: Fully

- convolutional densenets for semantic segmentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 2017: 11-19.
- [14]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, 2017, 40(4): 834-848.
- [15]Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015: 234-241.
- [16]Çiçek Ö, Abdulkadir A, Lienkamp S S, et al. 3D U-Net: learning dense volumetric segmentation from sparse annotation[C]//International conference on medical image computing and computer-assisted intervention. Springer, Cham, 2016: 424-432.
- [17]Kamnitsas K, Ledig C, Newcombe V F J, et al. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation[J]. Medical image analysis, 2017, 36: 61-78.
- [18]Kamnitsas K, Bai W, Ferrante E, et al. Ensembles of multiple models and architectures for robust brain tumour segmentation[C]//International MICCAI brainlesion workshop. Springer, Cham, 2017: 450-462.
- [19]Myronenko A. 3D MRI brain tumor segmentation using autoencoder regularization[C]//International MICCAI Brainlesion Workshop. Springer, Cham, 2018: 311-320.
- [20]Isensee F, Kickingereder P, Wick W, et al. No new-net[C]//International MICCAI Brainlesion Workshop. Springer, Cham, 2018: 234-244.
- [21]Zoph B, Le Q V. Neural architecture search with reinforcement learning[J]. arXiv preprint arXiv:1611.01578, 2016.
- [22]Zoph B, Vasudevan V, Shlens J, et al. Learning transferable architectures for scalable image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 8697-8710.
- [23]Liu C, Zoph B, Neumann M, et al. Progressive neural architecture search[C]//Proceedings of the European conference on computer vision (ECCV). 2018: 19-34..
- [24]Pham H, Guan M, Zoph B, et al. Efficient neural architecture search via parameters sharing[C]//International Conference on Machine Learning. PMLR, 2018: 4095-4104.
- [25]Bender G, Kindermans P J, Zoph B, et al. Understanding and simplifying one-shot architecture search[C]//International Conference on Machine Learning. PMLR, 2018:

- 550-559.
- [26] Cai H, Zhu L, Han S. Proxylessnas: Direct neural architecture search on target task and hardware[J]. arXiv preprint arXiv:1812.00332, 2018.
- [27] Liu H, Simonyan K, Yang Y. Darts: Differentiable architecture search[J]. arXiv preprint arXiv:1806.09055, 2018.
- [28] Xu Y, Xie L, Zhang X, et al. PC-DARTS: Partial channel connections for memory-efficient architecture search[J]. arXiv preprint arXiv:1907.05737, 2019.
- [29] Chen L C, Collins M D, Zhu Y, et al. Searching for efficient multi-scale architectures for dense image prediction[J]. arXiv preprint arXiv:1809.04184, 2018.
- [30] Liu C, Chen L C, Schroff F, et al. Auto-deeplab: Hierarchical neural architecture search for semantic image segmentation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 82-92.
- [31] Weng Y, Zhou T, Li Y, et al. NAS-Unet: Neural architecture search for medical image segmentation[J]. IEEE Access, 2019, 7: 44247-44257.
- [32] Zhu Z, Liu C, Yang D, et al. V-nas: Neural architecture search for volumetric medical image segmentation[C]//2019 International Conference on 3D Vision (3DV). IEEE, 2019: 240-248.
- [33] Zhao H, Qi X, Shen X, et al. Icnets for real-time semantic segmentation on high-resolution images[C]//Proceedings of the European conference on computer vision (ECCV). 2018: 405-420.
- [34] Yu C, Wang J, Peng C, et al. Bisenet: Bilateral segmentation network for real-time semantic segmentation[C]//Proceedings of the European conference on computer vision (ECCV). 2018: 325-341.
- [35] Johnson D R, Diehn F E, Giannini C, et al. Genetically defined oligodendroglioma is characterized by indistinct tumor borders at MRI[J]. American Journal of Neuroradiology, 2017, 38(4): 678-684.
- [36] Park Y W, Han K, Ahn S S, et al. Whole-tumor histogram and texture analyses of DTI for evaluation of IDH1-mutation and 1p/19q-codeletion status in World Health Organization grade II gliomas[J]. American Journal of Neuroradiology, 2018, 39(4): 693-698.
- [37] Chang K, Bai H X, Zhou H, et al. Residual convolutional neural network for the determination of IDH status in low-and high-grade gliomas from MR imaging[J]. Clinical Cancer Research, 2018, 24(5): 1073-1081.
- [38] Rathore S, Akbari H, Rozycki M, et al. Radiomic MRI signature reveals three distinct

- subtypes of glioblastoma with different clinical and molecular characteristics, offering prognostic value beyond IDH1[J]. *Scientific reports*, 2018, 8(1): 1-12.
- [39] van der Voort S R, Gahrman R, van den Bent M J, et al. Radiogenomic classification of the 1p/19q status in presumed low-grade gliomas[C]//2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017). IEEE, 2017: 638-641.
- [40] Korfiatis P, Kline T L, Erickson B J. Evaluation of a deep learning architecture for MR imaging prediction of ATRX in glioma patients[C]//Medical Imaging 2018: Computer-Aided Diagnosis. International Society for Optics and Photonics, 2018, 10575: 105752G.
- [41] Chang P, Grinband J, Weinberg B D, et al. Deep-learning convolutional neural networks accurately classify genetic mutations in gliomas[J]. *American Journal of Neuroradiology*, 2018, 39(7): 1201-1207.
- [42] Eichinger P, Alberts E, Delbridge C, et al. Diffusion tensor image features predict IDH genotype in newly diagnosed WHO grade II/III gliomas[J]. *Scientific reports*, 2017, 7(1): 1-9.
- [43] Bisdas S, Shen H, Thust S, et al. Texture analysis-and support vector machine-assisted diffusional kurtosis imaging may allow in vivo gliomas grading and IDH-mutation status prediction: a preliminary study[J]. *Scientific reports*, 2018, 8(1): 1-9.
- [44] Ren Y, Zhang X, Rui W, et al. Noninvasive prediction of IDH1 mutation and ATRX expression loss in low-grade gliomas using multiparametric MR radiomic features[J]. *Journal of Magnetic Resonance Imaging*, 2019, 49(3): 808-817.
- [45] Pan C, Liu J, Tang J, et al. A machine learning-based prediction model of H3K27M mutations in brainstem gliomas using conventional MRI and clinical features[J]. *Radiotherapy and Oncology*, 2019, 130: 172-179.
- [46] Li Z C, Bai H, Sun Q, et al. Multiregional radiomics features from multiparametric MRI for prediction of MGMT methylation status in glioblastoma multiforme: a multicentre study[J]. *European radiology*, 2018, 28(9): 3640-3650.
- [47] Arita H, Kinoshita M, Kawaguchi A, et al. Lesion location implemented magnetic resonance imaging radiomics for predicting IDH and TERT promoter mutations in grade II/III gliomas[J]. *Scientific reports*, 2018, 8(1): 1-10.
- [48] AMIA. Informatics summit - session details. 2018. Available at: https://informaticssummit2018.zerista.com/event/member?item_id¼7574587. [Accessed 17 September 2018].
- [49] Fetit A E, Novak J, Peet A C, et al. Three-dimensional textural features of conventional MRI improve diagnostic classification of childhood brain tumours[J]. *NMR in*

- Biomedicine, 2015, 28(9): 1174-1184.
- [50] Louis D N, Perry A, Reifenberger G, et al. The 2016 World Health Organization classification of tumors of the central nervous system: a summary[J]. *Acta neuropathologica*, 2016, 131(6): 803-820.
- [51] Eckel-Passow J E, Lachance D H, Molinaro A M, et al. Glioma groups based on 1p/19q, IDH, and TERT promoter mutations in tumors[J]. *New England Journal of Medicine*, 2015, 372(26): 2499-2508.
- [52] Pallud J, Mandonnet E, Duffau H, et al. Prognostic value of initial magnetic resonance imaging growth rates for World Health Organization grade II gliomas[J]. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, 2006, 60(3): 380-383.
- [53] Chamberlain M C. Is the volume of low-grade glioma measurable and is it clinically relevant?[J]. 2014.
- [54] Gevaert O, Mitchell L A, Achrol A S, et al. Glioblastoma multiforme: exploratory radiogenomic analysis by using quantitative image features[J]. *Radiology*, 2014, 273(1): 168-174.
- [55] Gutman D A, Cooper L A D, Hwang S N, et al. MR imaging predictors of molecular profile and survival: multi-institutional study of the TCGA glioblastoma data set[J]. *Radiology*, 2013, 267(2): 560-569.
- [56] Zhou H, Vallières M, Bai H X, et al. MRI features predict survival and molecular markers in diffuse lower-grade gliomas[J]. *Neuro-oncology*, 2017, 19(6): 862-870.
- [57] Zhang X, Tian Q, Wang L, et al. Radiomics strategy for molecular subtype stratification of lower-grade glioma: detecting IDH and TP53 mutations based on multimodal MRI[J]. *Journal of Magnetic Resonance Imaging*, 2018, 48(4): 916-926.
- [58] Jakola A S, Zhang Y H, Skjulsvik A J, et al. Quantitative texture analysis in the prediction of IDH status in low-grade gliomas[J]. *Clinical neurology and neurosurgery*, 2018, 164: 114-120.
- [59] Suh H B, Choi Y S, Bae S, et al. Primary central nervous system lymphoma and atypical glioblastoma: differentiation using radiomics approach[J]. *European Radiology*, 2018, 28(9): 3832-3839..
- [60] Zacharaki E I, Wang S, Chawla S, et al. Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme[J]. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 2009, 62(6): 1609-1618..

- [61] Citak-Er F, Firat Z, Kovanlikaya I, et al. Machine-learning in grading of gliomas based on multi-parametric magnetic resonance imaging at 3T[J]. *Computers in biology and medicine*, 2018, 99: 154-160.
- [62] Xing F, Cornish T C, Bennett T, et al. Pixel-to-pixel learning with weak supervision for single-stage nucleus recognition in Ki67 images[J]. *IEEE Transactions on Biomedical Engineering*, 2019, 66(11): 3088-3097.
- [63] Elsken T, Metzen J H, Hutter F. Neural architecture search: A survey[J]. *J. Mach. Learn. Res.*, 2019, 20(55): 1-21.
- [64] Baker B, Gupta O, Naik N, et al. Designing neural network architectures using reinforcement learning[J]. *arXiv preprint arXiv:1611.02167*, 2016.
- [65] Zoph B, Le Q V. Neural architecture search with reinforcement learning[J]. *arXiv preprint arXiv:1611.01578*, 2016.
- [66] Liu H, Simonyan K, Yang Y. Darts: Differentiable architecture search[J]. *arXiv preprint arXiv:1806.09055*, 2018.
- [67] He X, Zhao K, Chu X. AutoML: A Survey of the State-of-the-Art[J]. *Knowledge-Based Systems*, 2021, 212: 106622.
- [68] Buciluă C, Caruana R, Niculescu-Mizil A. Model compression[C]//*Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*. 2006: 535-541.
- [69] Hinton G, Vinyals O, Dean J. Distilling the knowledge in a neural network[J]. *arXiv preprint arXiv:1503.02531*, 2015.
- [70] Tao C, Wu W, Xu C, et al. Multi-representation fusion network for multi-turn response selection in retrieval-based chatbots[C]//*Proceedings of the twelfth ACM international conference on web search and data mining*. 2019: 267-275.
- [71] Vielzeuf V, Lechervy A, Pateux S, et al. Centralnet: a multilayer approach for multimodal fusion[C]//*Proceedings of the European Conference on Computer Vision (ECCV) Workshops*. 2018: 0-0.
- [72] 杨逍. 机器学习中隐式因子模型及其优化算法研究[D]. 哈尔滨工业大学, 2013.
- [73] 周志华. 机器学习[M]. 清华大学出版社, 2016年12月.
- [74] Dean J, Corrado G S, Monga R, et al. Large scale distributed deep networks[J]. 2012.
- [75] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[J]. *Advances in neural information processing systems*, 2012, 25: 1097-1105.
- [76] He K, Zhang X, Ren S, et al. Deep residual learning for image

- recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.
- [77] He K, Zhang X, Ren S, et al. Identity mappings in deep residual networks[C]//European conference on computer vision. Springer, Cham, 2016: 630-645.
- [78] 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.
- [79] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015: 234-241.
- [80] Liu C, Chen L C, Schroff F, et al. Auto-deeplab: Hierarchical neural architecture search for semantic image segmentation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 82-92.
- [81] Weng Y, Zhou T, Li Y, et al. NAS-Unet: Neural architecture search for medical image segmentation[J]. IEEE Access, 2019, 7: 44247-44257.
- [82] Tan M, Chen B, Pang R, et al. Mnasnet: Platform-aware neural architecture search for mobile[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 2820-2828.
- [83] Chen W, Gong X, Liu X, et al. Fasterseg: Searching for faster real-time semantic segmentation[J]. arXiv preprint arXiv:1912.10917, 2019.
- [84] Sandler M, Howard A, Zhu M, et al. Mobilenetv2: Inverted residuals and linear bottlenecks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 4510-4520.
- [85] Ma N, Zhang X, Zheng H T, et al. Shufflenet v2: Practical guidelines for efficient cnn architecture design[C]//Proceedings of the European conference on computer vision (ECCV). 2018: 116-131.
- [86] Stamoulis D, Ding R, Wang D, et al. Single-path nas: Designing hardware-efficient convnets in less than 4 hours[C]//Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Springer, Cham, 2019: 481-497.
- [87] Shaw A, Hunter D, Landola F, et al. SqueezeNAS: Fast neural architecture search for faster semantic segmentation[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops. 2019: 0-0.
- [88] Zhang H, Wu C, Zhang Z, et al. Resnest: Split-attention networks[J]. arXiv preprint arXiv:2004.08955, 2020.

- [89]Louis D N, Perry A, Reifenberger G, et al. The 2016 World Health Organization classification of tumors of the central nervous system: a summary[J]. *Acta neuropathologica*, 2016, 131(6): 803-820.
- [90]Molenaar R J, Botman D, Smits M A, et al. Radioprotection of IDH1-mutated cancer cells by the IDH1-mutant inhibitor AGI-5198[J]. *Cancer research*, 2015, 75(22): 4790-4802.
- [91]Sulkowski P L, Corso C D, Robinson N D, et al. 2-Hydroxyglutarate produced by neomorphic IDH mutations suppresses homologous recombination and induces PARP inhibitor sensitivity[J]. *Science translational medicine*, 2017, 9(375).
- [92]Beiko J, Suki D, Hess K R, et al. IDH1 mutant malignant astrocytomas are more amenable to surgical resection and have a survival benefit associated with maximal surgical resection[J]. *Neuro-oncology*, 2014, 16(1): 81-91.
- [93]Zhang M, Pan Y , Qi X, et al. Identification of new biomarkers associated with idh mutation and prognosis in astrocytic tumors using nanostring nCounter analysis system. *Appl Immunohistochem Mol Morphol*. 2018;26(2):101–107.
- [94]Zhang M, Pan Y, Qi X, et al. Identification of new biomarkers associated with IDH mutation and prognosis in astrocytic tumors using NanoString nCounter Analysis System[J]. *Applied Immunohistochemistry & Molecular Morphology*, 2018, 26(2): 101-107.
- [95]Matsui Y, Maruyama T, Nitta M, et al. Prediction of lower-grade glioma molecular subtypes using deep learning[J]. *Journal of neuro-oncology*, 2020, 146(2): 321-327.
- [96]Wegmayr V, Aitharaju S, Buhmann J. Classification of brain MRI with big data and deep 3D convolutional neural networks[C]//*Medical Imaging 2018: Computer-Aided Diagnosis*. International Society for Optics and Photonics, 2018, 10575: 105751S.
- [97]Feng X, Yang J, Lipton Z C, et al. Deep learning on MRI affirms the prominence of the hippocampal formation in Alzheimer’s disease classification[J]. *bioRxiv*, 2018: 456277.
- [98]Du J, Li W, Lu K, et al. An overview of multi-modal medical image fusion[J]. *Neurocomputing*, 2016, 215: 3-20.
- [99]Selvaraju R R, Cogswell M, Das A, et al. Grad-cam: Visual explanations from deep networks via gradient-based localization[C]//*Proceedings of the IEEE international conference on computer vision*. 2017: 618-626.

攻读学位期间主要的研究成果

一、发表的学术论文

[1] 本人第一作者. Brain Tumors Segmentation Based on 3D Dilated-U-Net[C].中南大学自动化学院第二届研究生学术年会论文集

[2] 本人第二作者（核心贡献）. Prediction of IDH Mutation Status of Glioma based on Multimodal MRI Images[C]. ACM 2021 3rd International Conference on Intelligent Medicine and Image Processing (IMIP 2021)(Accept)

[3] 本人第二作者（核心贡献）. A Novel Multi-modal Heterogeneous Neural Network for KI67 Index Prediction [C]. 2021 IEEE International Conference on Image Processing (ICIP 2021)(Accept)

二、申请国家发明专利

[1]一种多模态脑肿瘤特征分析方法: 201811597493.6[P]（申请）

[2]一种 ki67 指数确定方法及系统: 202110353954.0 [P]（申请）

三、硕士期间参与项目

[1] 教育部-中国移动医疗联合实验室开放课题：“多模态脑部图像分析”，2018.3-2019.7,主要技术人员，结题

[2] 教育部-中国移动医疗联合实验室开放课题：“基于非侵入式信息的胶质瘤分子分型预测系统及临床应用评估”，2019.7-至今,主要研究人员

四、硕士期间获得奖励情况

[1] 2018 年第四届中国“互联网+”大学生创新创业大赛全国总决赛银奖. 团队技术负责人.

[2] 2018 年中国计算机大赛-人工智能创意赛全国一等奖. 团队技术负责人.

[3] 2019 年中南大学格林美创新实践团队奖. 团队技术负责人.

致谢

时光转瞬即逝，在中南大学的校园已经生活了三年。在离开长沙这个城市前，我感到有太多的收获和成长。在这里，我有机会见识美丽的风景，结交优秀的朋友，学习前沿的知识，了解广阔的社会，经历难忘的奋斗历程。难忘与实验室的同学们度过的一个个默契合作、共同努力的日子，一次次交流借鉴、受益匪浅的例会，一场场欢声笑语、加深情感的团建。我真诚地感谢陪伴我、指引我、帮助我一路成长的师长同学们，感谢中南大学给我的平台和资源。首先要感谢我的研究生导师及人生导师。非常庆幸能有缘成为老师的学生。导师以渊博深厚的专业知识，求真务实的工作态度，高瞻远瞩的发展理念，诲人不倦的高尚师德，平易近人的处事风格深刻地影响了我。感谢老师在学术上给予我的专业的指导，指引着我探求学术的道路上找到方向、逐步前行。感谢老师指导我们完成项目的申请，在细节处给到我们准确的指点。感谢老师组织的温馨欢乐的团建活动，让我们在实验室中也感受到家一般的温暖氛围。衷心地祝愿谢老师及我们的团队发展得越来越好，每个人都迎来大好前程。感谢学院的老师们在此期间给我的指导和帮助，衷心祝愿各位老师工作顺利，身体健康！

感谢师兄师姐们的帮扶照顾。感谢我的同门，作为三年来一起学习、一起做项目的战友，祝你们未来都前程似锦！感谢同届的其他同学们，非常珍惜和你们一起努力一起欢笑的经历。感谢师弟师妹的相伴相助。

真诚地感谢父母二十多年来的养育之恩。是父母给了我宝贵的生命和美好的人生，用自己的青春换来了我的青春。如今终于要走入社会，用所学的知识回报父母，回报社会，我一定不会忘记父母的嘱托，努力为自己的事业与祖国美好的未来而奋斗打拼。

2021年4月14日

于家中