

硕士学位论文

基于放射基因组学的 Ki-67 指数  
非侵入式预测

Non-invasive prediction of Ki-67 index based on  
radiogenomics

专业学位类别

电子信息

专业领域

人工智能

作者姓名

李旭辉

指导教师

谢斌 副教授

2023 年 5 月

中图分类号 TP391.5

学校代码 10533

UDC 004.8

学位类别 专业学位

## 硕士学位论文

# 基于放射基因组学的 Ki-67 指数 非侵入式预测

Non-invasive prediction of Ki-67 index based on  
radiogenomics

作者姓名	李旭辉
专业学位类别	电子信息
专业领域	人工智能
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 副教授
副指导教师	

论文答辩日期 2023.5.24

答辩委员会主席 王勇

中南大学

2023年5月

## 学位论文原创性声明

本人郑重声明，所呈交的学位论文是本人在导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。

学位论文作者签名：李旭琦 日期：2023年5月30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，允许本学位论文被查阅和借阅。本人授权中南大学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和公开传播，可以采用复印、缩印或其它手段保存和汇编学位论文。本人同意按《中国优秀博硕士学位论文全文数据库出版章程》规定享受相关权益。本人保证：毕业后以学位论文内容发表的论文作者单位注明中南大学；学位论文电子文档的内容和纸质学位论文的内容相一致。

延缓公开论文延缓到期后适用本授权书，涉密论文在解密后适用本授权书。

本学位论文属于：(请在以下相应方框内打“√”)

公开

延缓公开，延缓期限（\_\_年\_\_月\_\_日至\_\_年\_\_月\_\_日）

学位论文作者签名：李旭琦

日期：2023年5月30日

指导教师签名：谭秋明

日期：2023年5月30日

(填写阿拉伯数字)

# 基于放射基因组学的 Ki-67 指数 非侵入式预测

## 摘要:

脑神经胶质瘤是颅内最常见的原发性高度恶性肿瘤，高级别胶质瘤的生存期极短，早期的及时诊断能有效降低其致死率。研究表明 Ki-67 指数与胶质瘤的临床预后及个性化治疗手段有紧密关联，但其诊断依赖于繁琐且耗时的组织活检，这会给本不富裕的医疗资源带来严峻挑战。若是能在术前得到 Ki-67 指数水平将带来重大的医学价值及临床意义。现有公开研究大多致力于统计特征的特征工程方法来预测 Ki-67 指数，这一过程需要大量的人力参与及时间成本，很难满足实际工程的需求。因此本文基于患者的核磁共振影像及临床元信息来预测其胶质瘤 Ki-67 指数水平，这种端对端的非侵入式预测方法，有望成为实际临床诊断中极为重要的辅助决策。本文具体的研究内容在以下方面：

(1) 本文依托湘雅医院，从真实就诊记录中整理得到湘雅胶质瘤分子分型 XiangyaGMS 数据集。目前已有的公开胶质瘤数据集暴露出规模小，分子病理标签不全等问题，凭借这些并不足以保障相关算法研究的工作开展。而 XiangyaGMS 数据集包含 619 例患者，清洗后其 MRI 影像规模接近 4 万，同时还包含有完整的分子病理标签及元信息，这些都有利于胶质瘤领域的相关研究开展。本文基于常见的卷积神经网络，在医学基准任务上评估了该数据集的可行性及适用性，验证结果表明 GMS 数据集有望成为同类研究实验数据的良好补充。

(2) 为实现胶质瘤病灶的精准分割，本文提出多任务信息引导的胶质瘤病灶分割网络 MIG-Net，其预测结果能接近专业医生的标注水平。肿瘤病灶无疑包含最丰富的肿瘤分子分型特征，因此胶质瘤病灶分割是后续分子标记物预测的重要预处理手段之一。本文希望在病灶分割模型的建模中，更多的赋予模型来自先验知识的可解释性。本文首先提出病灶概率图谱 FPG 损失，使模型更多聚焦到前景目标上；此外，通过引导胶质瘤恶性程度分级任务中同质的定位信息与语义信息，多任务信息引导网络 MIG-Net 能在交叉验证中取得 90.56% 的 Dice 相似系数指标。MIG-Net 在跨域数据集 BraTS 上的交叉验证实验，说明其预测结果具备较强的鲁棒性。

(3) **基于 XiangyaGMS 数据集，本文探索出一个多尺度多序列模型 KiNet，能首次实现对 Ki-67 指数的非侵入式预测。**核磁共振影像成像过程中会生成反馈不同肿瘤信息的序列数据，肿瘤分子分型预测研究的重要一步就是，设计多序列融合模型来有效捕捉不同序列影像之间的同质特征。得益于中期融合的有效性，本文设计出多尺度多序列模型 KiNet 来提取放射基因组学。此外，还提出一个新的二元多尺度 BMS 损失来监督同一患者序列间的相似性。放射基因组学特征结合患者性别年龄后，基于 LightGBM 在 Ki-67 指数水平预测任务上交叉验证取得 0.788 的 AUC 指标，最终将单张切片预测结果融合，在患者级预测上得到 **0.857 的 AUC 指标**。

(4) **本文展望胶质瘤相关算法的落地，设计并实现了胶质瘤 Ki-67 指数预测辅助诊断系统。**肿瘤分子分型辅助诊断系统在医学临床决策中有重要的应用前景，而目前符合性能的相关研究尚不够智能便捷。本文将前期的胶质瘤病灶分割和 Ki-67 指数预测算法级联成一个闭环框架，并把级联框架部署在客户端联网模式下，可实现 Ki-67 指数水平预测的功能及临床“阅片”的交互体验。

图 31 幅，表 13 个，参考文献 85 篇

**关键词：**神经胶质瘤；放射基因组学；Ki-67 指数；病灶分割；核磁共振影像序列

**分类号：** TP391.4

## **Non-invasive prediction of Ki-67 index based on radiogenomics**

**Abstract:** Glioma is the most common primary and highly malignant tumor in the brain. The survival time of high-grade glioma is very short, and early and timely diagnosis can effectively reduce its mortality. Studies have shown that Ki-67 index is closely related to the clinical prognosis and personalized treatment of glioma, but its diagnosis relies on tedious and time-consuming tissue biopsy, which poses a severe challenge to the poor medical resources. If Ki-67 index level can be obtained before operation, it will bring great medical value and clinical significance. Most of the existing public studies focus on the characteristic engineering method of statistical characteristics to predict Ki-67 index, which requires a lot of manpower participation and time cost, and is difficult to meet the needs of actual engineering. Therefore, in this thesis, Ki-67 index level of glioma was predicted based on MRI images and clinical meta-information of patients. This end-to-end non-invasive prediction method is expected to become an extremely important auxiliary decision making in actual clinical diagnosis. The specific research content of this thesis is as follows:

(1) Based on Xiangya Hospital, Xiangya glioma molecular typing XiangyaGMS dataset was collected from real medical records. Currently available publicly available glioma data sets reveal problems such as small scale and incomplete molecular pathological labels, which are not enough to guarantee the work of relevant algorithm research. The XiangyaGMS dataset contains 619 patients with MRI image size of nearly 40,000 after cleaning. Besides, it also contains complete molecular pathological tags and meta-information, which are conducive to the development of related research in the field of glioma. Based on common convolutional neural networks, this thesis evaluates the feasibility and applicability of this data set on medical benchmark tasks. The verification results show that the GMS data set is expected to be a good supplement to the experimental data of similar studies.

(2) In order to achieve accurate segmentation of glioma lesions, this thesis proposes a multi-task information-guided glioma lesion segmenta-

tion network MIG-Net, whose prediction results can be close to the labeling level of professional doctors. Tumor foci undoubtedly contain the most abundant tumor molecular typing characteristics, so glioma foci segmentation is one of the important pretreatment methods for subsequent molecular marker prediction. In this thesis, we hope to give the model more interpretability from prior knowledge in the modeling of focal segmentation model. In this thesis, we first proposed the FPG loss of the foci probability map to make the model focus more on the foreground target. In addition, by guiding the homogenous localization information and semantic information in glioma malignancy grading task, the multi-task information-guided network MIG-Net was able to obtain Dice similarity coefficient index of 90.56% in cross validation. The cross-validation experiments of MIG-Net on cross-domain data set BraTS show that its prediction results have strong robustness.

(3) Based on XiangyaGMS dataset, a multi-scale and multi-sequence model KiNet was developed in this thesis, which can realize the non-invasive prediction of Ki-67 index for the first time. In the process of MRI imaging, sequence data with different tumor information feedback will be generated. An important step of tumor molecular typing prediction research is to design a multi-sequence fusion model to effectively capture the homogenous features between different sequences of images. Thanks to the effectiveness of mid-stage fusion, a multi-scale and multi-sequence model KiNet was designed to extract radiogenomics. In addition, a new binary multiscale BMS loss is proposed to monitor similarity between sequences of the same patient. After radiogenomic characteristics combined with patient gender and age, an AUC index of 0.788 was obtained based on LightGBM cross-validation on the Ki-67 index level prediction task. Finally, the prediction results of single slices were fused, and an AUC index of 0.857 was obtained for patient-level prediction.

(4) This thesis looks forward to the implementation of glioma correlation algorithm, and designs and implements glioma Ki-67 index predictive auxiliary diagnosis system. Tumor molecular classification aid diagnosis system has an important application prospect in medical clinical decision making, but the current research on its consistent performance is not intelli-

gent and convenient. In this thesis, the previous segmentation of glioma lesions and Ki-67 index prediction algorithm are cascaded into a closed-loop framework, and the cascaded framework is deployed in the client networking mode, which can realize the function of Ki-67 index level prediction and the interactive experience of clinical "film reading".

**Keywords:** Glioma;Radiogenomics;Ki-67 index;Lesion segmentation;Magnetic resonance imaging sequence

**Classification:** TP391.5

# 目 录

第 1 章 绪论	1
1.1 研究背景与意义	1
1.2 国内外研究现状	3
1.2.1 肿瘤影像分割方法研究进展	4
1.2.2 分子分型预测研究现状	6
1.3 研究目的与主要研究内容	9
1.3.1 研究目的	9
1.3.2 研究内容及本文结构安排	9
第 2 章 湘雅胶质瘤分子分型数据集——XiangyaGMS	11
2.1 XiangyaGMS 数据集的建立	11
2.1.1 数据的获取与清洗	11
2.1.2 数据的标注与后处理	14
2.1.3 数据集的划分与架构	16
2.2 数据集的评估	17
2.2.1 数据集的基准实验	17
2.2.2 评价指标	18
2.2.3 基准实验性能结果	19
2.3 本章小结	19
第 3 章 基于多任务信息引导的胶质瘤病灶分割	21
3.1 病灶先验概率图谱	22
3.2 多任务引导的胶质瘤病灶分割模型	24
3.2.1 多任务学习	24
3.2.2 胶质瘤病灶分割模型	26
3.2.3 双重先验知识引导	27
3.2.4 逆向注意力门	29
3.3 实验结果与讨论	31
3.3.1 模型评估准则	32
3.3.2 与主流模型的实验结果对比	33

3.3.3	消融实验	34
3.3.4	分析与讨论	34
3.4	本章小结	36
<b>第 4 章</b>	<b>基于多序列融合的胶质瘤 Ki-67 指数预测</b>	<b>37</b>
4.1	多模态融合方式探索	37
4.1.1	多模态模型	38
4.1.2	实验细节及评价指标	39
4.1.3	实验结果及分析	40
4.2	基于多序列融合的 Ki-67 指数预测模型	42
4.2.1	多尺度多序列融合模型	42
4.2.2	二元多尺度损失函数	43
4.2.3	实验细节及评价指标	45
4.2.4	实验结果与分析	46
4.3	患者实例诊断	49
4.4	本章小结	50
<b>第 5 章</b>	<b>级联框架和诊断系统设计</b>	<b>52</b>
5.1	级联框架设计与实现	52
5.2	诊断客户端设计与实现	54
5.3	本章小结	56
<b>第 6 章</b>	<b>总结与展望</b>	<b>57</b>
6.1	论文总结	57
6.2	工作展望	58
	参考文献	59
	攻读学位期间主要的研究成果	66
	致  谢	68

## 插图索引

图 1-1	某患者的 MRI 切层图像	2
图 1-2	不同 Ki-67 指数的患者 MRI 多序列影像	3
图 1-3	全文框架图	10
图 2-1	XiangyaGMS 数据集建立流程	11
图 2-2	XiangyaGMS 数据集中某患者的示例。(b) 为 T1 序列某切层图像；(c) 为 T2 序列某切层图像；(d), (e), (f) 分别为 T1c 序列的冠状, 矢状, 轴向面切层图像	12
图 2-3	MRI 影像的颅骨剥离前后示意图	13
图 2-4	MRI 影像归一化前后的效果对比图	14
图 2-5	XiangyaGMS 数据集中胶质瘤病灶标注示意	15
图 2-6	XiangyaGMS 数据集的文件架构	17
图 2-7	Xception 的混淆矩阵	19
图 3-1	$W_{FPG}$ 在 0-1 上的分布	23
图 3-2	常见的多任务模型学习模式	25
图 3-3	MIG-Net 的概况	27
图 3-4	双重先验知识引导模块的概况	28
图 3-5	逆注意力门模块的概况	30
图 3-6	激活函数示意图	31
图 3-7	DSC 指标的几何意义	33
图 3-8	部分患者 MRI 影像的预测与真值的可视化。每行分别对应 T1c、T1 和 T2 序列	35
图 4-1	不同融合方式的多模态模型	38
图 4-2	ResNet 上针对不同输入组合的可视化结果。注：从蓝色到红色表示重要性的增加	41
图 4-3	KiNet 块的架构	43
图 4-4	KiNet 的总体框架（在测试阶段不考虑虚线标记）	44
图 4-5	相似度监测提供单序列输入的模型	45
图 4-6	Ki-67 指标不同模型的 ROC 曲线	47
图 4-7	训练完成的融合参数可视化（其中由内到外分别对应主分支的第 1-5 个阶段）	48
图 4-8	KiNet 基于不同患者级融合方法的概率校准曲线	50
图 5-1	诊断系统级联模型的概况	52

图 5-2	诊断系统细节流程 . . . . .	53
图 5-3	辅助诊断系统的客户端界面 . . . . .	54
图 5-4	数据输入的结构化目录说明 . . . . .	55
图 5-5	辅助诊断系统的流程图 . . . . .	56

## 表格索引

表 1-1	临床参考分子标记物的检测方式和诊断与预后价值 . . . . .	2
表 2-1	XiangyaGMS 临床元信息和分子分型诊断结果统计信息 . . . . .	16
表 2-2	混淆矩阵示意 . . . . .	18
表 2-3	三种经典模型在 XiangyaGMS 数据集上的性能 . . . . .	19
表 3-1	与主流分割模型性能的平均 DSC(%) 指标和 HD 指标对比 . . . . .	33
表 3-2	MIG-Net 模型的消融实验结果 . . . . .	34
表 3-3	XiangyaGMS 数据集与 BraTS 数据集交叉测试下的 DSC 指标 . . . . .	35
表 4-1	输入组合的不同构造 . . . . .	38
表 4-2	不同输入组合在 ResNet-50 网络上的预测性能 . . . . .	40
表 4-3	不同融合方式的 ResNet-50 预测性能对比 . . . . .	41
表 4-4	不同神经网络对 Ki-67 指标的预测性能 . . . . .	46
表 4-5	KiNet 消融实验的预测性能 . . . . .	47
表 4-6	KiNet 在 XiangyaGMS 测试集上的患者级预测性能 . . . . .	50

# 第 1 章 绪论

## 1.1 研究背景与意义

脑神经胶质瘤是人脑内部最常见的原发性恶性肿瘤，约占所有神经肿瘤的 32%[1, 2]，它由异常生长的神经胶质细胞构成，通常伴随有浸润性、弥散性、边界模糊等病理可观特征 [3]。不确定性的肿瘤位置和增殖速度，导致胶质瘤可能损害脑功能并危及生命。世界卫生组织 (World Health Organization, WHO) 依据组织病理学检查特征划分脑胶质瘤恶性程度为 I~IV 级 [4]，也是目前临床病理诊断的标准。其中 I/II 级属于低级别胶质瘤 (Low Grade Glioma, LGG)，III/IV 级则属于高级别胶质瘤 (High Grade Glioma, HGG)。根据中国胶质瘤协作组 (Chinese Glioma Cooperation Group, CGCG) 的统计资料表明，高级别胶质瘤（特别是 IV 级）占全部原发脑肿瘤的 50% 以上。即使给予最积极的治疗，由于高复发性，患者的中位生存期仍然小于 15 个月 [5]，而低级别胶质瘤的生存期长达 4.7-9.8 年。目前，尽管存在手术切除、放疗和化疗等治疗脑胶质瘤的方法，但是脑胶质瘤患者的生存率仍然非常低 [6]。

最近，神经肿瘤学和分子病理学研究已经观察到存在一些分子标记有助于胶质瘤的临床诊断。尽管如此，目前的 WHO 肿瘤分级大体上仍然是以组织病理学检查特征为基础。但是组织病理学提取的特征具有很强的主观性，很多时候并不能确保肿瘤组织的基因学背景和生物学特征足够客观、系统及准确。已经有足够的真实案例表明组织病理学在面对相似组织结构个体的局限，而存在相似组织结构个体间的分子特性是不同的，能反映出不同个体间预后的较大差异 [7]。随着人类基因组计划的完成和高通量技术的发展，基于肿瘤形态学和基因特征的分子病理和分子分型的概念已逐渐被神经外科学界认可并应用于胶质瘤的个体化治疗。在 2016 年，世界卫生组织 (WHO) 已将分子病理纳入胶质瘤病理诊断体系 [8]，证实已列入临床诊疗标准的分子标记可作为未来治疗的靶点，并成为预测预后的重要参考。

研究表明基于放射基因组学的分子病理（如 Ki-67 指数等）与胶质瘤的预后关系更加密切，并且能够为难以诊断的复杂肿瘤状况提供鉴别依据 [9]。放射基因组学是一种从分子水平研究放射学影像中基因组结构和功能特征的技术，其目标是预测组织的基因组特性的技术，医学成像结果与基因组特性之间的关联已经逐渐成为相关研究的趋势 [10]。

现有列入标准及有望成为新靶点的部分分子标记物已经在表 1-1 中罗列，并整理出其诊断与预后价值以及对应的临床检测手段。然而由表中得到，鉴别分子

水平基因组特性的金标准是对活检或手术切除获得的组织样本进行免疫组化或基因测序。参考癌症基因组图谱 (The Cancer Genome Atlas, TCGA) 报告上的描述, 患者提供的活检样本中仅约 35% 能够包含有达标的肿瘤细胞, 这些才被允许用于准确的分子鉴定 [11]。严格的化验周期以及医疗设备的繁琐操作无疑会给病理科医生带来较大的工作负荷, 同时也会对患者的自身健康造成不可逆的危害。

表 1-1 临床参考分子标记物的检测方式和诊断与预后价值

分子标志物	诊断与预后价值	临床检测手段
IDH 突变	在判定预后方面还没有定论, 但对于低级别胶质瘤的诊断贡献是可以肯定的	DNA 焦磷酸测序或免疫组化
MGMT 启动子甲基化	为胶质瘤判预后及治疗策略提供重要信息	焦磷酸测序或甲基化定性
染色体 1p/19q 缺失状态	针对少突胶质细胞瘤的诊断标准设定有一定作用, 是其对细胞毒疗法的预后标记	荧光原位杂交或阵列比较基因组杂交
EGFR 扩增	对胶质母细胞瘤的诊断、预后和治疗有指导意义	荧光原位杂交
PIEN 基因突变	对高级别胶质瘤来说是一个判定预后的因子	Sanger 测序
TP53 基因突变	已成为诊断的生物标记, 有可能成为药物靶点	Sanger 测序
Ki-67 指数	WHO 分级的重要参考之一, 可以作为评判预后的指标	免疫组化
miR-181d	对胶质母细胞瘤是一个预测预后的指标	原位杂交

近几年, 利用非侵入式的信息即在术前得出分子水平基因组特性的技术逐渐兴起并应用于脑肿瘤的诊断和治疗过程中。核磁共振图像 (Magnetic Resonance Imaging, MRI) 已成熟应用于颅脑的“无损伤”成像诊断, 临床上常使用 T1、T1c、T2 和 T2Flair 四个序列的 MRI 数据作为脑肿瘤诊断的依据, 某患者 MRI 切层影像如图 1-1 所示。

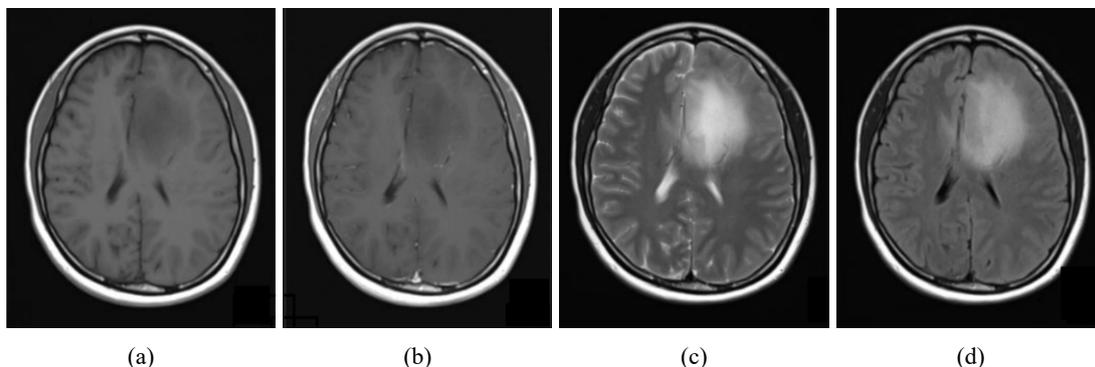


图 1-1 某患者的 MRI 切层图像

MRI 的 T1 序列能直观展现脑组织的解剖结构, 几乎包裹脑结构的脑脊液在其中表现为深色, 因此整体的切层影像是偏暗的, 如图 1-1-a。而此基础上加权后

的 T1c 序列能给脑部结构带来不同质地的对比层次,更突出病损区域,如图1-1-b。平扫 T2 序列能显示出肿瘤的不均匀性及不同程度的水肿现象,如图1-1-c,导致病损区域在影像中如“高亮”般突显出来,对病变的特征能更具体得到。T2Flair 序列中对病损更为敏感,其他区域会受到一定程度模糊,如图1-1-d。

胶质瘤亚型众多,不同亚型治疗方案差异显著。如图1-2是整理的不同 Ki-67 指数的患者 MRI 影像,其中可读知识对于非医务人员来说十分匮乏。专业医师有把握做到凭肉眼从 MRI 影像中获得脑胶质瘤的位置以及大小等,但仅仅这种程度,无法确定肿瘤内分子标记物的差异情况,给治疗方案的选择带来了极大挑战。因此,基于放射基因组学在术前预测出患者肿瘤的分子标记物情况能够在临床中起到重要的辅助作用,对于治疗方案的制定有极大参考意义。

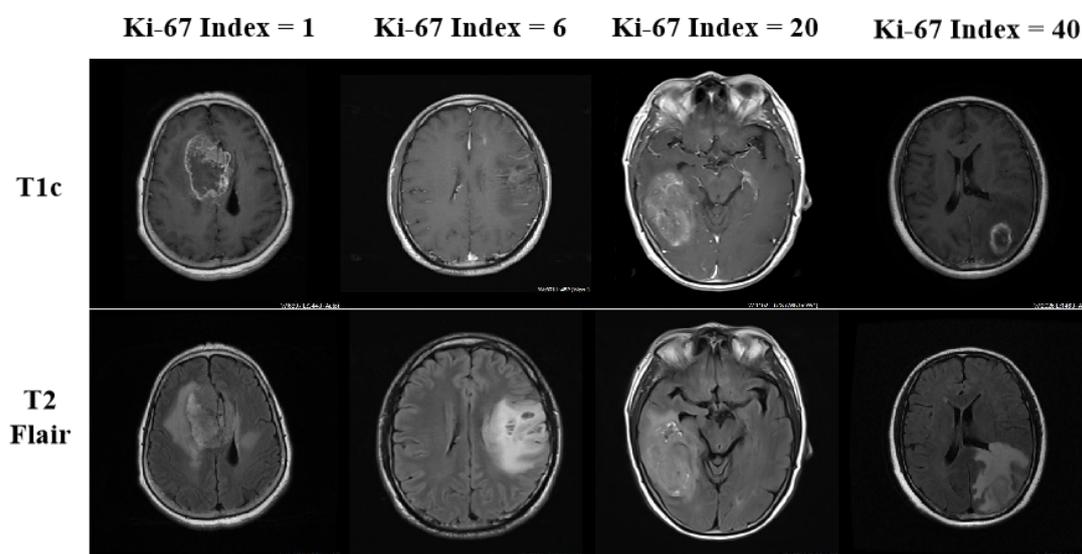


图 1-2 不同 Ki-67 指数的患者 MRI 多序列影像

对于实现分子标记物的术前非侵入预测,本项目与湘雅医院开展相关研究合作,通过深度学习技术对患者脑部的医学 MRI 影像及对应的临床问诊数据作基因组学特征提取,能够自动勾画脑部胶质瘤的病损区域,同时实现对标准分子标记物术前预测的功能,从而为临床治疗提供合理的智能辅助系统。

## 1.2 国内外研究现状

近年来,在精准医疗的推动下,其中的放射基因组学在目前医学文献研究中引起了广泛关注。传统放射基因组学研究着力从医学诊疗原理出发,提供足够多的预定义图像特征来描述医学影像信息。这种反馈信息可以是定性的,也可以使定量的,但其中步骤的繁琐对匮乏的医务资源带来十分严峻的挑战。与此同时,图形处理器等计算资源的快速增长为医学图像分析技术的发展带来重要的必要条件,相关的研究进入高速增长阶段。从临床辅助系统的设计角度出发,希望实

现胶质瘤的病损分割以及分子标记物的预测功能，本节将从两个角度出发介绍国内外研究工作。

### 1.2.1 肿瘤影像分割方法研究进展

从笼统来说，放射基因组学研究的标准步骤首先需要获取到高质量的设备成像，并通过预处理来整理成用于计算机读取的结构化数据集。整个预处理流程从色彩空间可以作标准化，但是与其他的大型公开数据集（如 ImageNet 等）不同，医学图像分析前能接收到具有可行性的先验医学参考。对肿瘤分子标记物的诊断任务来说，如 Ki-67 指数，其临床检测是统计肿瘤内部的抗体染色比例，因此 MRI 影像中的肿瘤区域是更加需要被关注到的。病损勾画在许多研究工作中都是由合作的医务团队完成，这一步骤在临床辅助系统中会被自动化标注，也可作为分子变异预测的准备工作。

目前已有的自动和半自动脑肿瘤分割方法主要分为两类：基于生成模型和基于判别模型 [12]。基于生成模型的方法通常需要大量的临床先验信息来进行脑肿瘤分割，本质上其实都是将数据映射到一个新的空间，并且这个空间的数据分布相对容易获得 [13]。另一方面，基于判别模型的方法能够基于图像特征，将图像中像素点分类为肿瘤或正常组织。但是这类分割方法的性能依赖于图像特征的提取以及分类模型。手动构造的用于肿瘤分割研究的主要包含多种基于图像的宏观特征，包括局部直方图 [14]、图像纹理 [15]、结构张量特征值 [16] 等。脑肿瘤分割早期研究中最常用的机器学习分类模型是支持向量机 (Support Vector Machines, SVM) 和随机森林 (Random Forest, RF)。

深度学习的盛行推动了语义分割方法的发展。现有的医学图像分割方法主要依靠 U 型结构的全卷积神经网络 (Fully Convolutional Neural Network, FCNN)[17–19]。2015 年，在 IEEE 国际生物医学成像研讨会 (International Symposium on Biomedical Imaging, ISBI) 挑战赛上，Ronneberger 团队 [17] 在果蝇细胞分割任务上开创性地提出 U-Net 模型。经典的 U 型网络 U-Net 由一个具有跳过连接结构的对称编码器-解码器组成。通过卷积层和连续下采样层来获取到深度特征，并将提取的特征上采样到输入分辨率来实现像素级的语义预测。其中跳过连接是指把来自编码器的不同尺度的高分辨率特征进行跳跃连接融合，以缓解下采样造成的空间信息丢失。凭借兼具有效性和效率的结构设计，U-Net 在各种二维和三维医学成像应用 [20] 中取得了巨大的成功。

在医学胶质瘤分割领域中，U-Net 同样也是常用模型之一。例如，Kamnitsas 等 [21] 制作了 U-Net 和 DeepMedic[22] 的组合模型，其中 DeepMedic 基于 DeepLab[23] 在 2017 年脑部肿瘤分割比赛 (Brain Tumor Segmentation, BraTS) 中

获得一等奖。BraTS 2018 的获胜者 Myronenko[24] 提出用变分自动编码器 (Variational Autoencoder, VAE) 来结合 U-Net 结构, 最终取得了挑战赛的最先进指标。Isensee 团队 [25] 提出, 经典的 U-Net 足以达成分割任务的基本性能需求。而基于 U 型结构的变种模型, 其结构合理性的探索仍然是一项艰巨的工作。在医学图像分割方面, V-NAS[26] 设计搜索策略来寻找 U-Net 整体性能最优的下采样和上采样模块, 并在肝脏计算机断层扫描 (Computed Tomography, CT) 2D 医学图像数据集上测试了模型的实际性能。按照这一技术路线, 已经开发了 Res-UNet[27]、UNet++[28] 和 UNet3+[29] 等模型, 用于各种医学成像方式的图像分割。

3D-CNN[20, 22, 24] 可以充分利用 MRI 数据通道维度的空间信息, 然而其模型参数量和计算量负荷也面临大幅增加的问题。在实际研究中可以集成多个 2D CNN 来分割 3D 医学图像。特别是, Prasoon 等 [30] 提出用于膝关节软骨分割的三平面 CNN。三平面网络使用 3 个 CNN 来处理从  $xy$ 、 $yz$  和  $zx$  平面提取的补丁, 并使用 Softmax 分类器层将它们融合。Setio 等 [31] 使用多视图卷积网络进行肺结节检测。他们提出的网络架构由多个 2D CNN 流组成, 在候选结节的特定角度提取的补丁特征输出最终被组合来检测肺结节。Dvorak 等 [32] 将多类脑肿瘤分割任务建模为 3 个二元分割子任务, 每个子任务都使用单一 CNN 解决。大多数脑肿瘤分割方法使用图像块感受野集合 (即 MRI 数据中的局部区域) 来训练 CNN。这些方法将每个图像块分为不同的类别, 例如健康组织、坏死组织、水肿、非增强核、增强核, 其中每个图像块的分类结果用于标记其中心像素来实现肿瘤分割。上述 CNN 脑肿瘤分割方法大多假设每个体素的标签是独立的, 为了考虑分割结果的外观和空间一致性, 马尔可夫随机场 (Markov Random Fields, MRFs), 特别是条件随机场 (Conditional Random Fields, CRFs), 已经在图像分割研究与深度学习技术相结合, 或者用作 CNNs 的后处理步骤 [23] 或制定为神经网络 [33, 34]。在后一种情况下, CNNs 和 CRFs 都可以用反向传播算法参与训练, 从而获得更好的分割性能。

尽管 CNN 展现较强的特征学习能力, 但是对于医学要求中的精准分割还不能够满足。卷积运算对远程依赖信息建模存在固有的局限性, 在输入分辨率较大的情况下, 缺失对于全局和远程语义信息交互的捕捉能力。针对卷积运算的固有局限性, 部分研究团队尝试从感受野角度做突破, 有提出空洞卷积 [35] 和图像金字塔 [36] 等方法来解决这个问题。许多研究试图通过基于特征图中所有像素的全局交互建模, 将自我注意机制整合到 CNN 中。Wang 等 [37] 设计了一个非局部算子, 它可以作为卷积层之间的过渡模块。Schlemper 等 [38] 在 U-Net 基础上, 在残差连接中集成了额外的注意门模块。然而, 这些方法在构造模型的长期依赖关系时仍会暴露其不足。Chen 等 [39] 将自注意力模块应用在 U-Net 的编码器阶

段，对实际任务能有更好的适用性。

最近，受 Transformer 在自然语言处理 (Natural Language Processing, NLP) 领域 [40] 的巨大成功的启发，研究团队试图修改 Transformer 的部分结构以适配其他领域任务的需求。他们试图首先将 Transformer 引入目标检测 (Object Detection) 视觉领域 [41]，取得超预期的结果。在模式识别 (Pattern Recognition) 视觉领域，在 Dosovitskiy 团队 [42] 中提出视觉转换器 (Vision Transformer, ViT) 方法来验证自注意力模型在数据上的延展性。ViT 以带有位置嵌入的二维图像补丁为输入，在 ImageNet 分类上取得了最先进的水平，直接将具有全局自注意力的 Transformer 应用于全尺寸图像。Liu 团队 [43] 提出一个分层的 Swin Transformer，以 Swin Transformer 块为骨干，在语义分割数据集上取得了最先进的性能，能证明 Transformer 在视觉领域的应用潜力。而针对图像分割视觉领域，也有一些基于基本 Transformer 结构改造的模型 [44] 同样取得有前景的性能。

通过整理以上的文献内容，肿瘤影像分割模型的研究现状可以总结为以下几点：

(1) 近期的肿瘤分割模型研究主要是以深度学习方法为主。以神经网络为代表的深度学习方法对于高维数据有很强的特征提取能力，依托于大规模的数据集，其性能的鲁棒程度要超过以阈值法为主的传统图像分割方法。

(2) 未能利用医学中存在大量且全面的临床信息。事实上，临床元信息在医生识别肿瘤病灶过程中起着重要参考意义。但是现有用于医学任务的分割模型中却很少关注到这些，很少利用它为模型赋予更多的医学解释。

(3) CNN 的特征捕捉能力还不能满足医学中的精准分割。医学影像不同于其他领域的数据集，不具备充分的语义信息。所以 CNN 的特征学习会受到感受野的约束，逐渐成为医学任务中性能提升的局限之一。

### 1.2.2 分子分型预测研究现状

放射基因组学其实是放射组学和基因组学的“结合”。在医学临床领域中，放射组学概括而言是从医学影像中获取到大量有用特征的学科方法，这些定性或者定量的特征都被认为是放射组学特征，包含某些肿瘤组织的生殖特征或者模式。当然首先需要点明放射组学的理论假设，患者所属的医学影像应该具备一定量的特征，也可以说是临床信息。这些特征有助于预测疾病的治疗和预后，对患者的专属医学诊疗提供足够有价值的帮助。而放射组学的相关研究是旨在挖掘医学影像中可能肉眼无法轻易获得的数据特征，甚至捕获到隐藏的可解释性医学指标。而基因组学是研究生物基因组信息与生物疾病联系的工作方向，已被证明能找到疾病有关的一些区别组织病理学的新的诊断以及治疗方法。基因测试能

得到病人更多的个性化治疗信息,从多个维度进行医学精准诊疗。方法论这块更多是依赖于精密设备的基因检测手段,这会对设备以及操作都有严格的要求。这一点对我国国情来说,在农村实行推广及普及是相当滞后的。

放射基因组学简言之,就是提取易获取到的医学临床影像中的有益信息,来预测肿瘤细胞基因组特性,具体地,是基因组的表现导致的染色体分子水平的变化。这类相关的研究表明,针对于基因组特性的预测工作在指标上不够准确,对图像数据的输入要求颇高,尚不能用于临床应用。而即使是取得出色的指标,仍然会被要求提供具体的医学可解释性说明。尽管如此,放射基因组学的研究工作仍有望发现高分辨率的、可挖掘的特征数据空间。

目前在相关领域已进行了大量研究,旨在探究弥漫性胶质瘤的成像表型与肿瘤分子标志物之间的关联。[45, 46],医学影像特征,比如 MRI,被发现可以预测 LGG 的预后生存和分子分型 [47]。因此,基于非侵入式机器学习方法来确定分子标记物的技术发展非常迅速。经典机器学习方法的效果在很大程度上依赖于影像中的宏观特征,如解剖区域,形状,纹理,强度等,同时被要求花费更多的人力财力。Gevaert 团队 [45] 在 2014 年利用 MRI 图像提取高通量的图像特征,描述了胶质母细胞瘤的影像学表型信息,并建立了相应的影像基因组学图谱,将这些特征与多种分子数据关联。研究发现,其中 3 个特征与存活期显著相关,7 个图像特征与分子分型相关 ( $P < 0.05$ ),这表明影像组学有可能无创地预测肿瘤的临床表现和分子标记物。Yao 等人 [48] 使用 CT 图像提取手工放射组学 (handcrafted radiomics, HCR) 特征,并对该信息进行建模,以无创性地预测表现良好的肺腺癌患者的 Ki-67 指数状况。经典机器学习已经成功地利用成像特征,从多模态 MRI 特征中预测 IDH 和 TP53 突变状态,预测精度达到了中等水平 [49, 50],MRI 预测肿瘤分级及区分原发性中枢神经系统淋巴瘤和胶质母细胞瘤的研究中,已采用过类似的技术 [51]。Zacharaki 等人 [5] 使用了纹理特征和监督分类器对不同类型的脑瘤和胶质瘤进行了分级,而 Citak 等人 [52] 则采用了多模态方法,支持向量机分类器的准确度达到 93%。区别于以上研究的数据集,低级别的胶质瘤中几乎没有成型的物理特征,这成为了制约经典机器学习方法发展的因素之一。

目前在基于医学影像的肿瘤分子标记物的预测领域,深度学习方法的关注度已经超越了传统的机器学习方法 [53]。最近,Chang 等 [54] 基于常规 MRI 图像使用深度学习来端对端地预测胶质瘤的 IDH1/2 突变状态,达到 94% 的准确度。同时,局部模型可能存在数据泄露的问题;即属于同一患者的不同切片可以分为训练集和测试集。Matsui 等 [55] 使用多模态 MRI、CT、正电子发射断层扫描 (Positron Emission Tomography, PET) 和病人年龄性别等多模态异构信息来预测 IDH 突变状态,准确度达 68%。在整理当地医院数据集的前提工作后, Li 等

[56] 探索出一种多模态切片图像的融合方法来预测胶质瘤 IDH 突变状态。Xing 等 [57] 采用染色侵入活检后的病理图像, 构建了一个弱监督的语义分割网络, 以获得颅内结构的分割结果, 并在临床经验指导下统计预测肿瘤的 Ki-67 指数。借鉴多模态融合的思路, Vielzeuf 等 [58] 假设每个影像序列可以由独立的深度卷积网络处理, 各个分支的特征在中心模块中逐像素级融合, 提高了该多模态融合方法的整体性能。Wang 等 [59] 将肿瘤病灶分割与 IDH 突变状态预测结合起来, 设计了一个新的多任务网络, 在 TCGA 公开数据集上取得了很好的诊断效果。这类方法通常采用 2D 图像 (切片) 分类方法, 而如何选择合适的融合策略来处理这类天生带有多个序列 (可以被认为是模态) 数据集的任务是面临的一个挑战。在研究的实际落地过程中, 如 Saha 等 [60] 提出多模态方法的临床应用可能是局限的。由于多模态 MRI 采集时扫描时间较长, 患者轻微的身躯运动可能会导致多模态图像在空间上的失配。因此, Feng 等 [61] 仅基于 TCGA 数据集中的 T1c 单一序列图像就构建了端对端的 IDH 分类 3D 网络, 其在交叉验证中达到 97% 的精度。Yogananda 团队 [62] 在胶质瘤的 1p/19q 状态识别中提出了一个 3D U-Net 用于分类, 达到 93.46% 的准确度。然而这样一来是丢失了医学图像任务中伴有的多模态数据集优势。还有就是, 在临床实践中, 密集的 3D 扫描序列获取成本较高, 很大程度上限制了 3D 网络的应用。

通过整理以上的文献内容, 肿瘤分子分型预测模型的研究现状可以总结为以下几点:

(1) 现有的分子分型预测研究仍停留在私有小规模数据集上。目前肿瘤分子分型相关的研究分散在医学各种类型的肿瘤数据集上, 导致研究领域内可用公开数据集一直处于缺失阶段。对于脑部胶质瘤领域, 相关数据集规模大多不足 300 例, 分子病理标签也存在大量缺失, 这样对于预测模型性能的泛化性而言, 有比较大的挑战。

(2) 目前尚缺少对胶质瘤 Ki-67 指数的端对端预测模型研究工作。肿瘤内 Ki-67 指数是 WHO 分级的重要参考之一, 可以作为评判患者治疗预后的重要指标。若能以非侵入式预测 Ki-67 指数水平, 对于患者的个性化治疗有显著的指导意义。尽管有基于手工特征提取的方法, 能有效预测 Ki-67 指数水平, 然而其在性能鲁棒及便捷实现上都远不如以神经网络为代表的深度学习方法。

(3) 深度学习模型对医学 MRI 序列影像间相关性的利用方式大多粗糙。现有的深度学习方法常采用早期融合的方式来处理医学 MRI 影像中的多序列图像, 这对于序列间的共性特征的捕捉很不全面, 甚至导致特征信息的冗余。

## 1.3 研究目的与主要研究内容

### 1.3.1 研究目的

高病发率的脑神经胶质瘤在后期阶段致死率极高，因此胶质瘤的早期诊断研究愈发重要。目前通过活检等组织病理学检测耗费巨大，且缺少临床辅助诊断的应用，相关的数据集及系统算法研究也存在空缺。本文旨在建立起一个依托湘雅医院的胶质瘤分子分型数据集，并在此基础上开展 Ki-67 指数的非侵入式系统诊断算法研究，致力于相关研究成果的落地部署。

### 1.3.2 研究内容及本文结构安排

本文研究的主要内容是如何建立完整的适用于胶质瘤非侵入式预测算法的胶质瘤分子分型数据集，并凭借这一条件，研究 Ki-67 指数的临床辅助诊断系统相关算法设计实现，具体分为下述的研究方面：

(1) 构建并整理出完整的胶质瘤分子分型数据集。本研究联合湘雅医院神经外科及影像科，以真实就诊患者的临床数据为基础，具备其临床元信息，MRI 影像及分子测序结果的胶质瘤分子分型数据集。在构建过程中，如何对数据进行清洗以得到更高质量数据集，如何采取研究必备的预处理手段来便捷适用相关的科研工作，都是本文工作的重点，此外还需要对构建的数据集在基准任务上评估其可行性及适用性。

(2) 胶质瘤病灶分割模型的可解释性建模研究设计。肿瘤病灶中包含最丰富的肿瘤信息，所以智能勾画出感兴趣区域 (Region of interest, ROI) 是前期工作。而如何利用医学临床中的先验知识来赋予分割模型更多的可解释性是研究重点。同时，病灶分割模型泛化性能也需要在跨域数据集上完成验证。

(3) 多序列胶质瘤 Ki-67 指数预测模型的研究设计。MRI 影像天生带有不同的序列数据，本文工作的重要一步就是，设计有效的模型结构来充分挖掘 MRI 不同序列影像之间的同质特征。同时，如何融合患者的临床元信息与医学高维影像，如何给出鲁棒的患者级诊断结论以实现更完整的诊断，这些都是值得探索的方向。

(4) 临床辅助系统的客户端应用研究设计。胶质瘤诊断算法的最终目的是应用于临床辅助诊断，提升医生的诊断效率。诊断算法的闭环设计以及诊断系统的落地形式都是本文工作的一部分。

全文内容共六章，全文框架图1-3所示，具体内容结构如下：

第一章：首先介绍脑胶质瘤分子分型（特别是 Ki-67 指数）诊断的研究背景与意义，从肿瘤病灶分割模型和肿瘤分子分型预测模型两条路线来概述国内外发

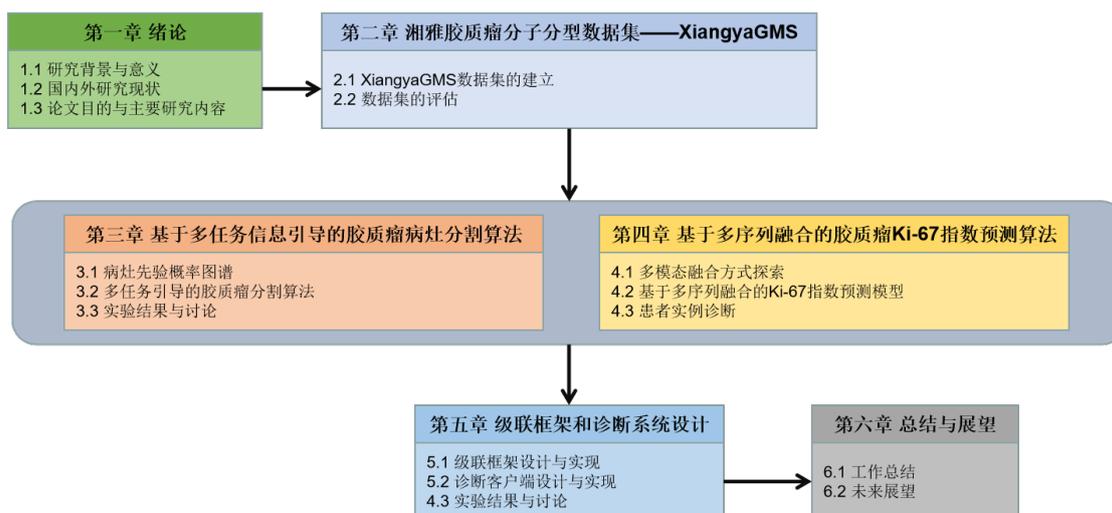


图 1-3 全文框架图

展现状，并归纳出该子领域中现有方法存在的问题及思路。最后，提出本文的研究目的，对主要内容进行总结，并概括论文的内容结构。

第二章：介绍胶质瘤分子分型数据集的建立过程及评估。本章介绍了数据集的获取途径及方式，对数据集的清洗、标注及预处理手段都给出科学且流程化的方法。最后，在本文构建的数据集上针对医学基准任务进行指标评价，验证其在该领域的可行性和适用性。

第三章：探索胶质瘤病灶分割模型的合理性。本章工作是希望在模型建模过程中结合医学的临床先验知识，赋予分割模型一定的可解释性，同时在公开的跨域数据集 BraTS 上验证分割模型的泛化性能。

第四章：研究 MRI 多序列胶质瘤 Ki-67 指数预测模型结构的有效性。本章层次性地验证 MRI 多序列影像之间的联系，并针对合理的融合方法，设计一种新的多序列模型来提升模型的预测性能。此外，提出了一个新的二元多尺度损失函数来监督来自同一患者数据的序列间相似性。最后，本章将根据单张 MRI 影像的预测结果，在不同的融合方法上验证其患者级预测结论的实际性能。

第五章：设计并实现临床辅助系统的客户端联网应用。本章将设计的 Ki-67 指数预测模型和胶质瘤病灶分割级联成一个闭环设计，同时，将级联后的诊断框架部署在客户端联网环境下，实现相关功能的设计，同时改善用户的使用体验。

第六章：本章总结了全文的主要工作，展望了下一阶段的研究方向。

## 第 2 章 湘雅胶质瘤分子分型数据集——XiangyaGMS

在胶质瘤分子分型研究领域，一直是缺少一个高质量且齐全的研究数据集。国内外已有的公开数据集（BraTS, TCGA 等），或是不包含完整的分子病理结果，或是缺少原始的医学影像，这对于整个胶质瘤领域的研究发展是十分遗憾的。本研究联合湘雅医院，以湘雅医院近四年的真实就诊患者为数据来源，构建了包含患者 MRI 数据，基因组学及临床问诊信息的湘雅胶质瘤分子分型数据集——XiangyaGMS (Xiangya Glioma Molecular Subtype, XiangyaGMS)。本章将从数据集的获取途径、数据清洗、标注及后处理方面来描述数据集的建立过程，同时在该数据集上验证了其在医学基准任务上的性能，审视建立 XiangyaGMS 的重要意义。

### 2.1 XiangyaGMS 数据集的建立

XiangyaGMS 数据集主要由患者的多序列 MRI 数据及其肿瘤病损区域标注、其对应的分子亚型结果和临床问诊元数据构成。本节从患者 MRI 影像的预处理及病损 ROI 标注来描述 XiangyaGMS 数据集的建立，具体流程见图2-1。

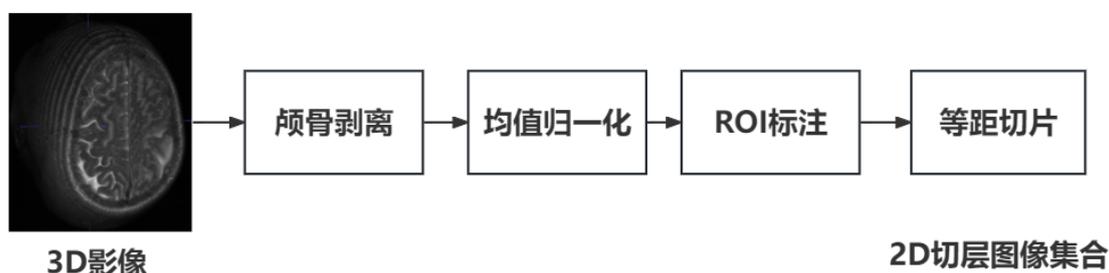


图 2-1 XiangyaGMS 数据集建立流程

#### 2.1.1 数据的获取与清洗

神经胶质瘤患者的数据来源于湘雅医院 2017-2020 年期间的真实就诊案。所有的 MRI 影像数据都是在湘雅医院的标准环境下捕获，导出的数据中包含以下几种分辨率：512×512, 640×640, 416×512 等。总体的患者病例数达到 619 例，包含 T1、T1c 和 T2 三个序列的横断面视角 DICOM 格式数据。XiangyaGMS 数据集中的患者不局限于 WHO 的 III 级或 IV 级，其分子标记物诊断结果都经过了严格的基因测序检验，并有相应的诊断报告作为依托。

图2-2给出了 XiangyaGMS 数据集的某患者的完整数据，其中包括有 T1, T1c 及 T2 的轴向面视角影像，此外还附有 T1c 序列的冠状面及矢状面视角影像。从图中可以看出，针对于轴向面（横断面）的序列影像来说，脑部 MRI 影像中除

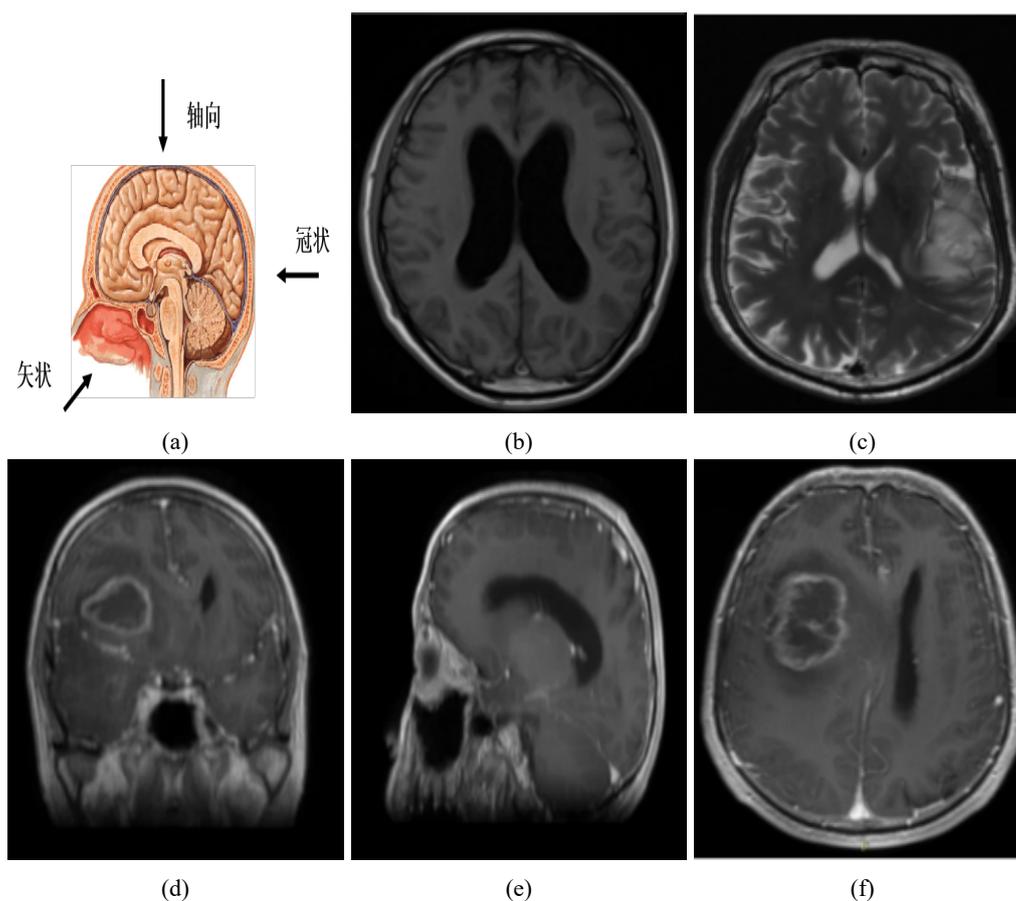
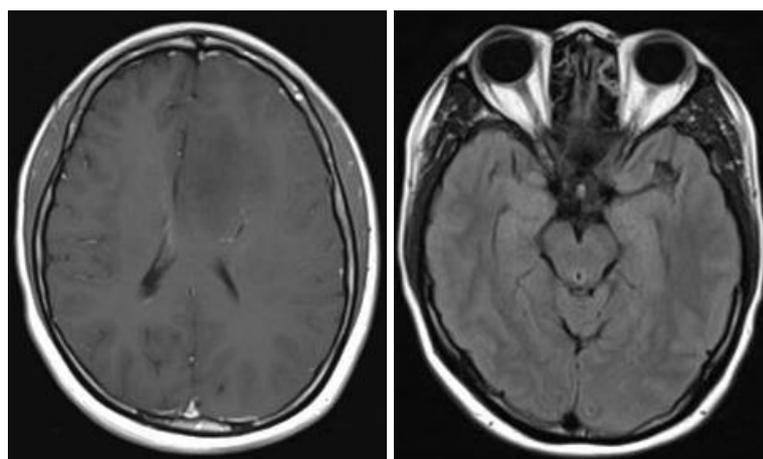


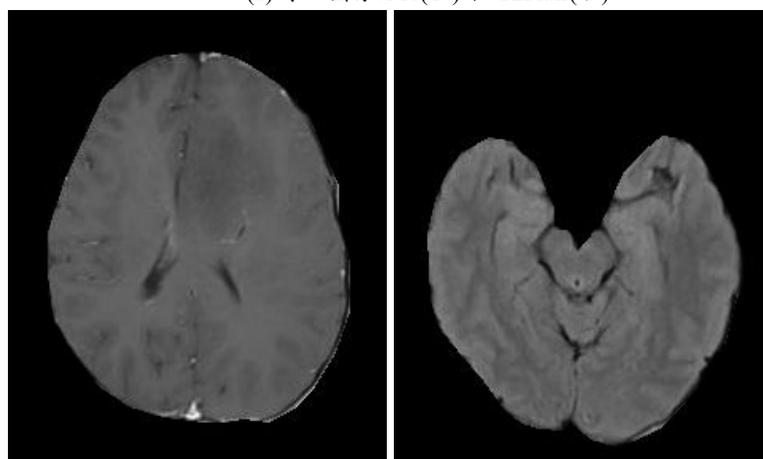
图 2-2 XiangyaGMS 数据集中某患者的示例。(b) 为 T1 序列某切层图像；(c) 为 T2 序列某切层图像；(d), (e), (f) 分别为 T1c 序列的冠状, 矢状, 轴面向切层图像

除了要提取的脑部组织之外, 还包含有以颅骨结构为代表的非脑部组织成分。在医学影像的标准化过程中, 由于这些非脑组织与真正的脑组织所属的灰度值十分相近, 这会极大影响脑组织的标准化结构, 给放射基因组学特征的捕捉增加难度。所以在医学影像研究任务中, 相比于其他的医学先验处理功能, 对数据进行颅骨剥离是研究开展的必要前提之一。本文没有采用略显繁琐的深度学习方法, 利用 3D Slicer 4.5 软件中的 Swiss Skull Stripper 拓展模块就可以快速实现自动剥离医学 MRI 影像中的颅骨结构。如图 2-3 所示, 脑部影像中的颅骨结构能够完全剥离, 符合后续研究预期的效果, 虽然能完全剥离颅骨结构, 但对于部分边缘的脑部结构也被除去, 这一点不会对后期研究工作造成较大影响。

原始的 DICOM 格式 MRI 影像调用 Pydicom 库读取后表示为单位厚度的灰度图, 医学影像往往通过不同的灰度值来反映脑部组织的病理或者解剖信息, 相关的研究通过分析灰度的密度及其分布可以对病损区域的现状有一定了解。其中灰度值的差异来源主要是两方面: 成像因素误差和成像对象本身。不同的医学成像因素差异会导致同一类脑部组织在 MRI 影像中的灰度值不一致。在放射组学研究中, 一般会考虑对医学 MRI 影像作一定的操作, 来抑制来自参数或人为的



(a) 原始图像:T1c(左)和 T2Flair(右)

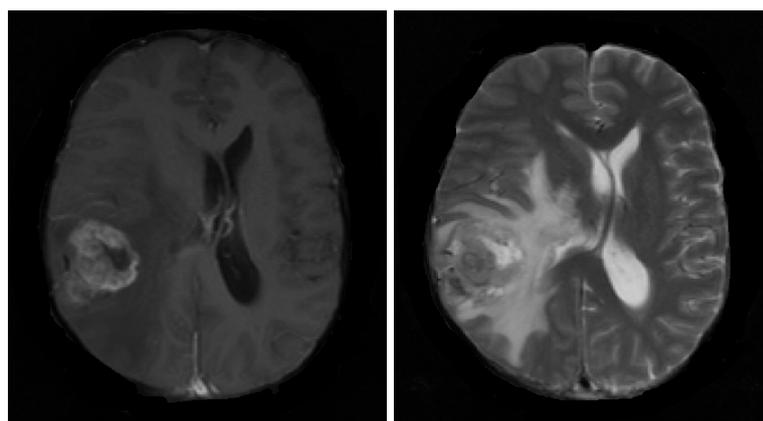


(b) 颅骨剥离后图像:T1c(左)和 T2Flair(右)

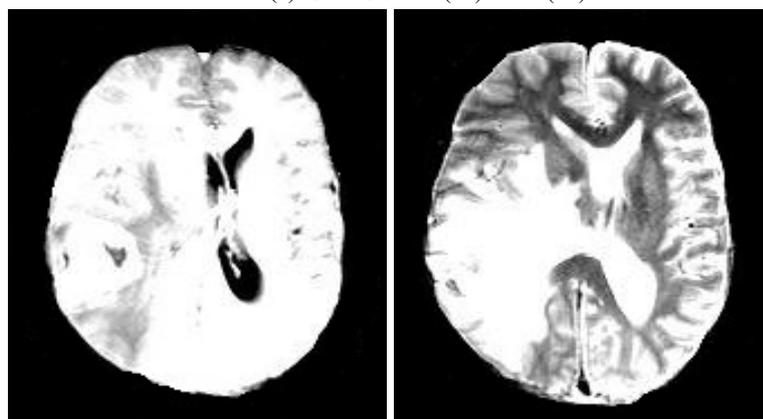
图 2-3 MRI 影像的颅骨剥离前后示意图

这种成像因素误差，使得在研究中更突出集中在成像对象本身（脑部组织）。这个操作就是均值方差归一化，旨在减小 MRI 影像中成像因素误差带来的灰度值不一致，同时保留脑部组织间灰度差异。

在医学 MRI 影像中的归一化操作需要先筛选出包含脑部组织的前景区域。本文将在同一患者的某个序列影像内操作归一化，这样做能把脑部组织间差异更精准地反馈到灰度值差异上，有利于后续的胶质瘤分子分型精准预测及诊断。如图2-4给出医学 MRI 影像归一化前后的效果对比图，其中为观感更明显。为可视化需要，图2-4将归一化后的影像像素矩阵与常数 255 执行向量数乘，可以看到脑部结构中脑脊液等组织与病损区域呈现可分的对比差异。除此外，归一化还有利于神经网络模型的整体收敛。



(a) 原始图像:T1c(左)和 T2(右)



(b) 归一化后图像:T1c(左)和 T2(右)

图 2-4 MRI 影像归一化前后的效果对比图

### 2.1.2 数据的标注与后处理

放射基因组学的常规步骤中包括对 ROI 区域的分割，在胶质瘤分子分型预测的研究中，ROI 区域对应的就是胶质瘤所在的病损区域，其中最有望捕捉到与分子标记物预测相关的诊断信息。对于 ROI 区域分割算法的研究，由于研究中需要处理的影像数据规模较大，手动对每张医学 MRI 影像进行分割工作量太大。目前的解决方案是倾向于自动分割算法，而 XiangyaGMS 数据集的分割标签标注更多是凭借专业放射科医师团队的经验完成，这些医生都有 4 年以上的实操经验。在标注过程中，标注团队成员会被分为两组，每组在负责各自的 MRI 影像标注的同时，还会复查来自另一组的影像标注结果，确保标注结果具有较高的可信度。标注的工具是可获取开放源码的图像标注软件 labelImg。分割标签的标注对象笼统来说是整体的病损区域，在 XiangyaGMS 数据集中，病损区域被更精准包含 GD-增强肿瘤和坏死肿瘤核心，如图2-5所示。区别于 BraTS 挑战赛中的分割标签，瘤周水肿（浸润组织）被排除在外。所得的分割标签标注经过或操作后创建，即共同注册到相同的 ROI 区域分割掩膜。

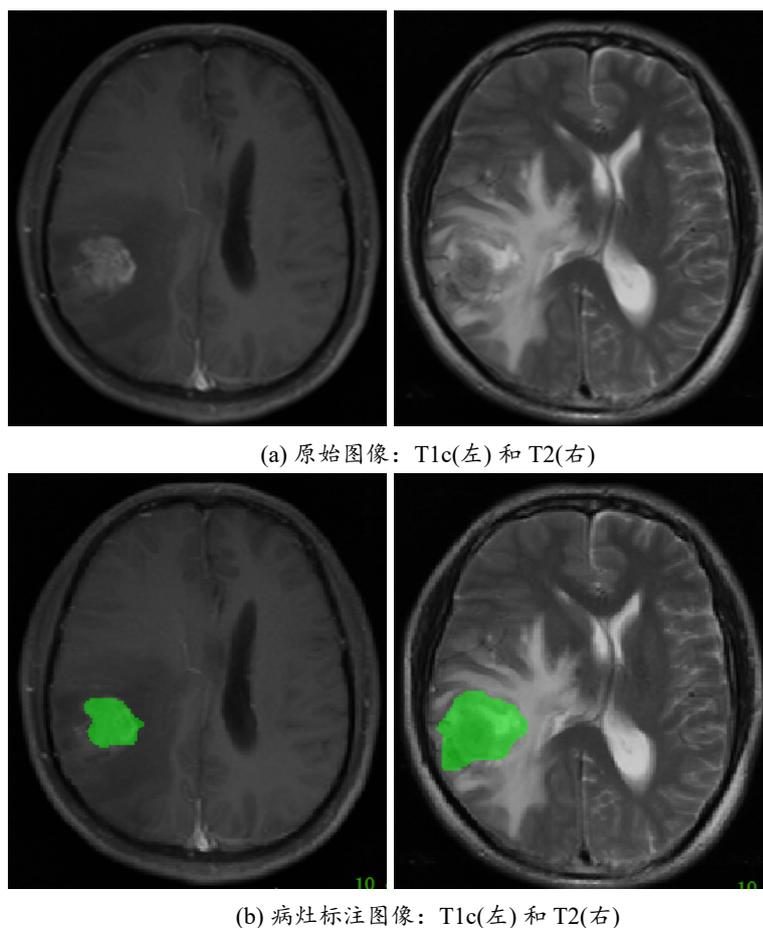


图 2-5 XiangyaGMS 数据集中胶质瘤病灶标注示意

上文有提到 XiangyaGMS 数据集中有整理到完整的 T1、T1c 和 T2 序列，病变区域的表现不同 MRI 序列影像明显不同。从 T2 影像中异常信号的手动描绘开始来主要定义整个肿瘤，之后再通过 T1c 影像对其中明显低信号的增强型和坏死型进行匹配。所以，不同 MRI 序列影像之间的交互信息再一次被证实和利用。对于获得的原始患者数据，其影像呈现 DICOM 格式，且同一患者对应的不同 MRI 序列影像的切层厚度不一致，导致 DICOM 格式影像的数据规模不一致，这一点会对本文的交互信息捕捉研究目的带来很大的干扰。在数据后处理过程中，本文首先将单个序列的体素插值到  $1\text{mm}\times 1\text{mm}\times 1\text{mm}$ ，并按照等距来得到空间配准的  $K$  个切层集合，本文选取的  $K = 20$ 。虽然 T1c 在成像上保持着与其他序列相同的脑部组织，但在其注射增强过程中几乎必然会带入偏移量。这里本文参考 T2 序列影像对 T1c 影像中的关键点进行匹配，这样的处理后的数据集在空间上默认为是配准的，所以能够为不同序列 MRI 影像的交互信息提供重要的数据保障。最终本文可以整理出一个包含 619 例患者，符合研究要求的 XiangyaGMS 数据集，其中 T1、T1c 和 T2 三个序列各自都达到了 12380 张。在专业医生团队经验建议下，本文中主要使用 T1c 和 T2 两个序列影像会富有更多医学诊断价值。

如表2-1给出了患者的临床元信息和分子分型诊断结果的统计结果。XiangyaGMS 数据集中所有患者的年龄均值为 45 岁，表中将小于 45 岁表示为 0 类，反之则为 1 类。同时把女性表示为 1 类，反之则为 0 类。

表 2-1 XiangyaGMS 临床元信息和分子分型诊断结果统计信息

	年龄	性别	IDH 突变状态	Ki-67 指数	MGMT 启动子甲基化
0类	236例	388例	499例	191例	356例
1类	383例	231例	120例	428例	263例

### 2.1.3 数据集的划分与架构

XiangyaGMS 数据集是之后胶质瘤 MRI 影像分子分型研究的必要基础工作。在本节中需要先确定 XiangyaGMS 数据集的固定划分，这样能更好的统一评估对比实验的效果，能衡量模型的好坏。为了保证不在模型训练中发生“信息泄露”的情况，本文需要以患者为单位去确立基准数据集。本文将 619 例患者数据按照 6: 2: 2 的比例相应地拆分成训练集、验证集和测试集。训练集包含 371 例患者数据，总计 7420(371×20) 组数据用于算法模型的训练阶段，而验证集和测试集中各自包含 124 例患者，合理的验证集设置会避免训练参数的复杂化，可以得到更好的模型泛化性能。

XiangyaGMS 数据集主要包含四个文件夹 GMSRaw、GMSBraTS、GMSGlioma 和 MetaMS，各部分简介见图2-6中。其中 GMSRaw 中包含来自湘雅医院的原始 DICOM 数据，除 XiangyaGMS 数据中的 T1、T1c 和 T2 序列影像外，还包含部分患者来自冠状面和矢状面视角的 T1c 序列 DICOM 数据，同时还有不完整的部分 T2Flair 数据。GMSBraTS 是病损区域分割任务使用的，首先建立 T1、T1c 和 T2 序列各自的文件夹，然后在各自文件夹中以患者编号创建文件夹。将其中的序列影像放缩到与 BraTS 2021 数据集相同的 240×240 的分辨率，其中用“.h5”格式来保存某一切层影像的字典，并以关键字分别保存影像矩阵以及其病损区域的掩膜矩阵，以方便算法调取。胶质瘤分子分型预测任务中使用的数据集是 GMSGlioma，对应的各个序列切层影像标注中包含有前景（实际的病损区域），其余与 GMSBraTS 文件夹的格式相同。而 MetaMS 文件夹中包含一份“.txt”文件，来源于专业医师团队，对应 619 例患者的临床元信息及分子分型基因测序结果。

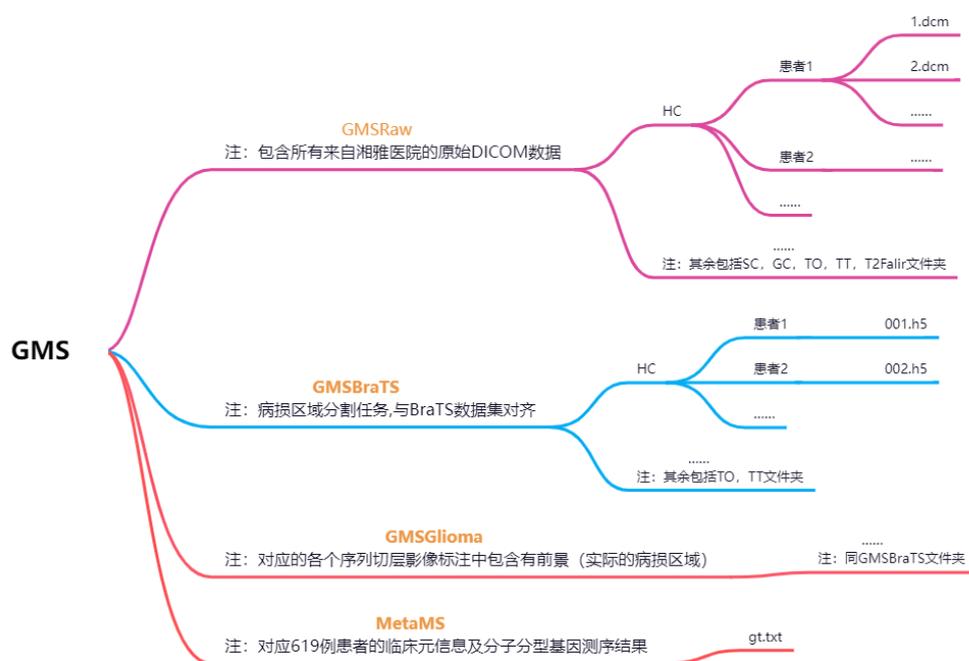


图 2-6 XiangyaGMS 数据集的文件架构

## 2.2 数据集的评估

### 2.2.1 数据集的基准实验

在上文工作中整理得到了用于胶质瘤分子分型研究的 XiangyaGMS 数据集，有望成为该领域后续进一步研究的基础。在开展本文后续的放射基因组学研究工作之前，需要验证基准任务在 XiangyaGMS 数据集上的性能，防止数据集的不规整性导致的性能崩溃。

本节设置的基准任务是胶质瘤存在性的诊断，一方面可以检验 XiangyaGMS 数据集的规整性，另一方面也可以为医师的先验判断作一轮标签的复查。本文希望能评估出 XiangyaGMS 数据集在常见 CNN 下的性能，获取其基准表现。本节选择了三种常用的具有不同基础结构的 CNN，包括 ResNet-50[63]、DenseNet-121[64] 和 Xception[65]，基于 XiangyaGMS 数据集就 MRI 中胶质瘤存在性实现二分类任务。这个任务其实等价于病损区域的存在性，所以这个基准任务的标签在实现病损区域标注后可以便捷得到。由于病损区域的临床判断还是凭借肉眼观察，辅以经验，所以 CNN 在这个任务上的可行性是存在预期的。

为了充分验证 XiangyaGMS 数据集的合理，本节采用五折交叉验证。整个训练过程是在两块 NVIDIA GeForce RTX 4090 的显卡上完成，实验条件均保持一致性，比如：均在 ImageNet 数据集上加载了预训练权重，最大训练次数 400，初始学习率 0.001，批次大小为 64，优化器设置为 Adam，损失函数采用经典的交

叉熵损失函数，如公式2-1所示：

$$L = -\frac{1}{N} \sum_i^N [y_i \cdot \log(p_i) + (1 - y_i) \cdot \log(1 - p_i)] \quad (2-1)$$

其中， $y_i$  表示样本  $i$  的标签，正类为 1，负类为 0。 $p_i$  表示样本  $i$  预测为正类的概率。

### 2.2.2 评价指标

本节使用到的评价指标为以下几种形式。

(1) 混淆矩阵 (Confusion Matrix, CM)。混淆矩阵，在机器学习领域，是一种特定的表格布局，用来可视化统计分类模型的性能。矩阵的每一行代表实际类中的实例，而每一列代表预测类中的实例，这样能很容易地看出算法模型在每个类别上的具体混淆比例，如表2-2所示。通过混淆矩阵可以轻易计算得到其他常用的评价指标，如准确度和召回率等。相比于其他指标，混淆矩阵能进一步分析具体的性能结果，比如模型对哪一种类别的预测会更弱势一些，这些具体的细节对模型或数据的进一步调整会有很大帮助。此外，混淆矩阵对角线上的实例越多，意味着算法模型的识别性能越好。

表 2-2 混淆矩阵示意

		预测标签	
		阳性	阴性
真实标签	阳性	正确阳性预测 (True Positive, TP)	错误阴性预测 (False Negative, FN)
	阴性	错误阳性预测 (False Positive, FP)	正确阴性预测 (True Negative, TN)

(2) 准确度 (Accuracy, ACC)。准确度通常可以作为分类任务中测试模型识别性能程度的统计度量。从医学上说，准确度是正确预测（真阳性和真阴性）在检查的病例总数中所占的比例。当然对于医学二维影像，评价对象更倾向于单独的影像，而非病例，如公式2-2所示：

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

(3) 召回率 (Recall)。召回率，也称灵敏度，查全率，它主要衡量的是模型对于正样本的预测结果。其具体表现为正确预测的正样本在所有正样本总数中所占的比例，公式2-3所示：

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

### 2.2.3 基准实验性能结果

最终的实验结果如表2-3所示，这三种经典的 CNN 都可以达到不俗的性能效果，其中 Xception 的性能会表现得更出色。图2-7是 Xception 的混淆矩阵，从中可以看出，混淆矩阵的深色区域主要分布在对角线上，而其他区域的颜色相对较浅。本节校对图中的浅色区域的实例发现，测试实例中的肿瘤病灶周围若出现瘤周水肿（浸润组织），其预测结果会有更多的误判。这可能是因为 XiangyaGMS 数据集在标注时并没有包含浸润组织区域，而这一部分与实验标注的标签之间的联系是紧密的，甚至在部分文献中是被归为标签类别的。值得一提的是，本节的目的并不是找到胶质瘤存在性的最优模型，而是验证 XiangyaGMS 数据集的规整性及有效性，确保其可以能够成为后续深入研究的基础条件之一。

表 2-3 三种经典模型在 XiangyaGMS 数据集上的性能

	Resnet50	DenseNet121	Xception
准确度	0.906	0.915	0.939
灵敏度	0.879	0.886	0.915

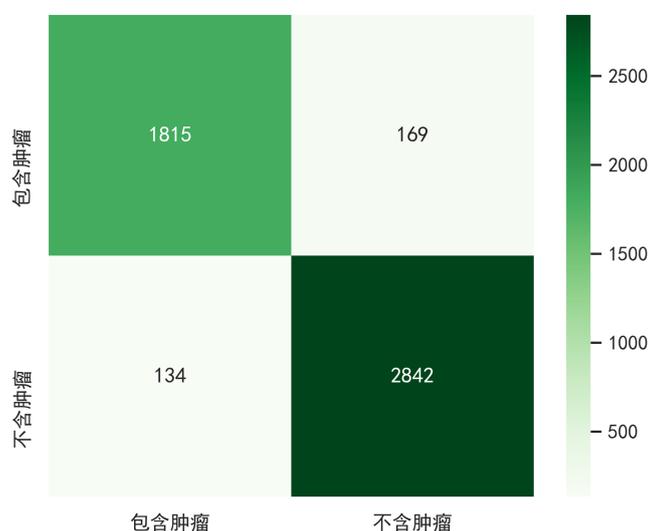


图 2-7 Xception 的混淆矩阵

## 2.3 本章小结

本章主要整理出一个全新的，带有胶质瘤分子标记物基因测序标签的胶质瘤分子分型数据集——XiangyaGMS，收据来源于湘雅医院近 4 年的 619 例真实

就诊患者，包含每个患者的 T1、T1c 和 T2 序列影像以及对应的临床问诊信息。此外，针对分子分型非侵入式预测，XiangyaGMS 数据集包含了完整的 IDH 及 Ki-67 病理标签，其中影像规模就接近 4 万。本章具体介绍了数据集的获取清洗、标注以及后处理的实现，并给出了 XiangyaGMS 数据集的文件架构，便于后续研究的展开。最后就 XiangyaGMS 数据集的规整性和有效性在基准分类任务上实现了评估，三种经典的 CNN 都能达到预期的识别性能，XiangyaGMS 数据集可以作为后续研究的一项重要支撑工作。

### 第 3 章 基于多任务信息引导的胶质瘤病灶分割

医学胶质瘤影像中的放射基因组学特征的提取是一个不易的工作过程。在医学 MRI 影像智能分析领域中，MRI 影像中的胶质瘤病灶（特别是核心区域）特征与 Ki-67 指数预测任务有重要的关联性。同样，专业医师团队及领域内研究文献中也给出相应的描述：相比于颅内其他的脑部组织，胶质瘤病损区域具有更重要和全面的诊断价值和临床意义。而目前在临床中，病损区域的标注一般是由专业的医师完成，大批量的临床影像会进一步增大医务人员的工作负担。本章将探索一种基于 MRI 的胶质瘤病损区域分割模型，该模型能够在 MRI 影像上找到并标注出胶质瘤病灶。这对于将其作为胶质瘤 Ki-67 指数预测模型的预处理手段，以此形成一个智能诊断的闭环结构有重要临床意义。

现有基于深度学习对医学影像中病灶分割的研究工作，已经有了很大的突破，更深更复杂的网络模型也被应用到医学图像分割任务中。但是从本文的研究目的出发，目前的同类研究工作中尚不能很好地覆盖到需求面。

其一，旨在用于临床辅助诊断的工作，比较重要的一点是其决策结果的医学可解释性，神经网络模型本身是不包含任何解释性的黑盒模型。领域内比较认可的模型可解释性主要是在内在结构与决策事后比较上，从内在结构上看，神经网络模型中大量的神经元固缺乏解释能力，即使在简单结构中被验证的有效节点，其代表的实际意义很难被道明，甚至在更多的时候，单个有效节点几乎不存在，模型的有效性需要更多的神经元节点之间的组合才能被实现。而从决策事后比较上来看，简单的端对端神经网络几乎不具备推敲性，输入端的细微改动会引发模型输出的巨大变化。

其二，真实条件下的临床影像成像操作会引入较大的成像误差，所以模型的实际泛化性能需要被验证。医学影像的成像质量会受到采集过程中的人员，设备，环境的因素影响，这就对于固有数据集训练得到模型的泛化性能提出比较大的考验。在辅助诊断系统的实际应用过程中，临床应用场合各异，复杂的成像因素导致的不确定性太大，所以对性能的要求十分苛刻。

在本文的前期工作中，在湘雅医院帮助下，整理得到有效的胶质瘤分子分型数据集——XiangyaGMS。本章的工作将基于 XiangyaGMS 数据集，探索 MRI 影像中脑部肿瘤的病损区域分割研究工作，目的是使模型分割水平接近专业医生的标注精度，且能从建模角度引入更多地临床先验知识，为模型赋予一定的医学可解释性。

本章首先通过挖掘部分医学先验知识，提升模型的分割性能，从而为模型的决策事后比较中带来一定的医学可解释性。同时，引入多任务模型，把胶质瘤的

病损区域分割任务和其 WHO 恶性程度分级任务统一起来，在模型训练中引导两者之间的相关信息。最后，将训练得到的最优模型在跨域数据集上进行分割性能的验证，充分考虑其临床诊断可行性。

### 3.1 病灶先验概率图谱

医学图像具有许多解剖学先验，例如器官呈现相似的形状和位置等，这一部分的工作已经在基于传统机器学习的器官分割中有所崭露头角 [66]。区别于器官，由于病灶本身带有的偶然性和突发性，其形状大小无法定型。而长期的医学文献分析统计以来，胶质瘤的病灶位置“似乎”有一些偏好发生区域，比如额颞顶枕叶。在本章的工作中，希望在深度学习神经网络中结合解剖学先验知识来准确医学图像分割，这本身是一项具有挑战性的任务。

概率图谱常被用来描述医学图像中器官空间分布的统计信息，广泛应用于医学图像研究领域。本章提出一种新的用于深度学习模型的概率图谱先验知识，称为病灶先验图谱 (Focus Prior Graph, FPG)。在病损区域的分割任务中，模型的学习注意力被期望更多在阳性样本（即病灶）上。本章通过引入不对称的 FPG 损失来提高阳性样本的固有权重，也可以部分改善阳性样本在医学图像中占比不高的训练偏见。而常用的 FocalLoss[67] 虽然也能一定程度改善样本标签不平衡的问题，其在训练过程中极易受到样本噪声的干扰，这一点会限制其进一步适配性。贝叶斯定理有明确给出先验条件和似然的说明，笼统来说就是，在获取到先验知识的条件下，将模型参数基于现有的观察样本来最大化似然函数。本章所构建的 FPG 损失能有效结合先验损失和似然损失，以有效表征 MRI 图像中的解剖学先验，这一点还在前人研究中没有被涉及到。

FPG 损失的基础要求是获取到训练数据集中所有患者的 MRI 影像病灶标注，这一点 XiangyaGMS 数据集可以满足。同时需要对所有患者的 MRI 影像进行配准操作，需要执行到包括不同患者，不同序列影像以及不同切层图像范围内，并从配准集中创建出其参考模板。在正式步骤中需要将配准后的 XiangyaGMS 数据集的病灶掩膜投影到参考模板中的相同位置。病灶概率图谱就是通过对参考模板上的，专家标注的病灶掩膜执行归一化操作得到的，这样其中胶质瘤病灶出现在某个空间位置的先验统计概率就转化成了这个参考模板上点的值。

FocalLoss 中通过定义不对称的权重损失来改变神经网络的训练注意力，而其本身对于难易样本的选择全权依赖于神经模型的学习水平。网络难以区分的样本会提升它的训练损失权重，这种选择思路在噪声数据上会十分脆弱，让网络的训练被带偏到局部样本上。而在 FPG 损失组成上，本章会利用 XiangyaGMS 数据中的胶质瘤病灶概率图谱，来预先定义 MRI 影像平面位置中的难易发生点。胶

质瘤病灶概率图谱中的每个值代表该点发生病损的概率，所以本章假设，平面上难分像素点在 FPG 上的值会靠近 0.5，0.5 的概率意味着该平面像素点是否发生胶质瘤病损的先验概率是接近的。为了在 FPG 损失中减轻易分样本的权重，本章中选择高斯函数将病灶概率图谱转换为其对应的训练损失权重  $W_1$ ，如公式3-1所示。高斯函数曲线的特性可以描述为具有对称性并且由中心向四周不断减小数值。

$$W_1 = \exp\left(-\frac{(P_i - 0.5)^2}{2\sigma^2}\right) \quad (3-1)$$

其中  $P_i$  是概率图谱中描述目标像素点  $i$  所在位置出现病灶的先验概率，是控制其权重偏向强度的超参数。 $P_i$  的值在 0 到 1 之间，如果  $i$  点位置在病灶概率图谱中值为 0，则表示该点从未出现病灶；如果值为 1，则表示在 XiangyaGMS 数据中在该点都有病灶出现。值为 0 或 1 的点是权重最小的模型易分现象像素点，值接近 0.5 的点是权重最大的难分像素点，因为很难从先验知识的基础上确定它们是病灶还是非病灶。此外，本章工作的感兴趣的是阳性目标像素点，所以进一步

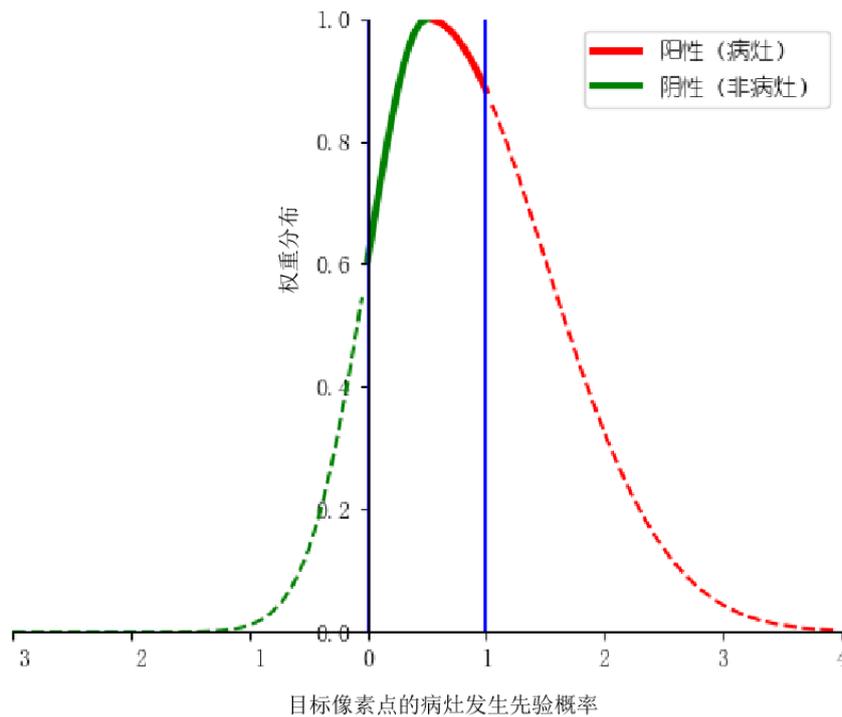


图 3-1  $W_{FPG}$  在 0-1 上的分布

将损失权重  $W_1$  描述成非对称高斯函数，以一定程度区分对阳性和阴性目标像素点的权重，用于训练的 FPG 损失如公式3-2所示：

$$W_{FPG} = \begin{cases} \exp\left(-\frac{(P_i - 0.5)^2}{2\sigma_{neg}^2}\right) & 0 \leq P_i \leq 0.5 \\ \exp\left(-\frac{(P_i - 0.5)^2}{2\sigma_{pos}^2}\right) & 0.5 < P_i \leq 1 \end{cases} \quad (3-2)$$

其中  $\sigma_{neg}$ ,  $\sigma_{pos}$  是控制阴性（非病灶）和阳性（病灶）目标像素点的权重偏向强度的超参数。由于本章研究着眼于胶质瘤病灶的精细分割，因此设置  $\sigma_{neg} < \sigma_{pos}$  确保阳性目标像素点的权重相对高于阴性目标， $W_{FPG}$  在 0-1 上的分布如图3-1所示。

## 3.2 多任务引导的胶质瘤病灶分割模型

目前大多数基于深度学习的脑部肿瘤分割模型，基本会被设定成用于独立任务的模型。独立任务这个前提假设在医学背景下是很难符合的，即使资深的专业医师在临床诊断时也需要“望闻问切”，基础研究中也同样需要结合病症的形状、位置等宏观信息。独立任务的这一假设之所以被普遍采纳，一方面是得益于深度学习模型出色的特征提取能力，另一方面是由于多数任务本身的实现复杂度不高。

对于较为复杂的，即需要从多个先验角度审视的医学任务，基于独立任务的模型容易忽略不同任务之间的相关性，比如肿瘤的分割和分级等，这样会阻碍深度学习模型对于目标任务的学习引导。除此之外，在目标任务的基础上设置医学辅助任务，也是一种借助医学先验知识的具体实现，能为用于目标任务的训练过程增加更多的可解释性，符合本章的研究目的。

### 3.2.1 多任务学习

多任务学习 (Multi-Task Learning, MTL) 属于机器学习的一个子领域，笼统将，其着力于同时训练多个学习任务，并利用不同任务之间的共性特征。就像 Caruana 等人 [68] 描述的那样，多任务学习是一种偏向归纳转移的学习机制，它利用训练相关任务中包含的任务间共性信息来提高每个任务的预测准确度和泛化性。鉴于此，考虑任务间相关性的多任务学习已成功应用于多种应用，如自然语言处理、计算机视觉和生物医学。多任务学习在不同研究领域中具有广泛适用性和可行性，其有效性体现在以下两个方面：

(1) 模型正则化：在 MTL 模型的过程中，训练量从原来的单个数据集仅被单个目标任务训练增长到同样的数据集被目标任务以及辅助任务训练。这种倍数增长的训练量对于模型来说，能真实有效地得到一个更具泛化性的特征集合。不同任务的训练过程中都会受到外面和内在因素导致的噪声问题，同时这类产生的噪声问题表现还各异。MTL 模型在把握整体损失最优的原则下，单一任务引入的单方面训练噪声会被“中和”部分，以保证模型在所有任务上都能取得出色的性能。相当于说，MTL 模型会在无形中引入归纳偏置作为正则化项，参数权重被合来满足其能在相关任务间捕捉共性特征。这种正则化是具备优势的，一般情

况下有望会胜过在目标任务上通过 L1 或 L2 正则化来惩罚其参数复杂性的效果。

(2) 模型训练引导：对于一个存在较多噪声的数据集或目标任务，从复杂的高维特征集合中挑选其中的有效特征对于模型来说是具有一定难度的。而辅助任务的训练中，可以作为重要特征筛选的一个额外参考。另一方面，辅助任务中相当于是从另一个优化目标去学习其对应的重要特征，寻到的重要特征是否有可能成为共性特征的一个相对可靠的候选集合，这对于目标任务的训练过程是一种正向建议的提示或引导。

虽然 MTL 的可行性及有效性已经被广泛认可，但其实际带来的性能提升却有差异，这也是目前所有领域中都会遇到的难题所在。一种具有普遍认可的说法是，如果引入的辅助任务和目标任务之间具有明显相关的共性，那么在随机欠采样的设定下会对目标任务的训练有比较突出的增益。而比较出乎意料的是，对于不相关的辅助任务设置，MTL 对目标任务的学习也能带来正面作用 [69]。

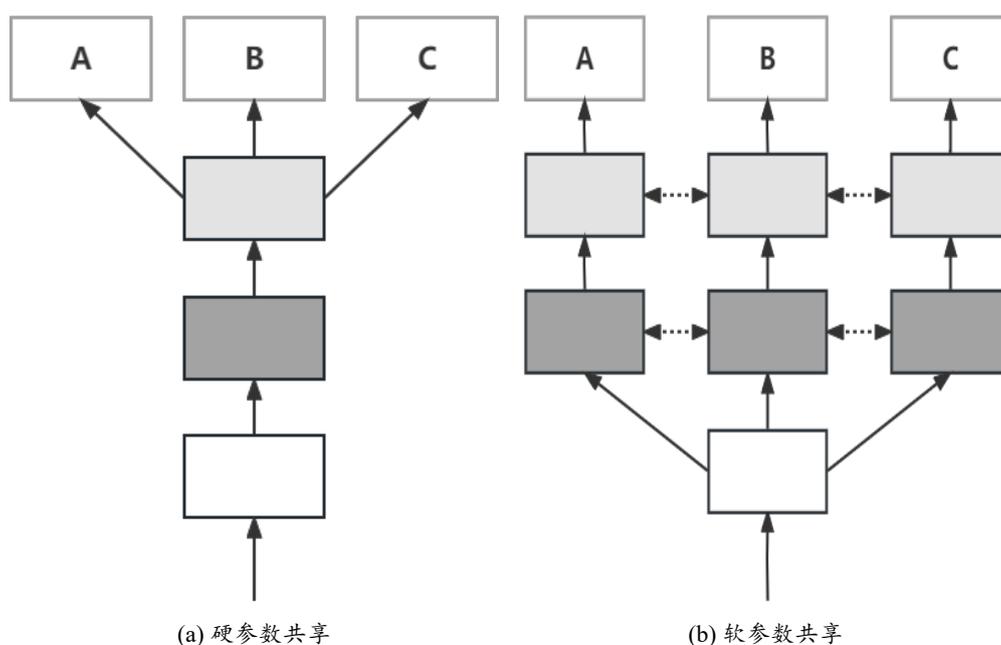


图 3-2 常见的多任务模型学习模式

目前研究文献中将 MTL 模型中针对隐层参数的学习模式总结为两种策略：硬参数共享和软参数共享，如图3-2-a和图3-2-b所示。其中比较最为常见好用的是硬参数共享，硬参数共享的形式，是通过所有任务共享一组隐藏层将来自多个任务模型的输入映射到同一语义空间，并保留几个独立隐藏层用于实现对应单一任务。原则上讲，MTL 模型中包含越多的辅助任务，能捕捉到共性特征的几率大一些，同时能降低在目标任务上的过合情况。相关研究中也结合癌症分割和病损区域分割，在皮肤癌症和新冠肺炎等方面取得了出色的模型性能。而在软参数共享的方式中，每个任务在独立的模型上各自训练，而模型中的参数信息是支持

共享的，在过程中可以获取其他任务的模型参数信息。软参数共享机制通常用于处理属于同一类型（如分类或分割等）的不同任务。然而，这类辅助任务的匹配性是非常有难度的，否则会影响目标任务的训练性能。

### 3.2.2 胶质瘤病灶分割模型

在本章中,探索一个多任务信息引导的网络模型 (Multi-Task Information Guidance Network, MIG-Net) 来实现胶质瘤的恶性程度分级和病灶分割。在专业医师的指导下,本章工作中会认为肿瘤是癌症恶性程度分级和病灶分割的直接因素。在传统的胶质瘤诊断过程中,专业医生往往通过会借助肿瘤特征(如形状和边缘)作为筛查肿瘤的恶性程度分级的证据之一,这些特征与肿瘤病灶分割中依赖的特征正好一致。MIG-Net 旨在基于胶质瘤的恶性程度分级任务,为目标任务病灶分割来提供包括粗糙定位和语义信息在内的双重先验知识指导策略,以增强任务相关性并促进信息流动。最后,对于分割模型中的解码器结构,本章采用一种注意力机制来增强上采样阶段的重要信息增益。

Swin-UNet[44] 在多器官分割任务上能取得杰出的性能和泛化能力。Swin-UNet 在编码器阶段能聚焦全局的自注意力,并在解码器中进一步探索上采样过程中的长程上下文信息,实现像素级分割预测。传统的 Transformer 模块由多头自注意力 (Multi Head Attention, MSA) 层 [40] 和多层感知器 (Multilayer Perceptron, MLP) 块组成,并通过层归一化 (Layer Normalization, LN) 来提高全局自注意力层的性能。Swin-UNet 模型的最重要主角是来自 [43] 中的 Swin Transformer 模块,其中提出的窗口注意力能将有关注意力的计算限制在每一个窗口内,而后也能凭窗口移位操作来做到独立窗口之间信息的交互学习,实现类全局的注意力。相比传统 Transformer 模块中动辄基于全局特征图来计算注意力的,模型的计算量能达到成倍缩减,对下游视觉任务的研究工作具有重要的指导价值。

如图3-3描述了 MIG-Net 的概况。MIG-Net 以 Swin-UNet 为神经网络分割骨架,在解码器结构中设计一个注意力堆来。同时, MIG-Net 在多任务分支上首先使用三个  $3 \times 3$  的深度可分离卷积来衔接特征提取器的共享特征向量。基于 FPG 的胶质瘤先验发生位置能够为各自任务分支的训练定位最佳的潜在表示信息。最后,将来自胶质瘤恶性程度分级的语义信息集成到病灶分割任务分支中,以引导胶质瘤病灶分割任务的学习。

对于 MTL 模型的共享结构,本章将 Swin-UNet 结构中的编码器模型作为特征提取器,用以对胶质瘤程度分级和胶质瘤病灶分割任务作联合训练。联合训练中,特征提取器的参数可以通过不同任务来更新。这样,将不同任务的数据表征嵌入相同的隐藏特征空间中,如此来自不同任务的噪声被部分“中和”,对于共

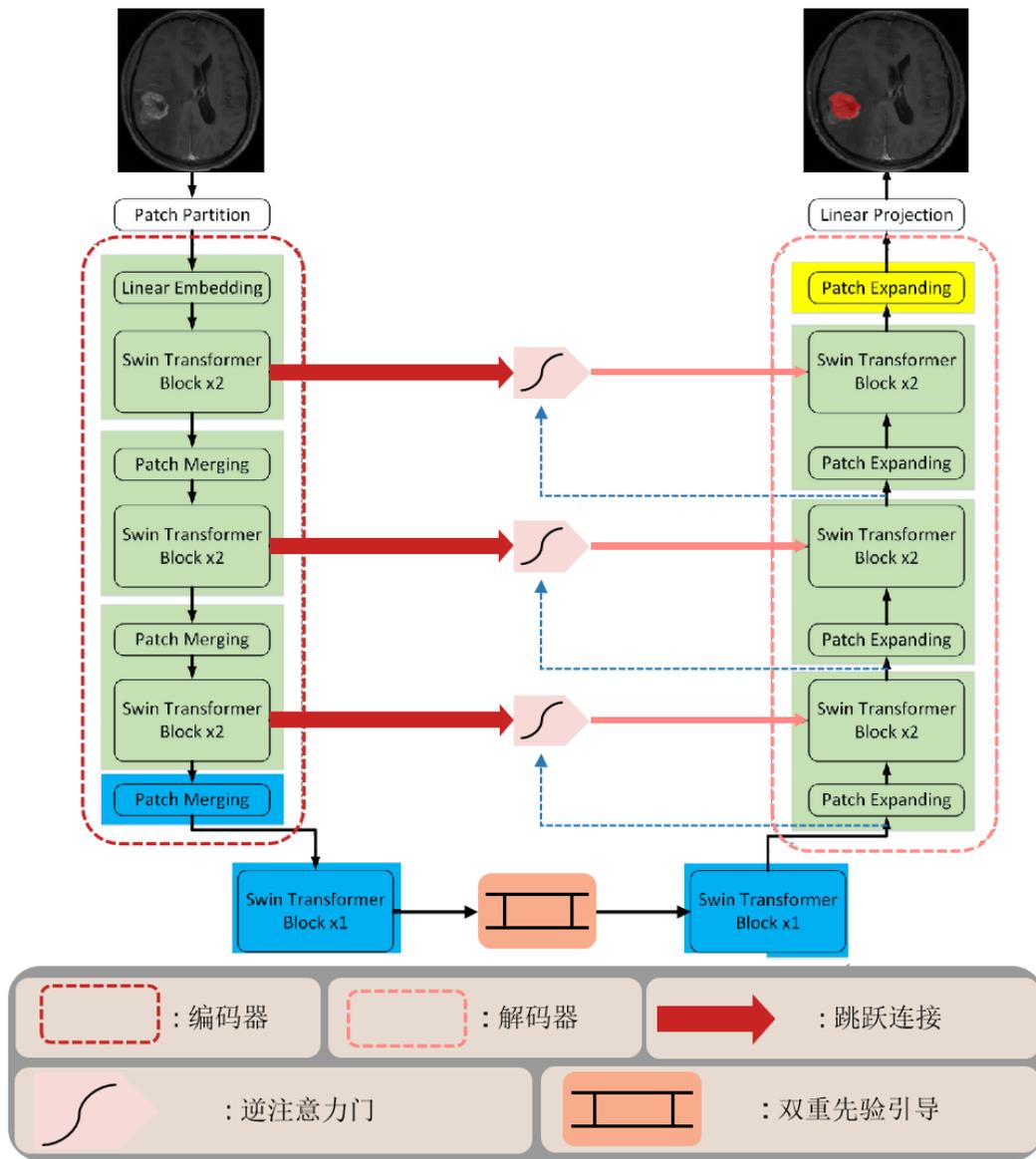


图 3-3 MIG-Net 的概况

性特征的捕捉是十分有利的。共享的隐藏特征表征会被分别输出到两个任务分支来实现具体任务的高级相关表征。

### 3.2.3 双重先验知识引导

在医学背景下，已经有不少团队已经开发了 MTL 模型，在指定的诊疗场合下能提升其目标任务的性能。他们中基本是采纳硬参数共享的学习模式，仅仅是通过共享隐藏层来学习共性特征，然后衔接一个输出层用于各自任务的训练。比较容易发现，这种 MTL 模型捕获的是早期隐藏层中的低级特征表示，也叫宏观特征，这会遗漏任务相关性涉及的高级特征表示。在 MTL 的模型结构上，通过调整其结构层次，可以满足在不同任务之间有选择性地共享相关性信息。一种改

进思路是，采用通道拼接或元素相加的操作来将辅助任务的信息表示引导到真正需求的目标任务上，一定程度可以带来额外的高级特征表示。但是这种高级特征表示的引导方式显然太过于直白，除需求的共性特征外，会伴随着侵入大量与辅助任务实现深度捆绑的“驳杂”特征，这些对于目标任务的性能并无好处。

旨在充分挖掘胶质瘤恶性程度分级和其病灶分割任务之间的相关性，基于本章的前期病灶先验图谱 FPG 工作，本节提出一种用于多任务的双重先验知识引导策略。该策略将通过 FPG 来引入胶质瘤的先验发生位置和先验语义信息，从而引导整体模型去关注胶质瘤病损区域，如图3-4。而在整个引导过程中，辅助任务和目标任务一样，有利于恶性程度分级任务的信息会被反馈到目标任务分支中，即这个过程中伴随着一个双向的训练引导作用。

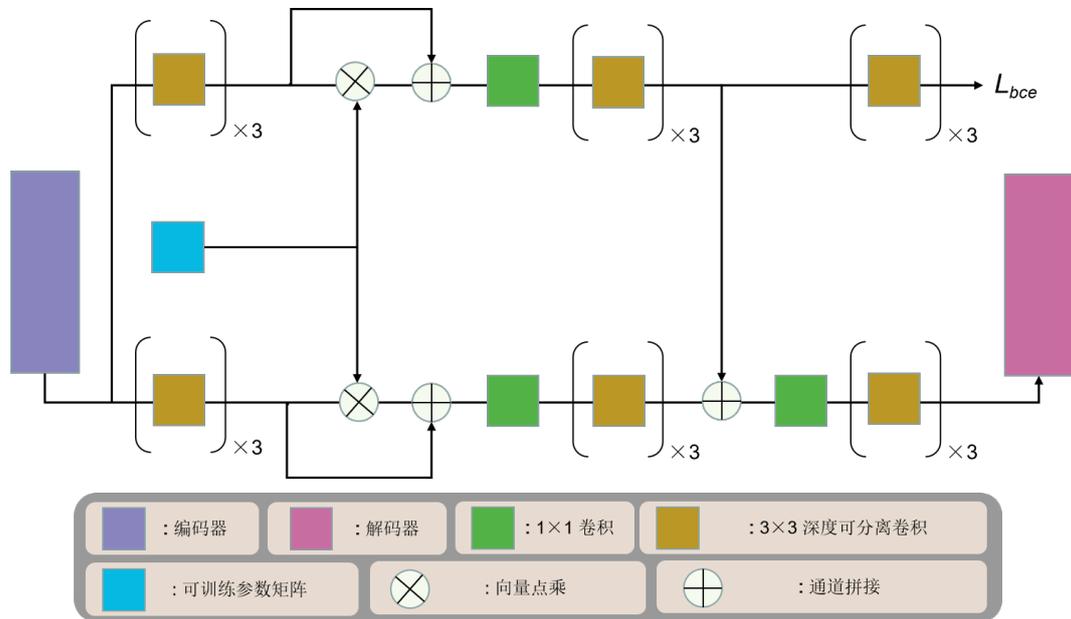


图 3-4 双重先验知识引导模块的概况

FPG，即病灶概率图谱，是记录胶质瘤发生位置的先验知识，天然可以作为模型训练过程中的定位引导。本节在两个任务分支中都应用 FPG 来促进模型对胶质瘤病灶的注意， $X_c$  和  $X_s$  是胶质瘤恶性程度分级和病灶分割任务分支，将共享特征提取器输出分别经三个可分离卷积层得到的特征图。矩阵参数  $\alpha$  是将 FPG 经与模型输入相同的分辨率操作得到的参考发生位置，研究中将其对应地初始化为包含 0-5 区间的概率映射数值，同时支持两个任务的训练更新。以公式3-3形式提供位置先验知识：

$$X_s = (\alpha \odot X_s) \oplus X_s \quad (3-3)$$

$$X_c = (\alpha \odot X_c) \oplus X_c$$

其中  $\odot$  是矩阵的元素乘法， $\oplus$  表示矩阵的通道拼接。然后，FPG 中会将具有较低先验发生概率的定位表示会被分配一个弱表示，对于下一个模块来说可以减少背

景的干扰。为维持原始的通道数，MIG-Net 采用  $1 \times 1$  卷积核来为特征维数降维。 $1 \times 1$  卷积核在实现通道降维的同时，通道之间信息的组合可以令网络能够学习更为丰富的表征。

包含更多与具体任务相关语义信息的高级特征表征都集中在输出层的输入中，本节将把胶质瘤恶性程度分级任务分支中的语义信息引导到病灶分割任务分支上，来增强不同任务在高级表征上的相关性捕捉。语义信息引导具体的实现如公式3-4所示：

$$Xs' = Xs' \oplus Xc' \quad (3-4)$$

其中  $Xc'$  和  $Xs'$  分别表示来自胶质瘤恶性程度分级和病灶分割分支的特征图。在以  $1 \times 1$  的卷积核进行通道降维输出到 MIG-Net 的解码器结构中。模型凭借两次知识引导，病灶分割任务分支能够在发生定位和高级语义两方面收到来自胶质瘤恶性程度分级任务的优化，无论是定位引导和语义引导都通过增强重要特征注意力来引入胶质瘤先验知识，从而聚焦于任务间的共性，同时保持目标任务的特异性。本章工作中，胶质瘤恶性程度分级任务的训练优化采用经典的交叉熵损失函数，其公式如2-1所示。

### 3.2.4 逆向注意力门

尽管以 U-Net 为代表的 U 型结构模型在医学分割中取得良好的性能，但这类多层级联的神经网络，容易造成模型参数量及计算量的冗余。如在 MRI 的 2D 切片中，胶质瘤的病灶必然存在形状、大小的较大差异，级联模型中多个层级提取的特征图中存在大片重复的低级特征 [70]。这种冗余的问题不仅带来资源的耗损，对模型的泛化性也会有所局限。旨在解决这个问题的相关研究工作更多是集中在 NLP 和知识图谱领域中，用于训练过程中的注意力分为硬注意力和软注意力。硬注意力是更多是一种区域的迭代建议生成，依赖的是出色的迭代策略 [71]。而软注意力则是利用反向传播关注到概率层面，比如通道注意力 [72]、空间注意力 [73]、自注意力 [74] 等，都能带来更鲁棒的模型性能。最近，Oktay 团队 [75] 的相关工作在医学器官分割领域取得突出的表现，能支持在 U 型结构上的扩展，以提高模型对前景目标的灵敏度。

针对特征冗余问题，本节将设计一种变种的逆向注意力门结构 (Inverted Attention Gates, IAGs)。IAGs 可以在执行时生成粗定位，突出局部有效特征，同时抑制不相关区域中的特征，有助于提高模型的灵敏度和准确度。同时 IAGs 结构本身不会带来太大的计算资源开销和庞大参数量。在本章提出的 MIG-Net 模型中，来自编码器的多尺度特征表示反映为逐层级的高维表征，其中  $l$  层的特征经非线性激活函数得到特征图  $X^l$ ，将通过跳过连接与补丁扩展层的解码器  $l$  层执

行特征融合操作，以恢复特征图的空间分辨率并实现像素级分割预测。本节的工作，在 MIG-Net 的基础上添加了 IAGs 模块，以增强对局部显著特征的注意力，如图3-3所示。

MIG-Net 会逐层对输出的特征图执行下采样操作，并通过提取上下文特征逐层解码得到像素级的分割预测。层级结构会导致在不同尺寸特征图输出时存在重复的初级特征，这对于形状大小差异较大的发生病灶，模型的灵敏度会受到影响。这类相似的问题，在目标检测领域，比较经典的方法是依赖于生成局部区域建议，然后对局部区域作单独的精准定位和识别任务。IAGs 也旨在逐步抑制不相关的背景区域特征响应，突出局部区域的显著特征。IAGs 模块的宏观输出表现为输入特征图和注意力系数矩阵的元素乘法操作，如公式3-5所示：

$$\hat{X}^l = X^l \odot \alpha^l \quad (3-5)$$

其中是  $X^l$ ， $\hat{X}^l$  分别对应着编码器和 IAGs 模块增强后的  $l$  层输出特征图， $\alpha^l$  是用于描述模型  $l$  层 IAGs 模块功能的注意力系数矩阵。 $\alpha^l$  将用于控制 IAGs 输出局部的增强显著特征，其对应的值域是 0-1。

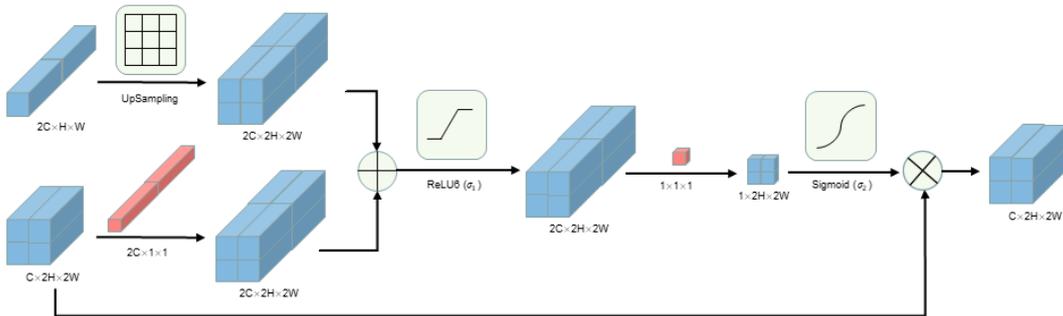


图 3-5 逆注意力门模块的概况

如图3-5所示，IAGs 模块会接收来自解码器  $l + 1$  层解码器上采样输出  $G^l$ （即 MIG-Net 中对应  $l$  层补丁扩展的输入）以及来自编码器  $l$  层的输出特征图  $X^l$ 。Oktay 团队 [75] 的工作是将整体数据降维，将经典的注意力门增强后的输出经重采样后得到注意力系数矩阵。本节的工作反而关注将整体数据作升维，之后配合 ReLU6 激活函数来减少非线性函数导致的信息损失，同时保持 IAGs 模块在低值运算中的鲁棒性。笼统来解释，升维后的数据在 ReLU6 这类非线性激活函数导致的截断域中有更好的容错率，而整个过程中由于对输入主动升维，信息的降维也可以接近“无损”。具体地，本节将注意力系数矩阵  $\alpha_l$  设置为加性注意力，可以计算为公式3-6：

$$\alpha^l = \sigma_2(\psi(\sigma_1(W_x X^l + W_g G^l + b_g)) + b_\psi) \quad (3-6)$$

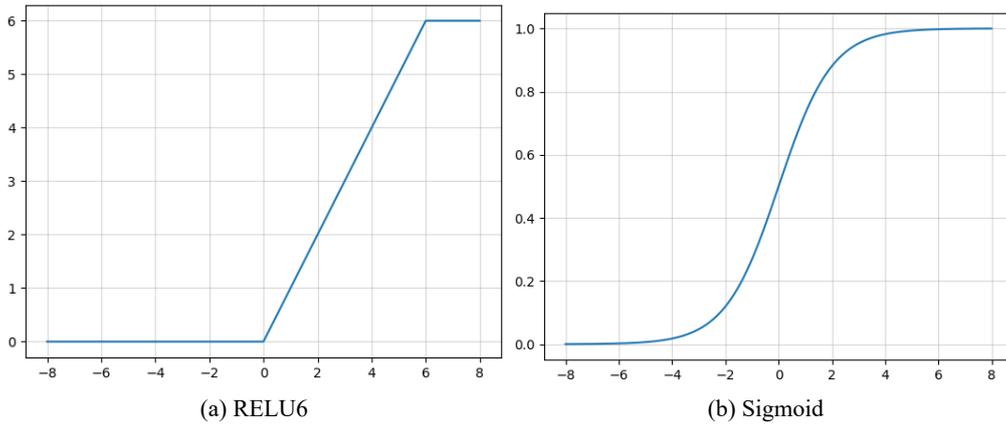


图 3-6 激活函数示意图

其中  $G^l$  是 MIG-Net 中对应  $l$  层的补丁扩展输入，能作为注意力门的门控信号向量，从更大的感受野来指导显著特征。 $\sigma_1$  被选为 ReLU6 函数，如公式3-7所示：

$$\sigma_1(r) = \min(\max(0, r), 6) \quad (3-7)$$

其分布如图3-6-a所示。 $\sigma_2$  设置为 Sigmoid 函数来归一化到 0-1，如公式3-8所示：

$$\sigma_2(r) = \frac{1}{1 + e^{-r}} \quad (3-8)$$

其分布如图3-6-b所示。其中  $W_x$ 、 $W_g$  和  $\psi$  是属于神经元权重线性变换， $b_g$  和  $b_\psi$  则是对应的神经元偏置项。本节中 IAGs 采用  $1 \times 1$  卷积核来实现对特征图  $X^l$  和门控信号向量  $G^l$  进行线性变换。由于 Softmax 激活函数会输出相对稀疏的结果，对于信息保留和训练收敛都是不利的，而 Sigmoid 激活函数在 IAGs 中能提供更好的训练性能。

本节的 IAGs 相关工作，能够根据不同尺度的特征图中包含的图像信息来调节显著特征的激活状态。同时，得益于级联结构，上层的 IAGs 会得到下层的来自不同尺度的特征图激活信号，这对于精准的预测提供了可行性和展望性。

### 3.3 实验结果与讨论

本章研究工作性能的验证过程将在第二章中整理的 XiangyaGMS 数据集上展开，将本章的模型与目前主流的医学分割模型作对比。同时，针对本章模型提出的创新点在消融实验上验证其重要作用。最后就本章模型的分割性能展示其可视化结果，并分析其个别案例下的性能分析。在验证其应用泛化性时，在跨域的 BraTS 公开数据集上评估本章模型的像素级预测结果，增强其作为辅助诊断系统的可行性。

所有实验的运行软件环境是 Ubuntu 20.04 和 Pytorch 1.13，所用的电脑硬件配置主频为 2.3GHz 的 Intel Xeon Platinum 8336C 处理器，GPU 为 NVIDIA 的

GeForce GTX 4090 显卡，在网络推理速度测试阶段使用了 TensorRT 加速。

实验所依赖的数据是 XiangyaGMS 数据集中的 XiangyaGMSBraTS 文件夹，其中的序列影像都统一放缩到与 BraTS 2021 数据集相同的  $240 \times 240$  的分辨率。通过用 “.h5” 格式来保存某一切层影像的字典，并以关键字分别保存影像矩阵以及其病损区域的掩膜矩阵，以方便算法调取。具体的操作中，本章将来自 T1c 和 T2 两个脑部 MRI 序列图像作为独立的输入，其维度表示为  $(240, 240, 1)$  的单通道图像。为了增加模型的泛化性能，对于训练阶段，本章会先将输入维度填补到  $(336, 336, 1)$ ，然后输入执行 [0 度, 90 度, 180 度, 270 度] 的随机旋转以及 10% 随机的高斯噪声来增加数据的复杂性。在每一次输入由 NumPy 数组转 Tensor 之前都会随机截取  $(224, 224, 1)$  维度的子图作为神经网络输入，能有效利用有限标注的数据集。

### 3.3.1 模型评估准则

由于本章的任务为胶质瘤病灶分割，所以在评估准则上会常用 Dice 相似系数 (Dice Similarity Coefficient, DSC)，是一个常用于衡量两个样本相似性的统计量，其定义如公式3-9所示：

$$DSC(P, T) = \frac{2 \times |P \cap T|}{|P| + |T|} \quad (3-9)$$

其中  $P$  代表模型预测的病灶， $T$  代表其真值标签。同样它可以看作是对集合的相似度量，如图3-7，可以很好描述 DSC 指标同 IOU 指标类似的几何意义。其中，蓝色部分为 TP 区域，红色部分为 FN 区域，黄色部分为 FP 区域。根据这样的划分，可以给出 DSC 指标的布尔数据公式，如公式3-10所示

$$DSC = \frac{2 \times TP}{(TP + FN) + (TP + FP)} \quad (3-10)$$

所以容易得出，DSC 指标越高，表明模型对病灶像素的预测相对越准确，越接近于 1，此时预测范围与真实范围完全重叠；而越接近 0，则说明模型的预测性能越弱，此时预测范围与真实范围相互分离。DSC 指标比较关注区域重叠面积，对预测的病灶区域填充比较敏感，而豪斯多夫距离 (Hausdorff, HD) 则可以用来评估病灶分割的边界，所以经常和 DSC 指标一起用于在图像分割任务中。HD 是描述度量空间中的两个子集彼此之间的距离，笼统来说，它是从一组中的一个点到另一组中最近点的所有距离集合中最大距离。如果一个集合的每个点都接近另一集合的某个点，则两个集合在豪斯多夫距离上很接近，其形式定义为：假设有两组集合 A、B，则这两个点集合之间的 HD 定义为式3-11：

$$H(A, B) = \max(h(A, B), h(B, A))$$

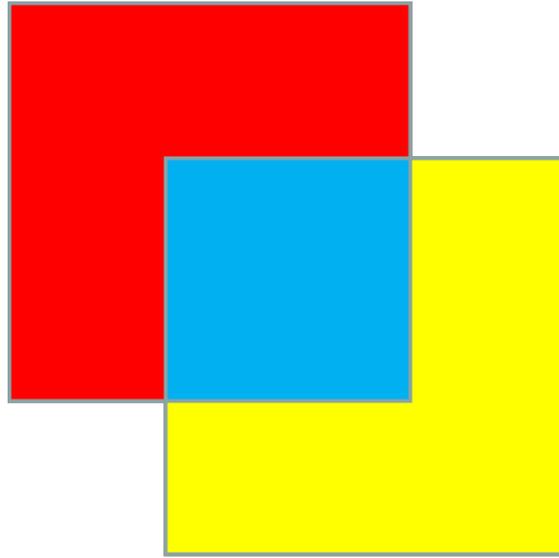


图 3-7 DSC 指标的几何意义

$$h(A, B) = \max_{a \in A} (\min_{b \in B} \|a - b\|) \quad (3-11)$$

$$h(A, B) = \max_{a \in A} (\min_{b \in B} \|a - b\|)$$

其中集合 A 和集合 B 之间的距离范式通常选用 L2 距离。集合间的 HD 指标是两个单向 HD 指标之间的较大值，这种双向的构成能够度量两个集合间的最大不匹配程度。

### 3.3.2 与主流模型的实验结果对比

在评估本章提出的 MIG-Net 模型在胶质瘤病灶分割任务上的性能时，选用三种比较主流的分割模型作对比实验。为了在完整的 XiangyaGMS 数据集上评估模型的泛化性能，实验采用了五折交叉验证的方法，表3-1展示了其验证的平均 DSC 指标和平均 HD 指标。

表 3-1 与主流分割模型性能的平均 DSC(%) 指标和 HD 指标对比

模型	DSC(↑)	HD(↓)
U-Net	83.16	36.70
Trans-UNet	84.84	28.69
Swin-UNet	86.92	21.54
<b>MIG-Net</b>	<b>90.56</b>	<b>15.34</b>

从表3-1可以对比得出，本章的胶质瘤病灶分割模型 MIG-Net，在 DSC 指标和 HD 指标上都要优于主流的经典网络。最重要的是，MIG-Net 的建模过程中存在医学先验知识的干预，能为其出色性能的带来部分解释性。

### 3.3.3 消融实验

本章设计的用于胶质瘤病灶分割的 MIG-Net 模型，借助其模型整体的有效性，能够在性能上超过主流的分割模型。而本节将通过消融实验来验证其中各个创新点设计的有效性，包括：① FPG 损失；② 双重信息引导策略；③ IAGs 模块在内，消融实验的实验环境与对比实验中保持一致，旨在彻底探索各模块对模型性能的关键作用。消融实验的结果如表3-2所示：

表 3-2 MIG-Net 模型的消融实验结果

Exp.	FPG 损失	双重信息引导策略	IAGs	DSC
1	✗	✗	✗	86.92
2	✗	✗	✓	88.13
3	✗	✓	✗	88.85
4	✗	✓	✓	89.74
5	✓	✗	✗	87.52
6	✓	✗	✓	88.76
7	✓	✓	✗	89.35
8	✓	✓	✓	<b>90.56</b>

由表3-2可知，实验 8（MIG-Net）中同时应用 FPG 损失，双重信息引导策略和 IAGs，相对于实验 1（Swin-Unet）能在 DSC 指标上能提升 3.6% 以上，其中三个消融对象都能在各自己的实验对比中验证其有效性，实验 1 和 5（2 和 6、3 和 7、4 和 8）的对比中能反映应用 FPG 损失模型的性能优势；根据实验 1 和 3（2 和 4、5 和 7、6 和 8）容易得出，应用双重信息引导策略能为模型带来较大的性能提升；同样，实验 1 和 5（2 和 6、3 和 7、4 和 8）的对比结果也能验证 IAGs 对模型整体性能提升的有效性。

### 3.3.4 分析与讨论

如图3-8给出 MIG-Net 模型与主流分割模型对部分患者 MRI 影像的预测可视化，每一行分别是 MRI 影像的 T1、T1c 及 T2 序列。如可视化结果可知，深度学习模型的预测结果与病灶真值之间，两者整体重合度较高。本章提出的 MIG-Net 模型相较于其他的主流分割模型，能够对形状大小等方面各异的脑胶质瘤病灶有一个更接近预期真值的分割预测。若将其鲁棒性能用于处理原始 MRI 影像，有望作为精准预测 Ki-67 指数研究诊断的重要预处理手段。然而，MIG-Net 会把部分 0 类区域识别为 1 类区域，会发现其假阳性区域是肿瘤的浸润组织，其常会伴随肿瘤病灶而出现，属于病灶的过渡区域，因此有比较强的迷惑性，甚至在有些

研究文献中会将其标记为 1 类。

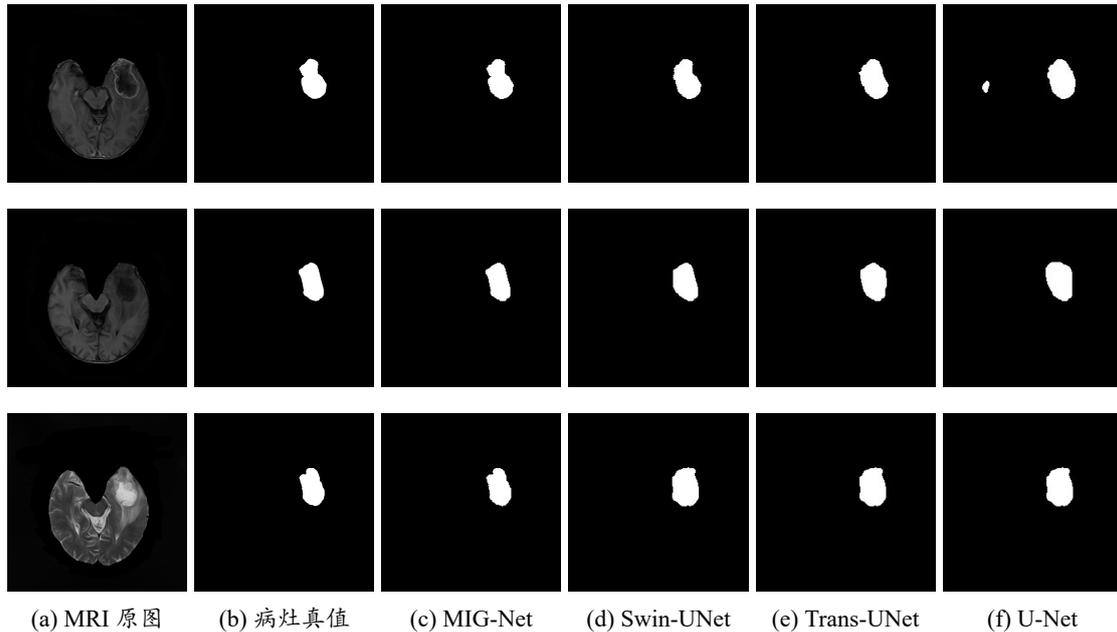


图 3-8 部分患者 MRI 影像的预测与真值的可视化。每行分别对应 T1c、T1 和 T2 序列

真实的临床医学影像的成像质量会受到采集过程中的人员，设备，环境的因素影响，这就对于用于固有数据集训练得到模型的泛化性能提出比较大的考验。在用于辅助诊断系统的实际应用之前，需要验证本章提出的 MIG-Net 模型在多中心医学影像数据集上的泛化性能。本节在 BraTS 2021 的病灶分割数据集上验证 MIG-Net 模型的性能。BraTS 数据集是在标准临床条件下从多个不同中心采集的脑肿瘤 MRI 影像数据集，不同中心由于使用不同的设备和成像协议，从而 BraTS 数据集中包含有不同临床操作导致的极不均匀的成像质量。

为了在模型层面上体现这种不同中心之间的成像质量差异，本节在不同的数据集间进行了交叉测试（在其中一个数据上训练，在其他数据集上测试），从而验证其泛化性能。表3-3给出了 MIG-Net 模型在多中心 MRI 影像上的泛化性能。

表 3-3 XiangyaGMS 数据集与 BraTS 数据集交叉测试下的 DSC 指标

		测试集	
		XiangyaGMS	BraTS
训练集	XiangyaGMS	90.56	86.23
	BraTS	82.64	93.73
	XiangyaGMS+BraTS	91.20	92.08

从表3-3的结果可以看出，病灶分割任务在同一数据集上训练和测试，反映的预测效果都比较出色，而当数据集之间有交互时，预测性能会有所下降。这部

分的性能下滑是由于不同中心临床操作和区域性引入的数据间差异性。为了缓解这部分的差异性带给模型的泛化性能降低状况，本节尝试在联合数据集上训练，即混合 XiangyaGMS 和 BraTS 数据集，从结果可以看出，要胜过于数据集间交叉训练的验证结果。但是来自多中心的类内差异增大，会导致整体预测性能会相较于同一数据集上训练的验证结果要差一些，在临床应用时需要就适用场合合作一个权衡。

### 3.4 本章小结

在本章中，为了避免临床上胶质瘤 Ki-67 指数预测的应用受到人工参与的限制，本章探索了一个多任务信息引导的网络模型 MIG-Net 来实现胶质瘤的恶性程度分级和病灶分割。同时，本章在 MIG-Net 模型中结合胶质瘤病灶先验发生位置来提供医学信息引导与赋予模型决策的部分解释能力。MIG-Net 能够基于胶质瘤的恶性程度分级任务，为目标任务病灶分割来提供包括粗糙定位和语义信息在内的双重先验知识指导策略，以增强任务相关性并促进信息流动。最后，对于 MIG-Net 模型中的解码器结构，本章采用一种变种的注意力门 IAGs 结构来增强上采样阶段的重要信息增益。本章设计的胶质瘤病灶分割模型在 XiangyaGMS 数据集上，取得了超过目前主流分割模型的出色性能，并结合 BraTS 数据集验证了 MIG-Net 模型良好的泛化性能，有望成为 Ki-67 指数预测模型的重要预处理手段。

## 第 4 章 基于多序列融合的胶质瘤 Ki-67 指数预测

高级别胶质瘤患者的中位生存时间不到 15 个月，而从其在各种治疗手段的预后来看，患者的生存率仍然很低。据现有文献调研及基因组分析鉴定结果来看，几种重要的分子生物标志物(如 IDH 突变状态、Ki-67 指数等)有望在预后关系中起到重要作用。但是这类分子标记物的分析都依赖于高精密且复杂的基因测序。这种侵入式的诊断手段，对于患者本身是有损伤的，同时其操作成本对于医疗资源不均的中国国情来说，是非常不友好的。目前医学影像领域中缺少 Ki-67 指数预测相关的研究工作，但实际中 Ki-67 指数水平具有良好的预后价值。因此本文将选用 Ki-67 指数来作为研究对象，利用深度学习的方法，以一种非侵入式的手段来获取到其相对准确的预测结果，在一定程度上帮助临床的决策。

MRI 的成像过程中，脑部组织可以用两个不同的弛豫时间来表征 T1 和 T2，T1 衡量的是自旋质子与外部磁场重新对齐所需的时间，而 T2 是自旋质子在垂直于主场旋转的原子核之间失去相位相干性所用时间的量度。笼统来说，T1 是纵向弛豫时间，T2 则属于横向弛豫时间。在这些序列图像中，图像的对比度和亮度主要由组织的 T1 或 T2 特性决定，这种差异化的组织表现对于影像中信息反馈是非常有利的。

基于 MRI 影像天然带有多序列的构成，临床诊断的习惯是会参考多个序列的信息来做出最终的决策。而对于本章开展的 Ki-67 指数水平的预测模型研究，也需要更深入探讨医学影像中的多序列图像融合的方法。在整章的工作中，本章整理基础结论并提出一种新颖的依赖多序列的神经网络预测模型，以无创的方式来预测得到 Ki-67 指数的高低水平。同时，根据医学任务特性设计出一种新的二元多尺度损失函数，希望从输入相似度方面给与预测模型另一维度的监督。

### 4.1 多模态融合方式探索

在临床诊断中，专业医生常凭借其经验来参考医学影像，主观地给出诊断决策。而怎样参考，参考比重等等，实际上，根本没有办法能将这种主观能动的部分精准地映射给深度学习模型。本节就医学 MRI 多序列影像数据展开相关研究，旨在寻到一种适合医学具体任务的输入组合方式，这对于构建 Ki-67 指数预测模型至关重要。一般地，在目前研究中，备受关注的多模态任务能够为本文研究的多序列 MRI 影像预测提供极大的参考价值。

### 4.1.1 多模态模型

MRI 影像中除胶质瘤病灶外，还包含大量的颅内组织。医学肿瘤领域中有大量的文献表明，用肿瘤病灶而不是整张 MRI 影像的方式，预计让模型训练过程得更加稳定，获得更鲁棒的预测性能。因此在探索多模态融合之前，本节首先需要给出医学 MRI 序列输入组合的方式，包括原始图像与其中胶质瘤病灶构成的输入组合。命名如表4-1所示。

表 4-1 输入组合的不同构造

	T1c/T2	
	原始 MRI 图像	只保留病灶
IMG	✓	—
ROI	—	✓
MIX	✓	✓

多模态模型中使用文本、图像、语音等信息提取特征，改善模型准确度，用于自然语言处理、图像识别、语音识别等领域。多模态模型可将不同模态特征结合，提升模型性能及准确度，从而实现更好的结果。现有研究中（侧重模式识别领域）将模态间不同的组合方式，将用于多模态数据组合方法分为早期融合，晚期融合和中期融合三种。基于 XiangyaGMS 数据集的三种融合方法模型结构图如4-1所示。

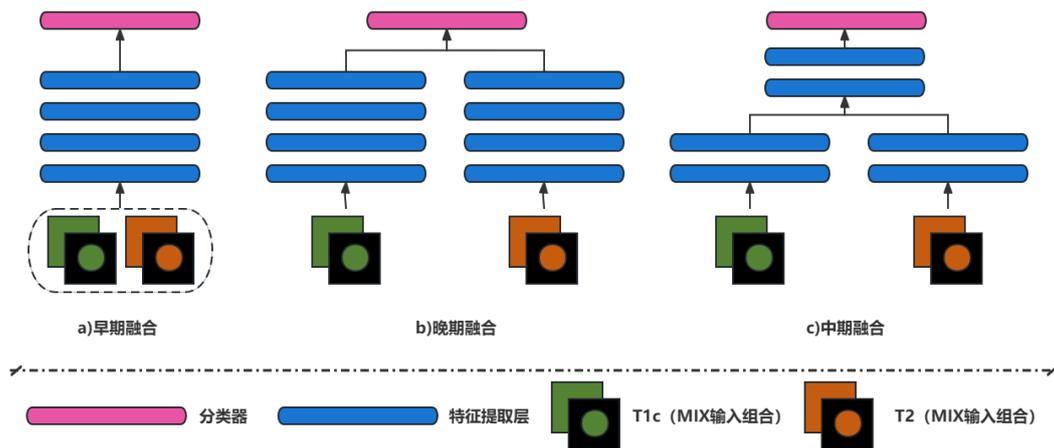


图 4-1 不同融合方式的多模态模型

比起其他方式，早期融合 [76] 方法可以更大程度保留和利用各种模态图像的初级特征，这对于关注图像本身初级特征的具体任务能很好地适用。但是对于本章旨在探索的 Ki-67 指数预测任务来说，初级特征会存在大量冗余信息，早期融合对强相关特征的挖掘没有明显的优势。

后期融合一般是用独立的分支分别训练不同模态的图像输入,并对各分支最终输出的特征图进行决策融合的方法。区别于早期融合是需要保留初级特征之间的共性,后期融合旨在获取到来自不同模态输入分支相对独立的抽象特征。比起利用初级特征之间的联系,更强调控制总损失以实现预测和决策的优化,这样可以尽可能均衡单个分类器的误差。在本节中,在设计实验模型的输入时,把对应的序列图像及其病灶掩膜读取成灰度图,并在通道维度上作堆叠操作。

中期融合,可以根据具体的任务需要有选择性地定位不同模态的融合阶段,能为模型的特征捕捉带来更高效的针对性设计。即,将图像输入执行卷积操作后得到对应的高维特征向量,以一种给定的方式在指定的模型阶段组合在一起。这些中期融合的方法,其多模态模型可以更容易捕捉和利用到不同模态数据之间的共同特征。因为对于大部分研究工作而言,想要达到预期的模型性能,绝不是关注数据的全部特征或仅仅其抽象特征可以做到的。而中期融合的可选择性给予了预期性能更大的容错性,更易于获得不同模态之间的共性。而这种共性,正是多模态模型所寻觅的能表现具体研究性能相关的特征表达。本节中对经典的中期融合模型建模,其中融合操作采用了通道维度的堆叠操作,旨在初步探索其特征之间的共性,验证融合方式的有效性。

本节基于多模态融合,确保模型的主干网络能够准确学习 MRI 图像中反映同一目标的不同层次的特征信息。ResNet 无疑是最受关注的 CNN 模型之一。

#### 4.1.2 实验细节及评价指标

T1c 和 T2 序列图像能在很大程度上反映脑部的真实组织状况,对胶质瘤的 Ki-67 指数预测诊断具有重要指导意义。本节实验的研究目的是基于 MRI 影像来实现无创预测胶质瘤中的 Ki-67 指数水平,术前得到 Ki-67 指数水平有利于对患者的后续治疗。本节实验选用的数据集是第二章整理得到的 XiangyaGMS 数据集,其中的数据均来自于真实就诊与湘雅医院的患者,包含有患者对应的 T1c 和 T2 两个序列图像,同时附带有病灶标注和 Ki-67 指数水平的基因测序结果。

为更客观地反映融合模型的性能,本节针对 MRI 影像中的胶质瘤 Ki-67 指数水平预测作了交叉验证,并取其平均性能来进行相关的统计比较。在本节实验中,考虑到胶质瘤 Ki-67 指数水平预测这一研究任务与日常的模式识别背景有很大的不同,并没有使用 ImageNet 预训练的权重来初始化本节的模型。本节的模型均采用 Kaiming 初始化 [77] 来适配于 ReLU 激活函数,实验统一在两块 NVIDIA GeForce 4090 GPU 上。在训练中,批大小设置为 16,初始学习率为  $4e-5$ 。

本节首先介绍几种常见的用于评价医学鉴定方法性能的指标。

(1) 灵敏度 (Sensitivity)。灵敏度,与召回率等价,它主要用于衡量的是,模

型对于阳性（实际患病）患者的预测情况。其值等于，正确预测的阳性患者数在所有阳性患者数中所占的比例，公式4-1所示：

$$Sensitivity = \frac{TP}{TP + FN} \quad (4-1)$$

高灵敏度测试的阴性结果可用于“排除”疾病，因为它很少误诊确实患有该疾病的人。然而，高灵敏度测试的阳性结果不一定对“诊断”疾病有用。

(2) 特异度 (Specificity)。特异度，衡量的是模型对于阴性（实际未患病）患者的预测情况。其值等于，正确预测的阴性患者数在所有阴性患者数中所占的比例，公式4-2所示：

$$Specificity = \frac{TN}{TN + FP} \quad (4-2)$$

具有高特异性的测试中的阳性结果可用于“诊断”疾病。值得说明的是，对于所有测试，包括诊断和筛查，通常在敏感性和特异性之间进行权衡，因此较高的敏感性将意味着较低的特异性，反之亦然。

(3) 曲线下面积 (Area Under the Curve, AUC)。可以通过设定多个阈值 T，得到一组敏感度和特异度的值。利用敏感度作为纵坐标，(1-Specificity) 作为横坐标，可以画出受试者工作特征 (Receiver Operating Characteristic, ROC) 曲线。在比较不同的分类模型时，通常会绘制每个模型的 ROC 曲线，并以 AUC 值作为模型优劣的评价指标。AUC 面积越大，表示算法的性能越好。

#### 4.1.3 实验结果及分析

本节在 ResNet-50 骨干网络上，基于 MRI 的单一序列，针对不同的输入组合进行了 6 组实验来预测 Ki-67 指数水平。对比实验的结果如表4-2所示。

表 4-2 不同输入组合在 ResNet-50 网络上的预测性能

输入	准确度	灵敏度	特异度	AUC
IMG	0.616	0.588	0.505	0.584
ROI	0.657	0.627	0.546	0.632
<b>MIX</b>	<b>0.678</b>	<b>0.630</b>	<b>0.637</b>	<b>0.679</b>

实验表明，病灶标注将为模型性能带来显著帮助。此外，与 IMG 相比，MIX 的输入提高了约 6-7% 的准确度，MIX 输入组合能更好地适用于本文的研究任务。故在来自 XiangyaGMS 的 T1c 和 T2 的多序列数据支撑下，本节将 ResNet-50 作为骨干模型，采用 MIX 输入组合来验证不同融合方式的模型在 Ki-67 指数预测上的性能。实验结果如表4-3所示。

表 4-3 不同融合方式的 ResNet-50 预测性能对比

融合方式	准确度	灵敏度	特异度	AUC
早期融合	0.682	0.646	0.641	0.695
晚期融合	0.666	0.635	0.635	0.682
中期融合	<b>0.691</b>	<b>0.693</b>	<b>0.689</b>	<b>0.724</b>

从表4-3可以看出，早期融合预测性能会优于晚期融合，中期特征可能包含有任务强相关的信息，仅凭借后期抽象特征可能很难学习到不同模态之间的关联信息。中期融合优于早期融合，说明初级特征中的冗余信息误导了预测模型的学习，相比晚期融合能够提高了 2-3% 的准确度。

实验结果来看，基于 MRI 的放射组学特征可以无创预测突变状态。本节对 IMG、ROI 和 MIX 输入的网络输出进行了 Grad-CAM[78] 分析。通过基于梯度的定位，可以获得关于注意力的粗糙视觉解释。图4-2给出了部分患者的 Grad-CAM 可视化结果。

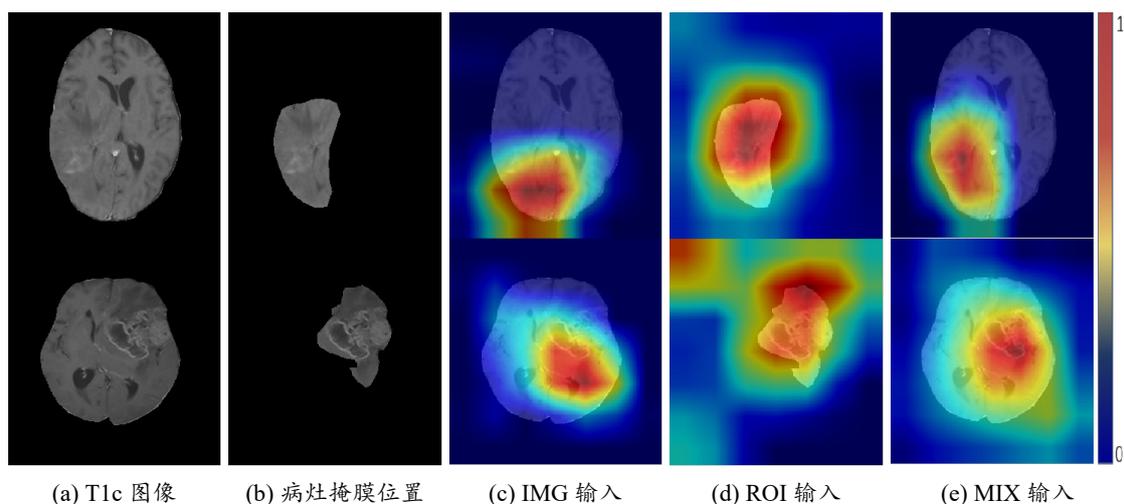


图 4-2 ResNet 上针对不同输入组合的可视化结果。注：从蓝色到红色表示重要性的增加

IMG 输入中的注意分布有时甚至无法覆盖肿瘤区域，这也与它的准确度较差相对应。而观察到在 ROI 输入中，模型会比较专注于病灶的局部边缘，而不是内部的核心区域。ROI 模型的性能不如 MIX 模型，本文推测 Ki-67 指数水平的预测更可能受到胶质瘤的病灶核心区域的影响。

事实上，多模态模型的性能取决于不同模态间的像素级匹配，本节采用的物理匹配或 Spm12[79] 等半自动算法来校准。然而，像素级的整体匹配仍然难以实现，为了解决这个问题，本章设计了不同融合阶段的多模态模型。早期融合作为主流方法之一，其模型性能非常依赖于像素级配准。晚期融合方法侧重于高级特

征,几乎不依赖像素级匹配,甚至可以融合文本等异构信息特征。而在本节的实验中,晚期融合的准确度却反而低于早期融合,这反映了 Ki-67 指数水平预测任务中,多序列间的特征相关性对于模型性能的影响更大。对于中期融合,单个模态的主要特征被编码到特征向量中。既能避免可能存在的模态不完全匹配所带来的干扰,又能充分利用多序列之间的特征相关性,从而获得较好的实验性能。

值得注意的是,如果不按患者划分训练集和验证集,这样构建的预测模型可能存在数据泄露的风险。来自同一患者的相邻切片图像通常共享很多信息,这可能是一些公开研究中相似方法获得不同结果的根本原因。在本节的研究中,若不考虑这种泄露情况,模型可以得到超越 90% 的准确度性能。但是为获得稳健的预测结果,本节在实验中十分严密地注意根据患者划分数据集,并使用交叉验证来测试其性能。

## 4.2 基于多序列融合的 Ki-67 指数预测模型

在本章之前的工作中,中期融合的多模态模型可以更好地适用于医学胶质瘤 Ki-67 指数预测任务。易得到一个简单的结论,胶质瘤 Ki-67 指数预测任务需要的不是单纯的初级特征共性或抽象特征共性,中间态特征之间共性特征的捕捉是一个复杂的过程。在上一节中采用的中期融合模型,在对比实验中取得最优的性能,而其中的融合方式只是一种初步的探索。在本节中,希望找到一种更优的融合方式来提升模型对于胶质瘤中 Ki-67 指数的预测性能。

### 4.2.1 多尺度多序列融合模型

多模态方法已成为基于医学图像进行相关性研究的关键技术。在前期工作中,中期融合模型能更好地适用于本文的任务,然而,至今为止,学术界仍未就具体的中期融合方式达成共识。比较主流的做法是采用通道拼接、向量元素积等方法为融合载体,将来自不同模态的信息特征映射到同一个空间维度 [80]。如此,该多模态模型的训练过程可视为学习一个联合表示。与此不同,协调表示具有更多的互补性,可以最大限度地挖掘不同的多模态数据信息表示的相关性。CentralNet 同时借鉴了两种表示形式,利用现有 CNN 独立处理单一模态的模型。该模型引入了一个中心网络来连接不同模态的特定网络,从而将不同模态的表示映射到同一个公共空间。

同时最近,Zhang 团队 [81] 提出了一个更全面的性能模型 ResNeSt。该模型集成了 ResNet 块的分组卷积和 ResNeXt [82],可以减少模型参数的数量,获得更大的接受域,使训练模型更适合特定的研究任务。此外,ResNeSt 还在包卷积中引入了 SENet [72] 和 SKNet [83] 的注意结构,增强了相关特征的表达。

本节受 CentralNet 的启发，提出 KiNet 模型用于预测 Ki-67 指数。该模型的构建涉及到一个主分支和两个独立的辅助分支，每个分支都包含四个特征提取阶段。KiNet 中的特征提取模块的结构类似于 ResNeSt 块，以深度残差学习方法提高模型的准确度，如图4-3所示，其中  $(h, w, c)$  分别为输入的高度、宽度、通道。

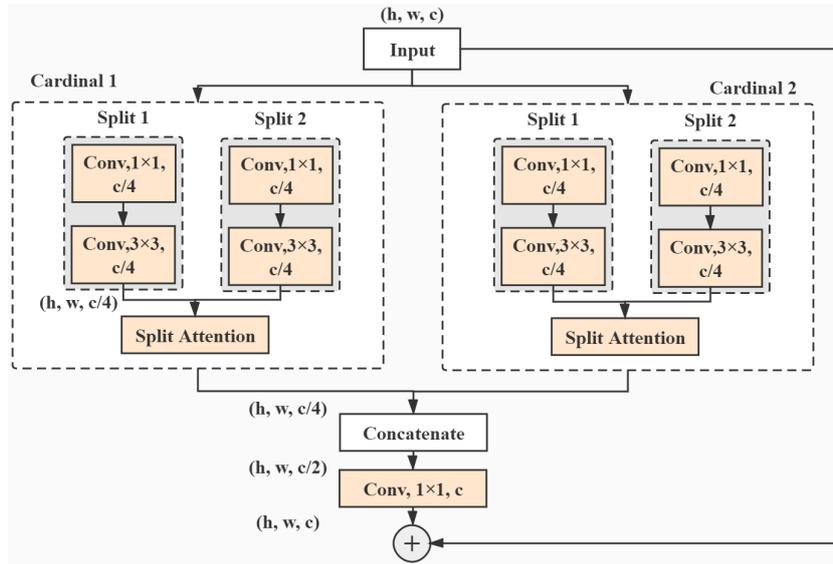


图 4-3 KiNet 块的架构

如下图4-4所示是本章设计的包含多尺度损失的 Ki-67 指数预测模型的整体框架。该模型中，输入独立分支网络的 T1c-IN 和 T2-IN 是 T1c 和 T2 两个序列各自的 MIX 输入，输入主干分支的 Merge-IN 则是两个序列 MIX 输入组合的通道拼接结果。独立分支网络中经不同阶段的 KiNet 结构块，能够提取到不同尺度的放射组学特征图。为捕捉不同序列间的异质相关性特性，主干分支拼接来自独立分支网络对应尺度的特征图，以映射多序列信息特征到同一个空间维度。主干分支跨阶段的放射组学特征输出经分类器（全局平均池化操作和全连接层组合）得到多尺度决策。最后，将图像的多尺度特征向量与患者元数据拼接送入 LightGBM[84] 分类器中，得到最终的预测结果。

#### 4.2.2 二元多尺度损失函数

本节构建了旨在实现 Ki-67 指标分类的多尺度多序列网络，包括单一序列分支和主要分支上多尺度的预测任务。本小节着重介绍了 KiNet 模型训练过程中的损失函数。

KiNet 模型的总损失函数  $L$  定义，如式4-3所示：

$$L = \alpha_1 \times L_{T1c} + \alpha_2 \times L_{T2} + \alpha_3 \times L_{master} \quad (4-3)$$

其中  $L_{si}$  为主分支中对应的第  $i$  级尺度下提取的特征图预测损失， $L_{cat}$  为主分支

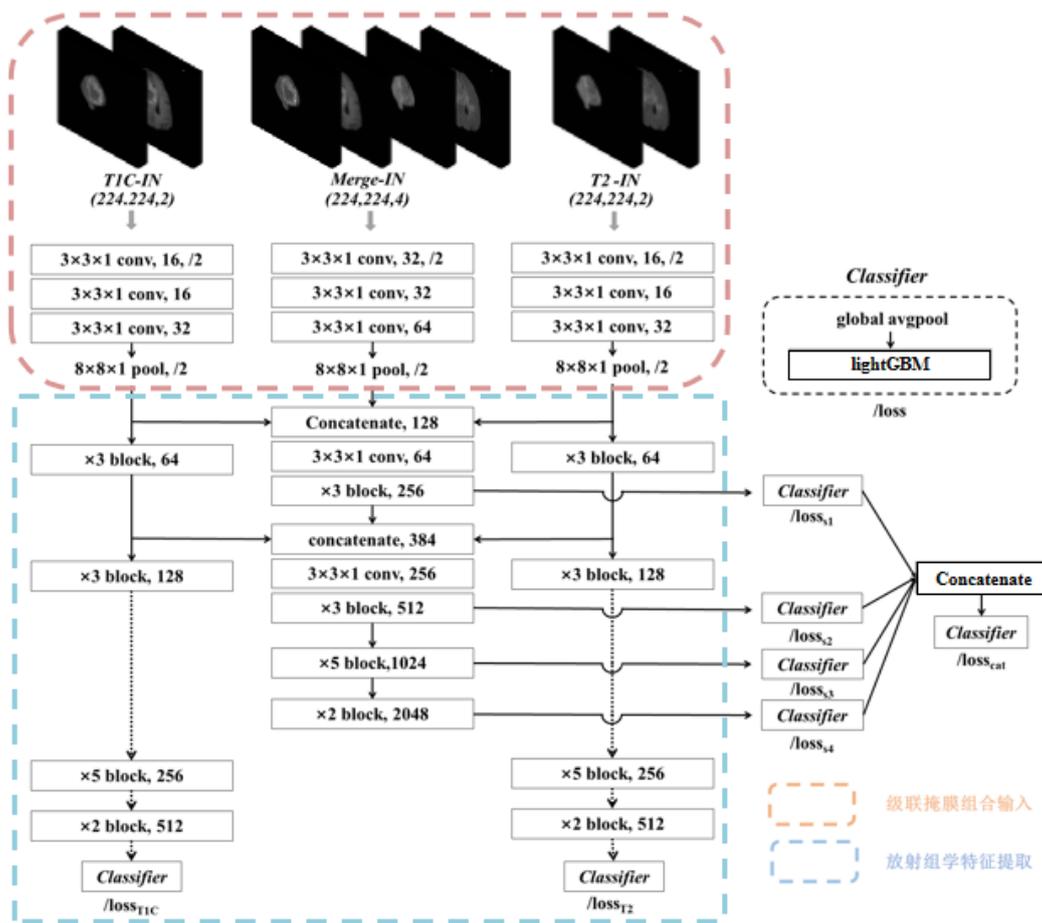


图 4-4 KiNet 的总体框架（在测试阶段不考虑虚线标记）

中融合不同尺度特征图后的预测损失。

本章选择将 Ki-67 指标预测任务视为高低水平识别的二分类任务，因此在独立分支  $L_{T1c}$ ,  $L_{T2}$  和  $L_{cat}$  上都采用交叉熵损失函数，如式2-1所示。而关于 KiNet 的主支路损失  $L_{master}$  的定义，如式4-4所示：

$$L_{master} = \left( \sum_{i=1}^4 \beta_i \times L_{s_i} \right) + \beta_5 \times L_{cat} \quad (4-4)$$

本节的实验设计中，会采用不同的训练策略来寻找符合模型最优性能的  $\alpha_{i=1,2,3,4}$  和  $\beta_{j=1,2,3,4,5}$  参数值。

不同层次的特征可以给模型的性能带来多尺度的监督，因此预测效果更加稳健。本节在现有模型上部署了多尺度损失监督，其中多尺度特征向量通过全局平均池化和全连接层连接。在探索多维特征融合在其他领域的思想时，本节注意到孪生网络 [85] 中相似度监督在语义场中的有效性。与语义语法中固定的已知结构差异不同，MRI 多序列图像更容易有效地监督类内相似性而不是类间差异。如

公式4-5所示。

$$L_{simaese} = \sum_{i=0}^3 \beta_i \frac{1}{2N} \sum_{n=1}^N \|A_n - B_n\|^2 \quad (4-5)$$

其中  $A_n$  和  $B_n$  是对应第  $N$  个样本的两个网络输出， $N$  是单序列图像的数量。构建的辅助监管损失模型如图4-5所示。

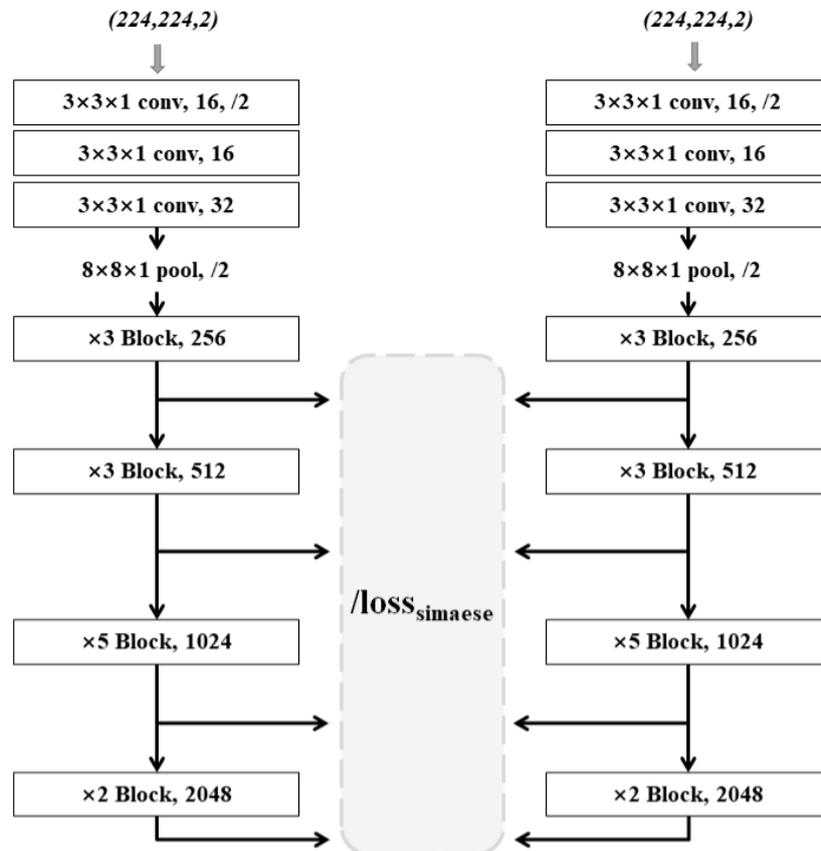


图 4-5 相似度监测提供单序列输入的模型

在此基础上,本节提出了更适合 Ki-67 指数预测的二元多尺度损失函数 BMS-Loss, 如公式4-6所示:

$$L_{BMS} = \gamma_1 \times L + \gamma_2 \times L_{simaese} \quad (4-6)$$

其中  $\gamma_1, \gamma_2$  分别为交叉熵多尺度损失和相似多尺度损失在 BMS-Loss 中的比例。在本实验中,将  $\gamma_1$  和  $\gamma_2$  都设置为 1, 以确保修改后的损耗不会放大原始损耗。

#### 4.2.3 实验细节及评价指标

本章最初认为 Ki-67 指数预测是一个回归任务,但临床获得的 Ki-67 指数本身是一个统计值,所以本章将其总结为一个分类任务。根据 WHO 指南, Ki-67 指数低于 10% 称为低水平 Ki-67 指数,标记为 0,其余称为高水平 Ki-67 指数,

标记为 1[4]。对于  $N$  个患者，其对应的年龄与年龄会在输入分类器前被编码为  $2 \times N$  向量。

在本节实验中，模型均采用 Kaiming 初始化来适配于 ReLU 激活函数，实验统一在两个 NVIDIA GeForce 4090 图形处理器上。在训练中，批大小设置为 16，初始学习率为  $4e-5$ 。为更客观地反映融合模型的性能，本节针对 MRI 影像中的胶质瘤 Ki-67 指数预测分级作了交叉验证，并取其平均性能来进行相关的统计比较。除常规的几个医学指标外，本节中还使用了 Kappa 系数来评价模型的实际临床性能。Kappa 系数的计算是基于混淆矩阵的，取值为 -1 到 1 之间，通常大于 0。其计算公式如 4-7 所示：

$$Kappa = \frac{p_o - p_e}{1 - p_e} \quad (4-7)$$

$$p_o = \frac{\text{对角线元素之和}}{\text{整个矩阵元素之和}}$$

$$p_e = \frac{\sum_i \text{第 } i \text{ 行元素之和} \times \text{第 } i \text{ 列元素之和}}{(\sum \text{矩阵所有元素})^2}$$

实际的分类问题中，不同类别所属的样本数量往往不太平衡。而根据 Kappa 的计算公式，越不平衡的混淆矩阵，Kappa 值就越低。

#### 4.2.4 实验结果与分析

借助 XiangyaGMS 数据集，本章引入了交叉验证方法进行相关实验。表 4-4 给出了在 Ki-67 指标预测任务中，KiNet 与三种主流神经网络模型的性能对比。ResNet-50 和 ResNeSt-50 的输入组合，是由不同序列的 MIX 组合在通道维度上

表 4-4 不同神经网络对 Ki-67 指标的预测性能

模型	准确度	灵敏度	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
<b>KiNet</b>	<b>0.752</b>	<b>0.755</b>	<b>0.749</b>	<b>0.788</b>	<b>0.472</b>

拼接得到，即网络输入为 Merge-IN。从表中数据可以看出，与仅采用不同序列 Merge-IN 这种方法相比，CentralNet 利用主分支来融合两个独立分支的多尺度特征对 Ki-67 指标的预测效果更好。与 CentralNet 不同的是，KiNet 能更有效利用不同序列输入在模型早期阶段的相关性，更有利于表达不同层次特征信息之间的相关性，因此在这项任务中更具优势。图 4-6 为不同神经网络对 Ki-67 指标的 ROC 曲线。

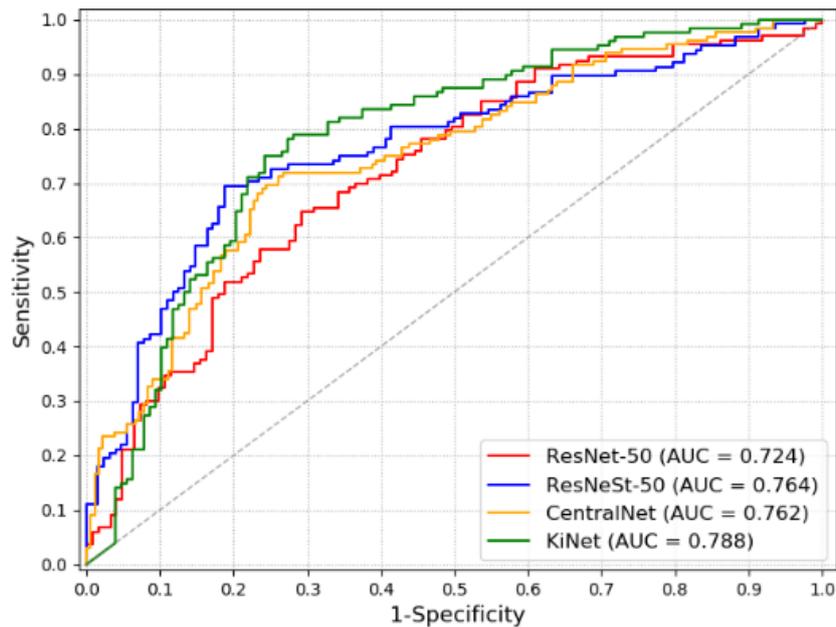


图 4-6 Ki-67 指标不同模型的 ROC 曲线

为了明确 KiNet 模型中性能提升的归因, 希望设置消融实验来验证 KiNet 中的子模块 (BMS-Loss, Merge-IN, Meta-data) 对模型整体性能的影响, BMS-Loss 描述的是主分支中由不同尺度特征图预测结果集成的损失, Merge-IN 则是主分支中的早期特征提取模块设计, 而 Meta-data 表示的是患者临床元信息的向量拼接。消融实验同样采用了交叉验证训练方法, KiNet 消融实验结果见表4-5。

表 4-5 KiNet 消融实验的预测性能

Exp.	BMS-Loss	Merge-IN	Meta-data	准确度	AUC
1	✘	✘	✘	0.735	0.766
2	✘	✘	✓	0.736	0.768
3	✘	✓	✘	0.742	0.773
4	✘	✓	✓	0.743	0.774
5	✓	✘	✘	0.746	0.778
6	✓	✘	✓	0.748	0.781
7	✓	✓	✘	0.750	0.785
<b>8</b>	✓	✓	✓	<b>0.752</b>	<b>0.788</b>

总的来说, 实验 8 在消融实验中能表现出最佳的预测性能, 这进一步证明了本章提出的 KiNet 在网络结构设计上的合理性。还容易得到以下结论:

(1) 通过对比实验 1 和实验 3 (2 和 4、5 和 7、6 和 8) 的结果发现, 添加额外的 Merge-IN 有助于提升 KiNet 模型对 Ki-67 指标的预测性能。Merge-IN 通过直

接组合不同序列的原始图像，可以有效地捕捉不同序列之间的早期融合关联。临床中涉及的人工选择特征，很多也大多是基于医学图像的早期特征。因此凭借这个输入组合，模型能更容易学习不同序列图像的位置、灰度等浅层特征之间的相关性；

(2) 在实验 1 和 2 (3 和 4、5 和 6、7 和 8) 的比较中发现，包括患者性别和年龄在内的元数据带来的性能提升比较小。对比结果表明，在预测胶质瘤患者 Ki-67 指数的具体任务中，直接对患者的性别和年龄进行向量编码及拼接图像深度特征的融合策略是有瑕疵的。此外，医生的部分解释原因包含可能是由于胶质瘤患者人群范围广，性别因素对胶质瘤增殖能力预测关联不大所致；

(3) 与此同时，实验 1 和 5 (2 和 6、3 和 7、4 和 8) 的结果表明，BMS-Loss 是有效的策略。因为不同尺度的特征图能反馈到大量的隐藏图像噪声和特征，同时，类中单序列样本的相似度可以对多序列模型进行弱监测，这些对于提高模型的鲁棒性和特征学习能力非常有益。

本文通过多次探索性实验，最终建立了主分支结构。在 CentralNet 的影响下，首先考虑了带权的元素和运算作为表示融合方法，其中权值为可学习参数。在 Ki-67 指标预测任务中，将元素和权重初始化为均匀分布。在训练的模型上进行参数可视化，其结果如图4-7所示。

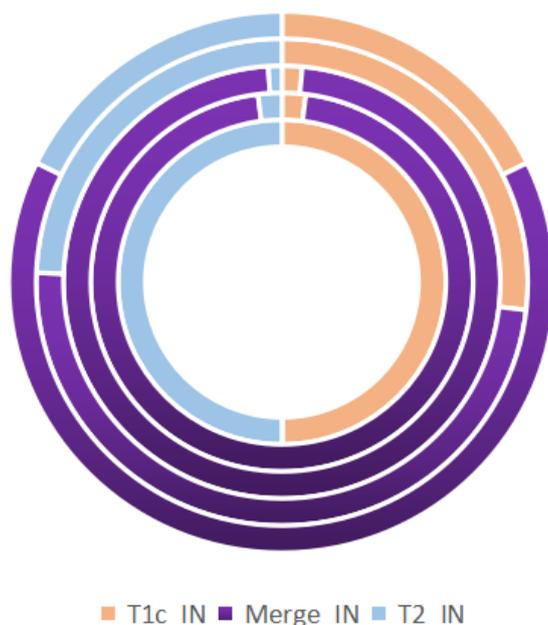


图 4-7 训练完成的融合参数可视化（其中由内到外分别对应主分支的第 1-5 个阶段）

从图4-7可以看出，对于主分支中比较深的两个阶段，可以发现来自独立分支输入的比重很小，因此本章只关联前三个较浅尺度的独立分支输入，以简化本章的 KiNet 模型。区别于直接使用元素和为融合手段，通道拼接可以减少融合后

信息的损失。在主分支的第一个阶段中，是由不同序列在通道维度上拼接得到的，所以序列间的比重自然而然为 1:1；而主分支的第二个阶段中，这里通过图中系数条长度估计得出，序列间的比重接近 1:2:1，同理，第三个阶段中的比重接近 1:4:1。根据这个结论后，本章尝试用向量通道拼接来替代可能会造成信息损失的向量元素相加的操作，对于主分支的二、三阶段按照对应的比重在特征图通道维度上执行向量拼接。

同时对于独立分支与主分支间的损失加权系数，专业医生团队给出建议，T1c 序列在 Ki-67 指数水平预测中也许能发挥更重要的作用，所以本章倾向于为 T1c 分支设定更大的损失权重。经过多次超参数给定及调试，最终将分支间损失权值  $\alpha_{i=1,2,3}$  设定为 0.35、0.4、0.25(T1c-IN 辅助支、Merge-IN 主支、T2-IN 辅助支)；分支内各阶段尺度的加权系数  $\beta_{j=1,2,3,4,5}$  分别设置为 0.025、0.025、0.2、0.35 和 0.4 时，网络性能最优。这个结果与本章将加权系数设置为可训练参数时的偏差不大，这在一定程度上表明深层抽象特征对模型预测结果的影响更大。

### 4.3 患者实例诊断

在上文的工作中，胶质瘤 Ki-67 指数预测系统能够给出单张切层图像的预测，对于医学研究具备一些意义。但是对于医学临床应用角度出发，本文更希望能得到患者级的决策结果，这样的模型结论才能真正具有实际意义。

回溯到 XiangyaGMS 数据集的标注获取途径，本是基于患者的组织来得到患者级的病理诊断结果，而在这之后才将病理诊断结果统一赋予该患者所属的所有切层影像。所以基于 Ki-67 指数预测模型输出的图像级诊断结果，本节会探索一种决策融合方式来得到患者级的预测结果。

比较容易知道的是，肿瘤病灶中是最有可能包含丰富的分子标记物预测相关特征的区域。在做了大量的文献调研以及专业医师团队的沟通之后，本节确定了一下几种具备可行性的方案：

(1) All: 选取所有包含胶质瘤病灶的图像，将他们预测的 Softmax 概率取平均，作为该患者的预测结论；

(2) Near: 选取胶质瘤病灶面积最大的图像及其上下的图像，将三者预测的 Softmax 概率取平均，作为该患者的预测结论；

(3) Large: 选取胶质瘤病灶面积最大的图像，将其预测结果作为该患者的预测结论。

针对提出的三种方案，本节在 XiangyaGMS 测试集上重新评估了其患者级的预测性能，如表4-6所示。容易得出，凭借 KiNet 对 MRI 影像的出色预测性能，融合后的患者级预测结论都能取得更好的性能。其中，Near 类融合得到的患者级

预测结论表现得更好一些,也得益于将病灶相对丰厚图像集合的预测结果进行一定的权衡之后,能获得更鲁棒的性能。此外,本节还给出了不同患者级融合方法的概率校准曲线(probability calibration curve),可以看出, Near 类融合方法的性能具备可靠性。本文的模型倾向于给高概率样本给出超过其实际概率的置信度,而对于低概率样本则赋予低于其实际概率的置信度,这种情况在医学临床中是可以被容忍度的。

表 4-6 KiNet 在 XiangyaGMS 测试集上的患者级预测性能

患者级结果融合方式	准确度	灵敏度	特异度	AUC
—	0.752	0.755	0.749	0.788
All	0.789	0.809	0.778	0.824
<b>Near</b>	<b>0.821</b>	0.819	<b>0.810</b>	<b>0.857</b>
Large	0.813	<b>0.827</b>	0.802	0.840

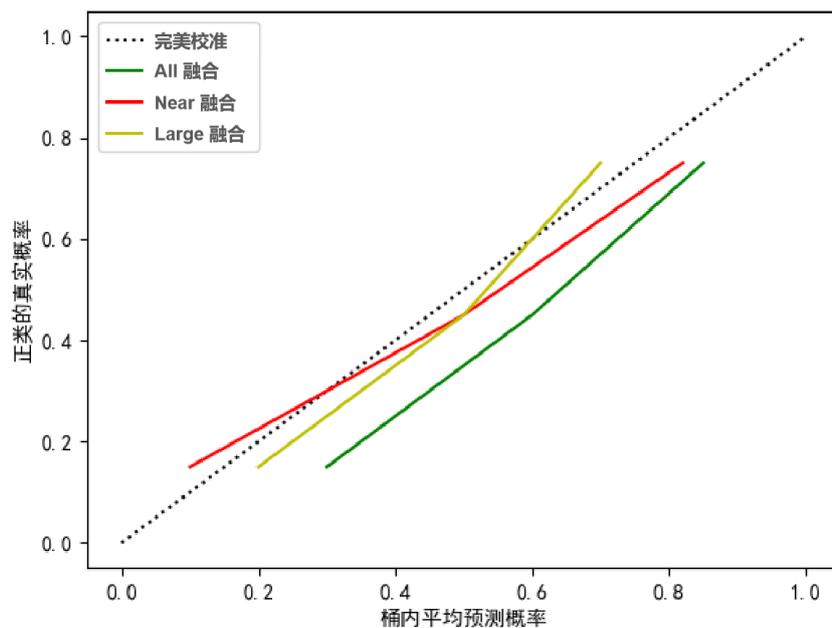


图 4-8 KiNet 基于不同患者级融合方法的概率校准曲线

#### 4.4 本章小结

总的来说,多模态 MRI 放射组学模型可以实现对 Ki-67 指数的稳健预测。肿瘤病灶是影响其分子标记物的重要参考,其中以肿瘤核心区域的视觉分析与预测 Ki-67 指数关联性最大。本文还提出了一种新的基于 MRI 多序列的多尺度神经网络 KiNet,用于无创预测胶质瘤 Ki-67 指数水平。KiNet 中引入两个独立分支用于 MRI 两个序列图像的特征提取,在主分支中集成来自独立分支的不同尺度特征,

将训练完的影响深层特征及患者元信息编码向量拼接，并输入到 LightBGM 中得到最终预测结果。此外，本章提出的一种新的二元多尺度损失函数 BMS-Loss 能为多序列模型提供更多维度的监督，带来性能的稳定增长。最终，本章探索合理的患者级融合方法来得到患者的诊断结论，该网络首次实现了对胶质瘤 Ki-67 指数的非侵入式预测，为医生提供更多的治疗方案选择。

## 第 5 章 级联框架和诊断系统设计

从本文前面的研究不难发现，深度学习能基本实现对胶质瘤的 Ki-67 指数预测效果。从应用角度考虑，在本章中会将胶质瘤病灶分割模型与 Ki-67 指数预测模型作级联设计，这样可以把实际的辅助诊断问题，转化为流程化智能模块的推理，为临床诊断的落地踏出第一步。最后，本章会将本文所设计的模型算法作一个整合，并将整合后的模型算法部署在客户端上作一个应用实现，对该临床诊断系统的落地应用有重要的参考价值。

### 5.1 级联框架设计与实现

关于临床诊断算法落地的展望，本文希望是能够具有比较高的智能化水平，因此在诊断过程中的一些预处理步骤，如肿瘤病损区域的分割及特征提取等，也是一个需要集成的功能之一。着眼于本文的研究工作，主要需要被集成的功能是 Ki-67 指数的诊断结果预测以及胶质瘤病灶的分割，分别是对应本文的第三章和第四章内容。

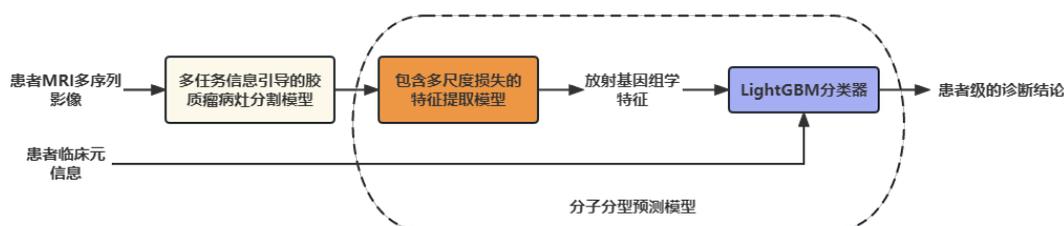


图 5-1 诊断系统级联模型的概况

临床辅助诊断系统主要是由上述的两个模块算法构成，本章将两个模块算法作级联操作，如图4-8所示。级联模型主要包含以下步骤：

(1) 首先，患者的 MRI 影像会被输入到胶质瘤病灶分割模型中，来获取对应的胶质瘤病损区域的掩膜输出。

(2) 然后，将原始 MRI 图像与其经分割模型得到的掩膜组合成满足胶质瘤 Ki-67 指数预测模型的输入格式。

(3) 然后，将合理的输入组合输入到预测模型中来得到具备临床参考价值的放射基因组学特征。

(4) 最后，利用集成分类器从放射基因组学中预测得到具备临床价值的结论。诊断具体的细节见图5-2。其中，本文就胶质瘤病灶分割任务提出了一个多任务信息引导的模型 MIG-Net，能够根据输入的 MRI 影像来给出其对应的掩膜输出。MIG-Net 已在本文第三章中被验证其出色的预测性能，同时具备优秀的泛

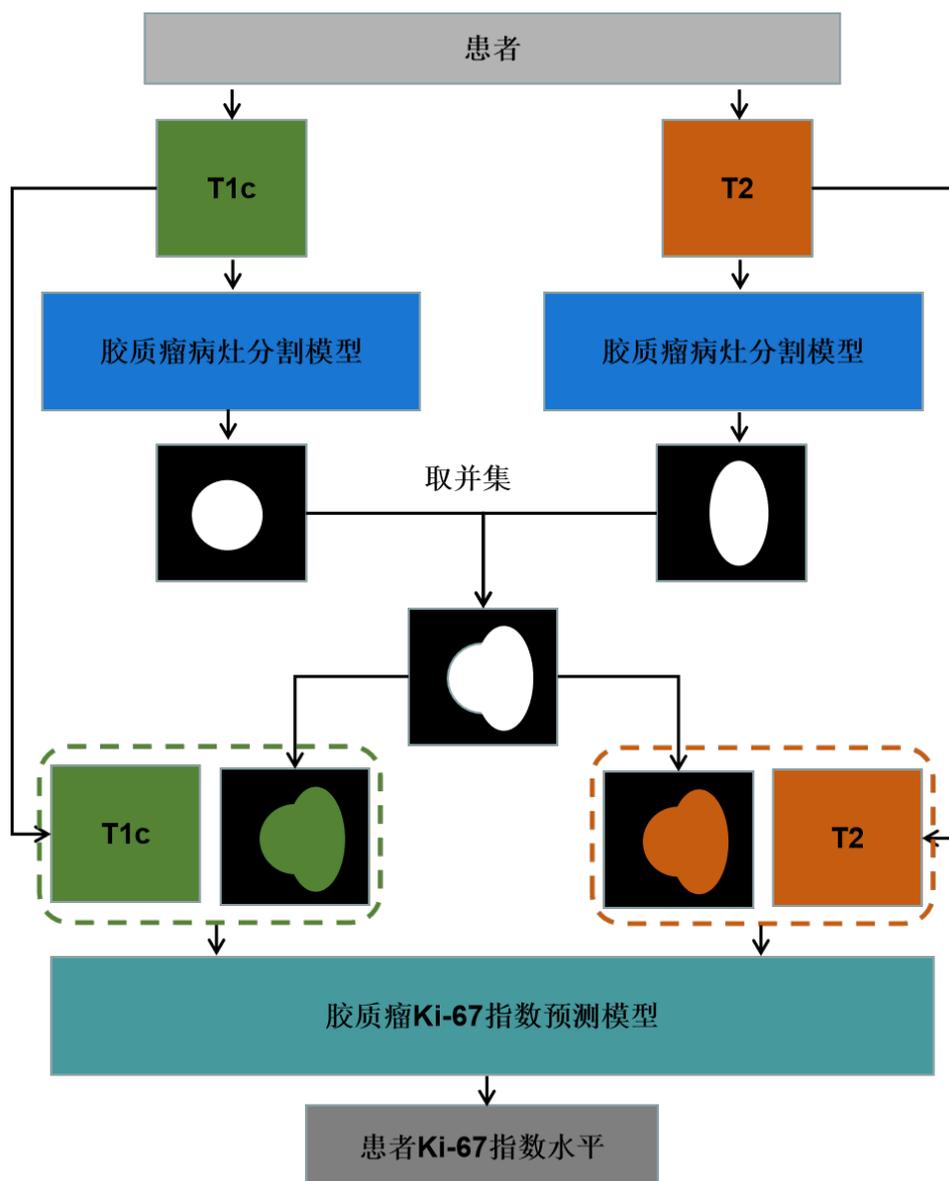


图 5-2 诊断系统细节流程

化性能，能够提供到 Ki-67 指数预测模型一个较为精准的胶质瘤病灶定位，来减少诊断系统应用时的人为干预。而观察到 Ki-67 指数预测模型的输入组合，其设置为来自同个患者的不同序列及其对应的病灶定位输入，这对于病灶分割模型 MIG-Net 的输出来说，尚不能直接适配。下面会给出适配的详细步骤：

(1) 首先，需要对来自同个患者中，不同序列相对应位置切层影像的输出作一个决策融合。因为胶质瘤病灶的标注，是专业医生根据具体的 MRI 影像标注得到的，而其中起重要作用的是 T1c 和 T2 序列。T1c 序列要在操作之前往血液打造影剂（颜料），这样会让血液流动快的地方显示明亮，肿瘤病灶区域也会明显呈现亮块，而 T2 序列能够作为肿瘤组织解剖的重要依据。所以融合的目的不是为了精准得到对应切层的病灶，而是想要将病灶尽可能完成的包含进来，所

以本章采用了将不同序列病灶掩膜的并集，以确保能关注到一整个肿瘤病灶，不至于造成重要信息的遗漏。

(2) 然后，需要把病灶掩膜投影到原图上来得到病灶定位，与原图一起在通道维度上拼接来得到 MIX 输入组合，这样构成的输入可以符合胶质瘤 Ki-67 指数预测模型的要求。这一点在原则上，病灶掩膜被要求接近专业医生的标注精度。但是从工程实用角度，本文的肿瘤病灶分割模型在保证相对高的预测精度前提下，可以被允许将周围的背景区域错误预测为病灶（即假阳率偏高），确保病灶信息的完整性。

## 5.2 诊断客户端设计与实现

应用的落地，涉及到界面设计，数据传输及算法部署等方面，是一个需要较长时间开发的模块。在本章的进度安排中，暂时只实现整个应用的客户端展示初代版本，能在联网主机上使用，满足当前合作项目的交付要求。如图5-3是使用 Qt 4.2 开发实现的辅助诊断系统客户端界面，其中的 Ki-67 指数模块是本文研究工作的产出成果。

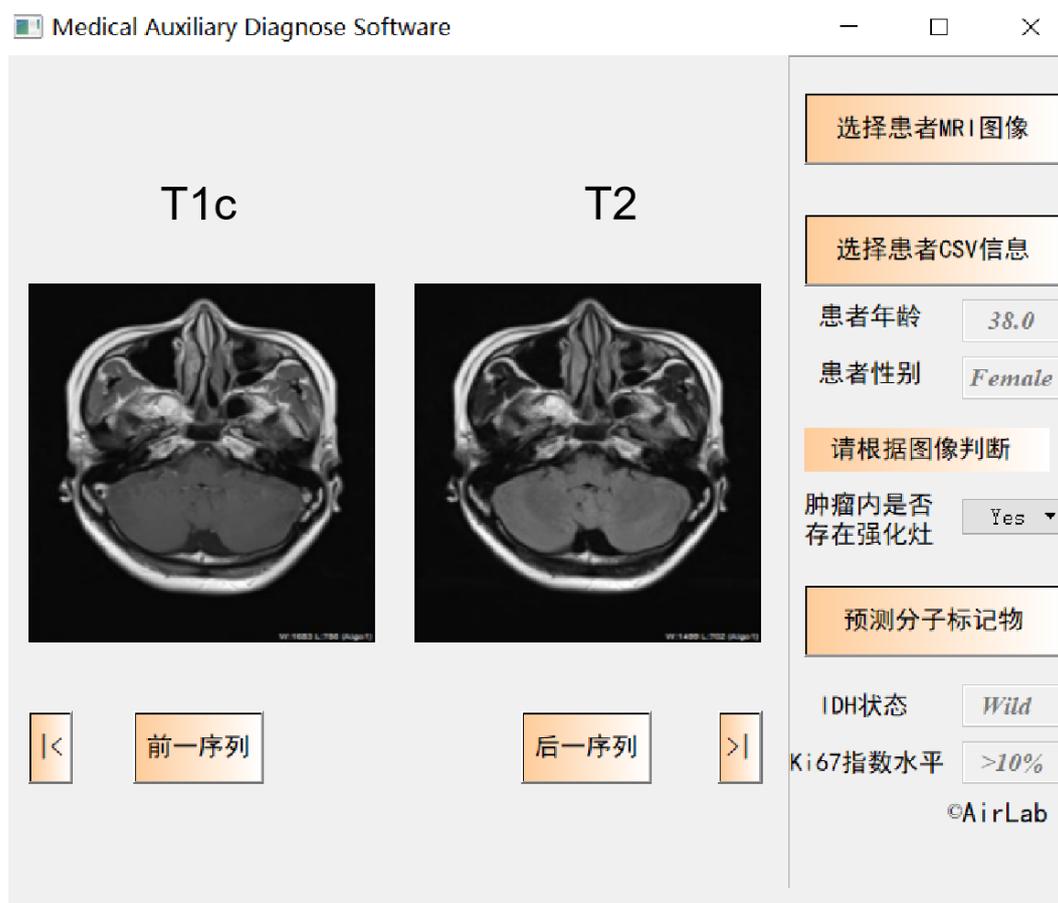


图 5-3 辅助诊断系统的客户端界面

首先，该系统需要一个比较结构化的数据输入，具体的文件格式如图5-4所示。对于临床元信息的统计表，由三列构成：编号，年龄，性别，其中注意性

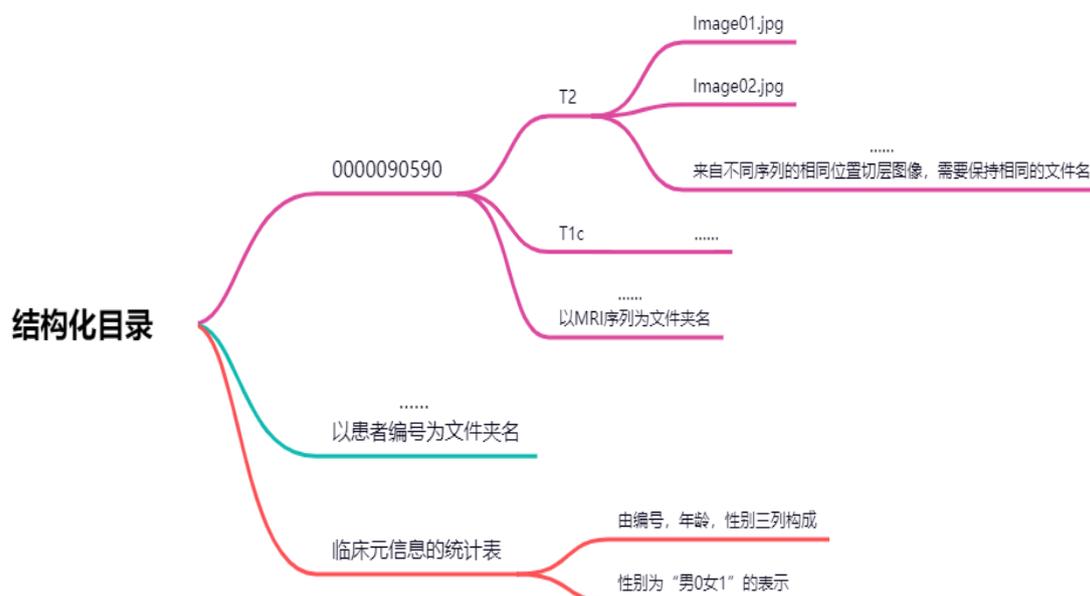


图 5-4 数据输入的结构化目录说明

别为“男 0 女 1”的表示。而对于图像输入，一级目录是以编号为文件夹名的文件夹集合，二级目录是以 MRI 序列为文件夹名的文件夹集合，最后是属于某一 MRI 序列的“.jpg”格式图像，需要注意的是，来自不同序列的相同位置切层图像，需要保持相同的文件名。

对于分子分型的预测，其流程如图5-5所示，需要完成几步操作：

(1) 首先点击“选择患者 MRI 图像”，最终选择的文件夹是对应患者的“T1c”文件夹，T2 序列会被自动识别选中，可以通过”箭头“按钮实现前后 MRI 序列的看片翻阅，能够模拟出临床评估时的实际需要；

(2) 首先点击“选择患者 csv 信息”，选择包含对应患者临床元信息的统计表，模型会自动提取到属于对于患者的年龄与性别，用于输入到预测模型中；

(3) 然后可以根据看片翻阅的结果，判断肿瘤内是否存在强化灶，并在对应的下拉框中选择”yes”或者”no”。据沟通结果表示，强化灶与 IDH 的突变存在强相关性，所以在使用时，若人为选择出强化灶的存在性，则对于 IDH 突变状态的预测无需调用深度学习预测网络，即可反馈精准的结果；而对于没有人为干预的情况，则与 Ki-67 指数一般调用深度学习模型进行预测；

(4) 最后点击“预测分子标记物”，系统便会在 5 秒内给出对应患者的 IDH 突变状态和 Ki-67 指数水平的患者级预测结论。

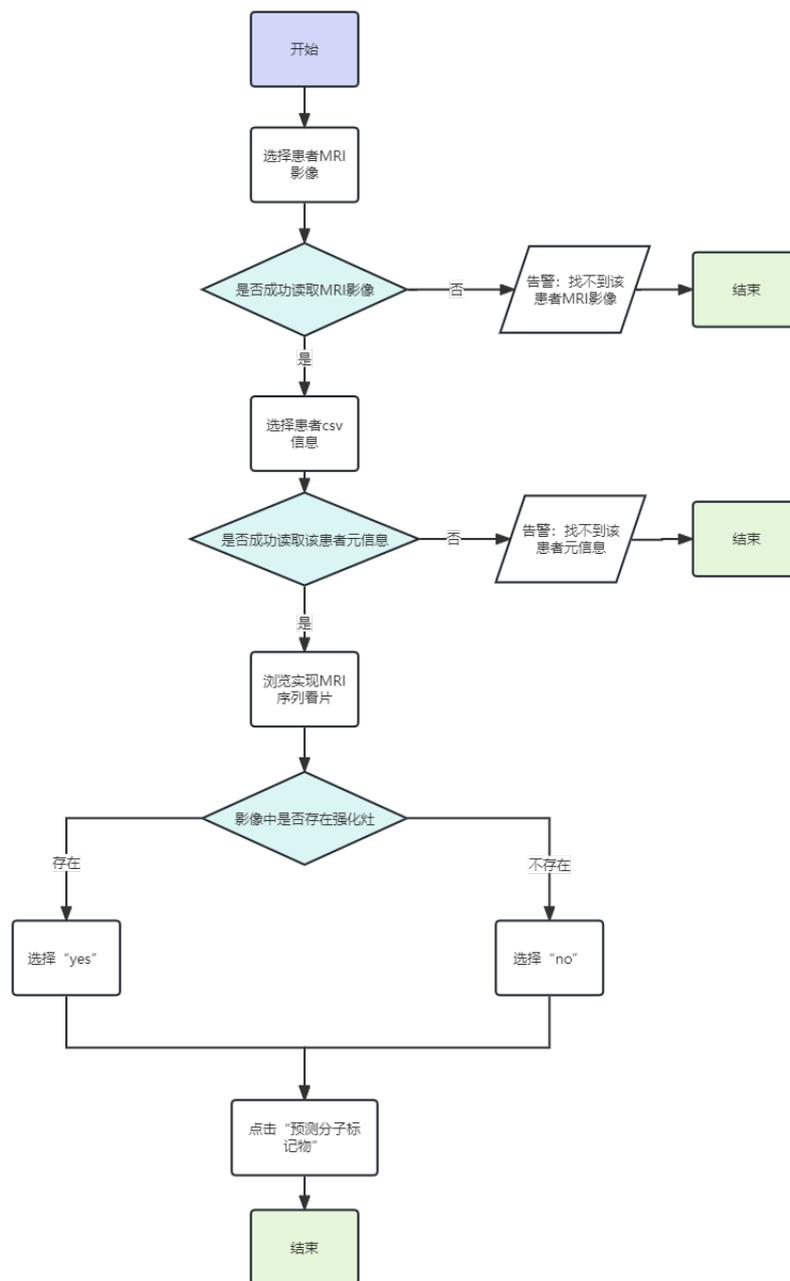


图 5-5 辅助诊断系统的流程图

### 5.3 本章小结

本章首先为诊断辅助模型落地完成了整个模型模块之间的级联设计；然后从将胶质瘤 Ki-67 指数预测算法在临床辅助诊断中落地的应用角度出发，介绍了其以客户端联网模式下的应用设计架构与具体实现。详细描述了辅助诊断系统的结构化数据目录以及使用步骤，能结合临床决策习惯来结合人为干预及人机交互的方案。本章的工作能够为辅助诊断系统真正的落地化应用带来重要的参考意义。

## 第 6 章 总结与展望

### 6.1 论文总结

脑神经胶质瘤在人体最常见的肿瘤之一，伴随着高病发率，高级别的胶质瘤患者在后期阶段的存活率非常低。然而受限于不均衡的医患比例和复杂的组织病理检测手段，导致胶质瘤的早期诊断困难重重，研究一个非侵入式预测胶质瘤 Ki-67 指数的诊断模型有突出的实际意义。受限于数据集规模与黑匣建模，相关端对端模型应用的探索是比较空缺的。在这个背景下，本文致力于相关研究成果的落地部署，主要集中在以下工作：

(1) 构建并整理了一个完整且可适用的湘雅胶质瘤分子分型 XiangyaGMS 数据集。现有的胶质瘤领域中，已有的公开数据集存在患者病例缺少，影像数据不足或分子标记物标签不齐整等问题，不足以支撑起胶质瘤相关算法的研究工作。本文基于近四年湘雅医院真实就诊记录，整理得到完整标签的 619 例患者，每例都包含有临床问诊信息，MRI 多序列影像及病灶勾画，病理检验依托的分子分型结果等。在构建过程中，本文会对影像数据作有效的清洗及预处理步骤，包括图像配准、颅骨剥离及均值归一化等方法，清洗后获得 MRI 影像规模接近 4 万。同时为确认该数据集的可行性验证，本文安排其在医学基准任务上进行实验来评估其性能，认定 XiangyaGMS 数据集能够作为相关领域的切实补充，有望为整个领域的研究工作带来重大贡献。

(2) 提出了一个基于医学先验知识补充的胶质瘤病灶分割算法。本文在专业医师团队的引导下，希望能够对分割算法的建模过程中，结合医学临床中的先验知识，赋予其更多的可解释性。具体的，病灶概率图谱 FPG 损失被提出来给到模型对胶质瘤易发生区域像素更多的注意力；同时，提出一个多任务信息引导网络 MIG-Net，MIG-Net 在建模过程中结合医学多任务间的定位及语义信息引导，能为整个模型带来一定的可解释性，平均取得 90.56% 的 DSC 指标。最后考虑到病灶分割模型性能的鲁棒性之于诊断系统的首要性，在跨域数据集 BraTS 上交叉验证了 MIG-Net 的性能，能满足泛化性能的要求。

(3) 探索了一种用于 MRI 多序列影像的胶质瘤 Ki-67 指数预测模型。MRI 影像天生带有不同的序列数据，本文工作的重要一步就是，设计有效的模型结构来充分挖掘 MRI 不同序列影像之间的同质特征。首先，本文采用不同的多序列融合方法来预测 Ki-67 指数，然后基于最优的融合方法，提出一个新的多序列多尺度模型 KiNet 来预测 Ki-67 指数的水平。此外，还针对医学 MRI 背景，提出一个二元多尺度 BMS 损失来监督来自同一患者数据的类内相似性。KiNet 能结合患

者的临床元信息，在集成分类器中平均取得 0.788 的 AUC 指标。基于端对端的单张 MRI 预测结果，在不同的决策融合方法中来验证得到可适用的患者级诊断结论，在患者级的 Ki-67 指数水平预测任务上达到平均 0.857 的 AUC 指标。

(4) 设计并实现了临床辅助系统的客户端联网应用。目前符合性能的相关研究大多停留在特征组合上，尚达不到应用部署的智能便捷要求。本文将多个模型以级联框架集成为整体，以实现诊断系统的闭环设计。最终将辅助诊断系统部署在客户端联网模式下，可实现 Ki-67 指数水平预测的功能及临床“阅片”的交互体验。

## 6.2 工作展望

本文针对胶质瘤 Ki-67 指数非侵入式预测算法及其应用进行了深入的研究，取得了一定的进展与成果，但目前仍存在一些有前景的探索方向：

(1) 目前建立的 XiangyaGMS 数据集中，仍有大量的患者缺少部分 MRI 序列影像，此外，现有整理的患者临床元信息只停留在性别年龄等基础信息，所能提供的特征有限。所以更精准的临床元信息是未来研究的倚仗之一。

(2) 本文提出的 KiNet 模型虽然能提升预测的性能，但是其包含过多的结构分支，这会加重模型的计算负荷。计算量与参数量会导致应用端部署的巨大麻烦，所以在未来一个比较明确的反向是模型的剪枝等轻量化操作。

(3) 对于目前的多任务信息引导网络 MIG-Net，建模思路虽然能被解释，但是仍没有给出足以支撑其可解释性的建模细节。比如其中的定位与语义信息构成需要被进一步解析，以给出更明确的解释性成分。

## 参考文献

- [1] Ostrom Q T, Cioffi G, Gittleman H, et al. Cbtrus statistical report: primary brain and other central nervous system tumors diagnosed in the united states in 2012–2016 [J]. *Neuro-oncology*, 2019, 21(5): 1-100.
- [2] 吴劲松, 毛颖, 姚成军, 等. 术中磁共振影像神经导航治疗脑胶质瘤的临床初步应用 (附 61 例分析) [J]. *中国微侵袭神经外科杂志*, 2007, 12(3): 105-109.
- [3] Lapointe S, Perry A, Butowski N A. Primary brain tumours in adults [J]. *The Lancet*, 2018, 392(10145): 432-446.
- [4] 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.
- [5] Zacharaki E, 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.
- [6] Ganau L, Paris M, Ligarotti G, et al. Management of gliomas: overview of the latest technological advancements and related behavioral drawbacks [J]. *Behavioural neurology*, 2015, 2015(3): 862634.
- [7] Tom M, Cahill D, Buckner J, et al. Management for different glioma subtypes: are all low-grade gliomas created equal? [J]. *American Society of Clinical Oncology Educational Book*, 2019, 39(39): 133-145.
- [8] Eckel-Passow J, Lachance D, Molinaro A, 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.
- [9] Buckner J, Giannini C, Eckel-Passow J, et al. Management of diffuse low-grade gliomas in adults—use of molecular diagnostics [J]. *Nature Reviews Neurology*, 2017, 13(6): 340-351.
- [10] 蒋杰杰, 关丽明, 徐克. 基于常规 T2WI 的放射基因组学模型对胶质瘤 p53 状态预测效能初探 [J]. *中国临床医学影像杂志*, 2021, 32(9): 609-612.
- [11] Atlas T C G. Comprehensive genomic characterization defines human glioblastoma genes and core pathways [J]. *Nature*, 2008, 455(7216): 1061-1068.

- [12] Menze B H, Jakab A, Bauer S, et al. The multimodal brain tumor image segmentation benchmark (brats) [J]. *IEEE transactions on medical imaging*, 2014, 34(10): 1993-2024.
- [13] Gooya A, Pohl K M, Bilello M, et al. Glistr: glioma image segmentation and registration [J]. *IEEE transactions on medical imaging*, 2012, 31(10): 1941-1954.
- [14] Goetz M, Weber C, Bloecher J, et al. Extremely randomized trees based brain tumor segmentation [C]//*Medical Image Computing and Computer Assisted Interventions: volume 14*. 2014: 6-11.
- [15] Reza S, Iftekharuddin K. Improved brain tumor tissue segmentation using texture features [C]//*Medical Image Computing and Computer Assisted Interventions: volume 27*. 2014: 30.
- [16] Kleesiek J, Biller A, Urban G, et al. Ilastik for multi-modal brain tumor segmentation [C]//*Medical Image Computing and Computer Assisted Interventions: volume 12*. 2014: 17.
- [17] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation [C]//*Medical Image Computing and Computer Assisted Interventions: volume 9351*. 2015: 234-241.
- [18] Isensee F, Jaeger P F, Kohl S A, et al. nnu-net: a self-configuring method for deep learning-based biomedical image segmentation [J]. *Nature methods*, 2021, 18(2): 203-211.
- [19] Jin Q, Meng Z, Sun C, et al. Ra-unet: A hybrid deep attention-aware network to extract liver and tumor in ct scans [J]. *Frontiers in Bioengineering and Biotechnology*, 2020, 8: 1471.
- [20] Çiçek, Abdulkadir A, Lienkamp S S, et al. 3d u-net: learning dense volumetric segmentation from sparse annotation [C]//*Medical Image Computing and Computer Assisted Interventions: volume 9901*. 2016: 424-432.
- [21] Kamnitsas K, Bai W, Ferrante E, et al. Ensembles of multiple models and architectures for robust brain tumour segmentation [C]//*Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries Workshop*. 2018: 450-462.
- [22] Kamnitsas K, Ledig C, Newcombe V F, et al. Efficient multi-scale 3d cnn with fully connected crf for accurate brain lesion segmentation [J]. *Medical image analysis*, 2017, 36(1): 61-78.
- [23] 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.
- [24] Myronenko A. 3d mri brain tumor segmentation using autoencoder regularization [C]//Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries Workshop. 2019: 311-320.
- [25] Isensee F, Kickingereder P, Wick W, et al. No new-net [C]//Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries Workshop. 2019: 234-244.
- [26] Zhu Z, Liu C, Yang D, et al. V-nas: Neural architecture search for volumetric medical image segmentation [C]//International conference on 3d vision. 2019: 240-248.
- [27] Xiao X, Lian S, Luo Z, et al. Weighted res-unet for high-quality retina vessel segmentation [C]//international conference on information technology in medicine and education. 2018: 327-331.
- [28] Zhou Z, Rahman Siddiquee M M, Tajbakhsh N, et al. Unet++: A nested u-net architecture for medical image segmentation [C]//Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support Workshop. 2018: 3-11.
- [29] Huang H, Lin L, Tong R, et al. Unet 3+: A full-scale connected unet for medical image segmentation [C]//IEEE International Conference on Acoustics, Speech and Signal Processing. 2020: 1055-1059.
- [30] Prasoon A, Petersen K, Igel C, et al. Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network [C]//Medical Image Computing and Computer Assisted Intervention. 2013: 246-253.
- [31] Setio A A A, Ciompi F, Litjens G, et al. Pulmonary nodule detection in ct images: false positive reduction using multi-view convolutional networks [J]. IEEE transactions on medical imaging, 2016, 35(5): 1160-1169.
- [32] Dvořák P, Menze B. Local structure prediction with convolutional neural networks for multimodal brain tumor segmentation [C]//Medical Computer Vision: Algorithms for Big Data Workshop. 2016: 59-71.
- [33] Zheng S, Jayasumana S, Romera-Paredes B, et al. Conditional random fields as recurrent neural networks [C]//IEEE international conference on computer vision. 2015: 1529-1537.
- [34] Liu Z, Li X, Luo P, et al. Semantic image segmentation via deep parsing network

- [C]//IEEE international conference on computer vision. 2015: 1377-1385.
- [35] Gu Z, Cheng J, Fu H, et al. Ce-net: Context encoder network for 2d medical image segmentation [J]. IEEE transactions on medical imaging, 2019, 38(10): 2281-2292.
- [36] Zhao H, Shi J, Qi X, et al. Pyramid scene parsing network [C]//IEEE conference on computer vision and pattern recognition. 2017: 2881-2890.
- [37] Wang X, Girshick R, Gupta A, et al. Non-local neural networks [C]//IEEE conference on computer vision and pattern recognition. 2018: 7794-7803.
- [38] Schlemper J, Oktay O, Schaap M, et al. Attention gated networks: Learning to leverage salient regions in medical images [J]. Medical image analysis, 2019, 53: 197-207.
- [39] Chen J, Lu Y, Yu Q, et al. Transunet: Transformers make strong encoders for medical image segmentation [J]. ArXiv, 2021, 2102: 04306.
- [40] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need [J]. Advances in neural information processing systems, 2017, 30.
- [41] Carion N, Massa F, Synnaeve G, et al. End-to-end object detection with transformers [C]//European Conference on Computer Vision. 2020: 213-229.
- [42] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16x16 words: Transformers for image recognition at scale [C]//International Conference on Learning Representations. 2021.
- [43] Liu Z, Lin Y, Cao Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows [C]//IEEE/CVF international conference on computer vision. 2021: 10012-10022.
- [44] Cao H, Wang Y, Chen J, et al. Swin-unet: Unet-like pure transformer for medical image segmentation [C]//European Conference on Computer Vision. 2023: 205-218.
- [45] 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.
- [46] Gutman D A, Cooper L A, 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.
- [47] 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.
- [48] Yao W, Liao Y, Li X, et al. Noninvasive method for predicting the expression of ki67 and prognosis in non-small-cell lung cancer patients: Radiomics [J]. *Journal of Healthcare Engineering*, 2022, 2022: 7761589.
- [49] 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.
- [50] 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.
- [51] 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: 3832-3839.
- [52] 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.
- [53] 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.
- [54] 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 imagingneural network for determination of idh status in gliomas [J]. *Clinical Cancer Research*, 2018, 24(5): 1073-1081.
- [55] 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: 321-327.
- [56] Li X, Xu Y, Xiang F, et al. Prediction of idh mutation status of glioma based on multimodal mri images [C]//International Conference on Intelligent Medicine and Image Processing. 2021: 39-44.
- [57] 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.
- [58] Vielzeuf V, Lechervy A, Pateux S, et al. Centralnet: a multilayer approach for multimodal fusion [C]//European Conference on Computer Vision Workshops. 2018:

- 575--589.
- [59] Wang Y, Wang Y, Guo C, et al. Sgpnnet: A three-dimensional multitask residual framework for segmentation and idh genotype prediction of gliomas [J]. Computational Intelligence and Neuroscience, 2021, 2021: 1-9.
- [60] Saha C, Hossain M F. Mri brain tumor images classification using k-means clustering, nsct and svm [C]//IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics. 2017: 329-333.
- [61] 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, 1: 456277.
- [62] Yogananda C G B, Shah B R, Yu F F, et al. A novel fully automated mri-based deep-learning method for classification of 1p/19q co-deletion status in brain gliomas [J]. Neuro-oncology advances, 2020, 2(4): 42-48.
- [63] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition [C]//IEEE conference on computer vision and pattern recognition. 2016: 770-778.
- [64] Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks [C]//IEEE conference on computer vision and pattern recognition. 2017: 4700-4708.
- [65] Chollet F. Xception: Deep learning with depthwise separable convolutions [C]//IEEE conference on computer vision and pattern recognition. 2017: 1251-1258.
- [66] Tong T, Wolz R, Wang Z, et al. Discriminative dictionary learning for abdominal multi-organ segmentation [J]. Medical image analysis, 2015, 23(1): 92-104.
- [67] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection [C]//IEEE international conference on computer vision. 2017: 2980-2988.
- [68] Caruana R. Multitask learning [J]. Machine learning, 1997, 28: 41-75.
- [69] Hajiramezanali E, Zamani Dadaneh S, Karbalayghareh A, et al. Bayesian multi-domain learning for cancer subtype discovery from next-generation sequencing count data [J]. Advances in Neural Information Processing Systems, 2018, 31.
- [70] 周志华. 机器学习 [M]. 清华大学出版社, 2016.
- [71] Girshick R. Fast r-cnn [C]//IEEE international conference on computer vision. 2015: 1440-1448.
- [72] Hu J, Shen L, Sun G. Squeeze-and-excitation networks [C]//IEEE conference on computer vision and pattern recognition. 2018: 7132-7141.
- [73] Woo S, Park J, Lee J Y, et al. Cbam: Convolutional block attention module [C]//

- European conference on computer vision. 2018: 3-19.
- [74] Jetley S, Lord N A, Lee N, et al. Learn to pay attention [C]//International Conference on Learning Representations. 2021.
- [75] Oktay O, Schlemper J, Folgoc L L, et al. Attention u-net: Learning where to look for the pancreas [C]//Medical Imaging with Deep Learning. 2018.
- [76] Du J, Li W, Lu K, et al. An overview of multi-modal medical image fusion [J]. *Neurocomputing*, 2016, 215: 3-20.
- [77] He K, Zhang X, Ren S, et al. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification [C]//IEEE international conference on computer vision. 2015: 1026-1034.
- [78] Selvaraju R R, Cogswell M, Das A, et al. Grad-cam: Visual explanations from deep networks via gradient-based localization [C]//IEEE international conference on computer vision. 2017: 618-626.
- [79] Ashburner J, Friston K J. Voxel-based morphometry—the methods [J]. *Neuroimage*, 2000, 11(6): 805-821.
- [80] Tao C, Wu W, Xu C, et al. Multi-representation fusion network for multi-turn response selection in retrieval-based chatbots [C]//ACM international conference on web search and data mining. 2019: 267-275.
- [81] Zhang H, Wu C, Zhang Z, et al. Resnest: Split-attention networks [C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022: 2736-2746.
- [82] Xie S, Girshick R, Dollár P, et al. Aggregated residual transformations for deep neural networks [C]//IEEE conference on computer vision and pattern recognition. 2017: 1492-1500.
- [83] Li X, Wang W, Hu X, et al. Selective kernel networks [C]//IEEE/CVF conference on computer vision and pattern recognition. 2019: 510-519.
- [84] Ke G, Meng Q, Finley T, et al. Lightgbm: A highly efficient gradient boosting decision tree [C]//International Conference on Neural Information Processing Systems. 2017: 3149–3157.
- [85] Chopra S, Hadsell R, LeCun Y. Learning a similarity metric discriminatively, with application to face verification [C]//IEEE Computer Society Conference on Computer Vision and Pattern Recognition: volume 1. 2005: 539-546.

## 攻读学位期间主要的研究成果

### 一、学术论文

[1] **Li Xuhui**, Xu Y, Xiang F, et al. KINET: A Non-Invasive Method For Predicting Ki67 Index Of Glioma[C]// IEEE International Conference on Image Processing, 2021: 150-154.(CCF C 类会议, 对应本文第三章内容)

[2] **Li Xuhui**, Xu Y, Xiang F, et al. Prediction of IDH Mutation Status of Glioma Based on Multimodal MRI Images[C]// International Conference on Intelligent Medicine and Image Processing, 2021: 39-44.(对应本文第五章内容)

[3] **Li Xuhui**, Zhao Xinyu, Ma Haoran and Xie Bin\*. Image Analysis and Diagnosis of Skin Diseases - A Review [J]. Current Medical Imaging, 2023, 19(3): 199-242. (Q3,IF:1.315)

### 二、发明专利与软件著作权

[1] 谢斌, 万思远, 徐勇, 李旭辉. 一种 ki67 指数确定方法及系统. 专利号: ZL202110353954.0, 授权公告号: CN113096086B.

[2] 中南大学. 基于异构信息的胶质瘤分子分型预测系统 V1.0. 登记号: 2021SR0693511, 证书号: 软著登字第 7416137 号.

### 三、主持和参与的科研项目

(1) 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”, 2020.9-至今, 主要参与成员。

(2) 企、事业单位委托项目: 综合场景 AI 事件检测, 2021.1-2022.7, 主要参与成员。

(3) 中南大学研究生自主探索创新项目: 基于放射基因组学的 Ki67 指数非侵入式预测, 2021.12-2022.12, 项目负责。

### 四、个人获奖情况

(1) 国家级, 中国高校计算机大赛——人工智能创意赛一等奖, 2021;

(2) 省部级, 第一届湖南省研究生人工智能创新大赛一等奖, 2020;

(3) 校级, 中南大学 2021-2022 学年优秀学生, 2022;

(4) 校级，中南大学 2021-2022 学年校级一等奖学金，2022；

#### 五、参加学术组织及活动

(1) 2021.07.25, 2021 IEEE International Conference on Image Processing.

(2) 2021.04.23, 2021 3rd International Conference on Intelligent Medicine and Image Processing.

## 致 谢

匆匆三年，时光马车又驶到人生的转角，于我个人而言，在中南的求学生涯，精彩而又略带遗憾！回忆起当我行至前一个转角时，周围确是充斥着分歧声。当时我的真实状态是满腔热血，故学府中南，属实是一个跳脱舒适圈的大胆决定。初到中南时，民主楼的斑驳红墙，麓山脚的葱郁林木，以及 AIRLab 的热情氛围，都能成为我中南生活的胸腔温暖。

毫无疑问，谢斌老师是我在中南的指引明灯。在我尚未正式进入中南学习前，谢老师已经就升学目的以及就业打算与我耐心讨论并阐明研究方向的培养方案。这个阶段让我全面了解到 AIRLab 的多领域布局，科研与实践落地的强延展性，是谢斌老师给每一位同学提供的求职技能包。专业上，谢斌老师多次在我的学术研究上一针见血地点明关键逻辑，在论文撰写上耐心为我逐字推敲其合理性；而在生活中，谢斌老师能其乐融融地参与 AIRLab 的团建活动，寓教于乐。谢斌老师给我印象最深的一点是，他非常愿意与每一位同学深入交流，主动询问同学们的想法，并给与相同的平等和尊重。

我取得的“成绩”同样离不开周围优秀伙伴的帮助。首当其中的自然是 AIRLab 的同门，“同门”的概念，我更多会将“同”描述在研究方向上。从高年级师兄身上，我亲身感受到的是一种真正意义上的传帮带。对于我”好奇宝宝“般的发问，师兄们甚至愿意停下手头的工作，启发式地引导我的思路以及实操。他们面对课题调研及项目交付时的从容与严谨，一旁的我耳濡目染，能为我后续的工作习惯提供良好的参考。毕业之际，回神一想，真的感谢师兄们的陪伴与教导，也感谢师弟师妹们的相处。此外，我也很感激中南的同学朋友们给我提供的帮助。最后，我想也需要对父母在我研究生期间的教诲有更多的感恩，相信父母实际付出的心血一定比我了解到的更多，更细致。

也是到了告别的日子，回忆起来这三年，在科研，生活体验上都还留有不少的遗憾。不过我相信，有遗憾并不是不完美，正是让我在今后能正视到这些维纳斯断臂。最后的最后，在这里我希望身边的所有人都身体健康，在自己的人生中活出自己想要的精彩。