中枢神经影像学・

MRI 序列以及注意力门、残差网络对 U-Net 脑肿瘤分割模型的影响

张巨,朱文珍,张顺,朱虹全,吴迪,刘栋

【摘要】目的:探讨 MRI序列选择以及注意力门、残差网络对 U-Net 脑肿瘤分割模型准确性的影响。方法:使用 BraTS 2021 的脑肿瘤病例 1251 例,排除经载剪后缺乏肿瘤图像特征的病例,以7:2:1 的比例分为训练组、验证组和测试组,使用 7 种不同序列组合(所使用序列包括 T₁WI、T₂WI、T₂ FLAIR、T₁ 增强)训练 U-Net 模型,使用 Friedman 检验和成对比较(经 Bonferroni 校正法调整显著性值)对比测试集的 Dice 系数。使用 4 通道图像以相同方法对比基线 U-Net、注意力 U-Net、残差 U-Net 以及注意力残差 U-Net 对脑肿瘤分割的准确性。结果:在分割肿瘤强化区域、核心区域时,无 T₁ 增强 序列的 3 通道组的 Dice 系数显著低于其他组;在分割全肿瘤时,无 T₂ 及 T₂ FLAIR 的 2 通道组的 Dice 系数显著低于其他组;在分割全肿瘤时,无 T₂ 及 T₂ FLAIR 的 2 通道组的 Dice 系数显著低于其他组;在分割全肿瘤时,无 T₂ 及 T₂ FLAIR 的 3 通道组显 著高于其他组,余组间差异无统计学意义。4 种 U-Net 模型仅在分割全肿瘤时存在显著差异,在进一步的成对比较中差异无统计学意义。结论:MRI 序列对 U-Net 模型化显著影响;与注意力门相比, 残差网络可能一定程度提高了 U-Net 模型的分割准确度。

【关键词】 图像分割;注意力门;残差网络;脑肿瘤

【中图分类号】R445.2;R739.41 【文献标志码】A 【文章编号】1000-0313(2023)07-0825-05

DOI:10.13609/j. cnki. 1000-0313. 2023. 07. 003

开放科学(资源服务)标识码(OSID):



Effects of MRI sequence and attention gates, residual networks on U-Net brain tumor segmentation model ZHANG Ju,ZHU Wen-zhen,ZHANG Shun, et al. Department of Radiology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China

[Abstract] Objective: To investigate the impact of MRI sequence selection and attention gates, residual networks on the accuracy of the U-Net brain tumor segmentation model. Methods: A total of 1251 brain tumor cases from the BraTS 2021 dataset were used, excluding cases lacking tumor image features after cropping. The cases were divided into training, validation, and test groups in a ratio of 7: 2:1. Seven different sequence combinations (including T1WI, T2WI, T2 FLAIR, and T1 contrast-enhanced) were used to train the U-Net model. The Friedman test and pairwise comparison (significance values were adjusted by the Bonferroni correction) were used to compare the Dice coefficients of the test set. The baseline U-Net, attention U-Net, residual U-Net, and attention-residual U-Net were compared using four-channel images in the same way. **Results**: The Dice coefficients of the 3-channel group without T_1 -enhancement sequence were significantly lower than those of the other groups when segmenting the Gd-enhancing tumor and tumor core. When segmenting the whole tumor, the Dice coefficients of the 2-channel group without T_2 and T_2 FLAIR were significantly lower than those of the other groups. The Dice coefficients of the 3-channel group without T_2 FLAIR were significantly lower than those of the other remaining groups, while the 4-channel group and the 3-channel group without T_1 were significantly higher than the other groups. There was no significant difference between the other groups. There were significant differences among the four U-Net models only when segmenting the whole tumor. There was no significant difference in the corrected pairwise comparisons. Conclu-

作者单位:430030 武汉,华中科技大学同济医学院附属同济医院放射科

作者简介:张巨(1989-),男,陕西延安人,博士研究生,主治医师,从事脑肿瘤成像及深度学习方面的研究。

通讯作者:刘栋, E-mail:liudongshole@163.com

基金项目:国家自然科学基金联合基金重点支持项目(U22A20354)

sion: The impact of MRI sequences on the segmentation performance of the U-Net model may be related to the annotation protocol and the feature information contained in the sequence. In this study, the exclusion of the T_1 sequence alone did not have a significant impact on the U-Net model. Compared to attention gates, residual networks may have improved the segmentation accuracy of the U-Net model.

[Key words] Image segmentation; Attention gate; Residual network; Brain neoplasms

兴趣区(region of interest, ROI)勾画是脑肿瘤影 像研究及临床定量分析的基础。手动勾画兴趣区工作 量大,效率低,受人员经验及主观判断影响较大,具有 明显的局限性。一些简单的算法如阈值分割、区域生 长^[1,2]等在勾画边界分明、特征简单、体素值差异明显 的目标时可以显著提高效率,是传统分割方法中常用 的辅助方式。然而医学图像往往具有极高的复杂性和 多样性,传统手段在包括脑肿瘤分割等多种使用场景 下效果欠佳,常需大量手工调整。深度学习技术实现 了特征工程的自动化,相对传统技术具有明显优势。 语义分割(semantic segmentation)是目前使用深度学 习对医学图像进行分割的常用的方式[3]。其中相较传 统全卷积网络(fully convolutional networks, FCN), U-Net 结构包含了一个与编码路径对称的解码路径并 通过相应层级拼接的方式提供了更为精确的定位[4], 该结构因为简单可靠被广泛运用于图形分割领 域[5,6]。部分研究认为对 U-Net 的结构调整可能可以 进一步提高模型的准确性,Oktay 等^[7]提出的注意力 门机制可以增加目标区域的权重,He 等^[8]提出的残差 网络有效的应对了网络退化的问题。深度学习模型的 训练、使用及数据的储存、传输均需要一定的硬件及时 间成本,该成本可通过降低数据量来有效控制,然而充 足的数据量是充分训练模型的基础。识别并剔除无法 显著提升模型分割准确率的冗余数据则具有实用意 义。本研究使用 BraTS 2021 的数据,探索了不同序列 组合对 U-Net 模型的影响,并对比了基线 U-Net 以及 添加了注意力门、残差网络模块的模型在分割脑肿瘤 及亚区方面的差异。

材料与方法

1. 图像数据处理

使用 BraTS 2021 脑胶质瘤数据 1251 例,每个病例包括 T₁、T₁ 增强、T₂、T₂ FLAIR 四个序列的 图像以及一个标签图像,标签包 括:0,背景(非肿瘤区域);1,肿瘤 坏死区域;2,水肿、浸润区域;4,肿 瘤强化区域。将所有 MRI 图像归 一化处理,标签值4改为3。将 MRI 图像使用7种序列组合(4通 道组、无 T₁ 的 3 通道组、无 T₂ 的 3 通道组、无 T₂ FLAIR 的 3 通道组、无 T₁ 增强的 3 通道组、无 T₁ 及 T₂ 的 2 通道组,无 T₂ 及 T₂ FLAIR 的 2 通道组)堆叠 为多通道图像,对图像进行裁剪,保留图像中心范围为 $128 \times 128 \times 128$ 的体素。由于部分病例经裁剪后目标 体素较少且缺乏肿瘤图像特征,这部分病例被排除。 剩余的病例以 7:2:1 的比例分为训练组、验证组、和测 试组。

2. 模型构建

基线 U-Net 模型:使用 3D U-Net 模型,结构如图 1,编码器包括 5 层子模块,每层子模块包括 2 个卷积 层(首个子模块 filters=16,后续子模块 filters 依次乘 以 2;kernel_size=3)、2 个激活函数层(activation= LeakyReLU, alpha=0.01)、以及一个随机失活层,前 4 个子模块包括一个最大池化层(pool_size=2)。解 码器包括 4 层子模块,每层子模块包括一个反卷积层 (首个子模块 filters=128,后续子模块 filters 依次除 以 2;kernel_size=2)、2 个卷积层(首个子模块 filters =128,后续子模块 filters 依次除以 2;kernel_size= 3)、2 个激活函数层(activation=LeakyReLU, alpha= 0.01)、以及一个随机失活层,其中反卷积层的输出接 受来自相应编码层的跳跃连接。

注意力 U-Net 模型:编码器与基线模型一致。解码器与基线模型间的区别在于卷积层前的注意力门结构(图 2):来自深层的信号经卷积(首个注意力门结构filters=128,余依次除以 2,kernel_size=1)、激活函数(ReLU)、卷积(同前)后,与来自相应编码层的跳跃连





图1 基线 U-Net 模型。





接的信号经卷积层(首个注意力门结构 filters=128, 余依次除以 2,kernel_size=1,strides=2)后相加,经 过激活函数 ReLU、卷积层(filters=1,kernel_size= 1)、激活函数 Sigmoid 后,上采样至前述跳跃连接的大 小并与之相乘,经卷积层(首个注意力门结构 filters= 128,余依次除以 2,kernel_size=1)、批量归一化后输 出。

残差 U-Net 模型:与基线模型的区别在于各子模 块第 2 次卷积后的输出接受第 1 次卷积前的输入逐点 卷积(filters 与相应子模块卷积层一致,kernel_size= 1)后的跳跃连接(图 3)。

残差注意力 U-Net 模型:即结合上述注意力门及 残差模块至基线 U-Net 的模型。

3. 模型训练

使用前述 7 种不同多通道图像的训练组和验证组 训练基线 U-Net,使用 4 通道图像的训练组和验证组 训练残差 U-Net、注意力 U-Net、残差注意力 U-Net。 损失函数使用 Categorical Focal Loss 结合 Dice Loss,





优化器使用自适应矩估计 (Adaptive Moment Estimation, Adam),初始 Learning rate 为 0.0001,Batch size 为4,每种序列 组合、模型至少训练 1000 个 Epoch,保存每次验证 IOU 分数提高 的模型。训练使用一块 NVIDIA GeForce RTX 4090 显卡,在 TensorFlow 框架(2.10.1)/Keras API(2.10.0)下完成。

4. 模型对比

选取每次训练 1000 个 Epoch 中验证 IOU 分数最高的模型进行 对比,使用各模型对测试组(n= 116)进行图像分割,计算分割结果

的 Dice 系数。使用 Friedman 检验和成对比较(经 Bonferroni 校正法调整显著性值)对比不同序列组合 以及不同模型间 Dice 系数的差异,采用 IBM SPSS Statistics 26 进行统计学分析。

结 果

经过裁剪,最终 1151 例患者被纳入研究,其中训 练组 805 例,验证组 230 例,测试组 116 例。

使用不同序列组合训练基线 U-Net 模型分割肿 瘤各区域的 Dice 系数均值、标准差及秩均值见表 1。 在分割肿瘤强化区域(Gd-enhancing tumor, ET)时, 无增强序列的 3 通道组 Dice 系数低于其他 6 组(P < 0.001),余组间差异无统计学意义;在分割肿瘤核心区 域(tumor core, TC)时,同样是无增强序列的 3 通道组 Dice 系数低于其他 6 组(P < 0.001),余组间差异无统 计学意义;在分割全肿瘤(whole tumor, WT)时,无 T₂ 和 T₂ FLAIR 的 2 通道组 Dice 系数显著低于其他组 (P < 0.001),无 T₂ FLAIR 的 3 通道组 Dice 系数低于 除前述 2 通道组以外的其他组(P < 0.001),4 通道组 和无 T₁ 的 3 通道组的 Dice 系数高于其他组(P < 0.05),余组间无差异无统计学意义。

基线U-Net、注意力U-Net、残差U-Net、注意力

| 衣I 谷伊州组合 Dice 糸剱母狼、孙准左及伏, | 〒1 谷序列组合1 | Dice糸致り | 吲隹、杯准 | 主左 戊 秖 | 131 |
|---------------------------|-----------|---------|-------|--------|-----|
|---------------------------|-----------|---------|-------|--------|-----|

| 序列 | ET(n=111) | TC(n=116) | WT(n=116) |
|---------------------------|------------------|------------------|-------------------------|
| 4C | 0.876±0.113,4.70 | 0.893±0.158,4.67 | 0.911±0.077,5.31 |
| 3C(无 T ₁) | 0.877±0.115,4.61 | 0.901±0.116,4.71 | 0.911±0.078,5.33 |
| 3C(无 T ₂) | 0.875±0.128,4.59 | 0.884±0.170,4.39 | 0.896±0.105,4.38 |
| 3C(无T ₂ FLAIR) | 0.879±0.105,4.53 | 0.903±0.115,4.59 | 0.884±0.823,2.96 |
| 3C(无 T ₁ 增强) | 0.490±0.224,1.02 | 0.645±0.248,1.38 | 0.900±0.818,4.34 |
| $2C(ftarrow T_1, T_2)$ | 0.868±0.131,4.02 | 0.876±0.193,4.00 | 0.896±0.110,4.33 |
| $2C(ft T_2, T_2 FLAIR)$ | 0.876±0.111,4.54 | 0.882±0.164,4.25 | $0.813 \pm 0.139, 1.35$ |

注:C为通道数;ET组因5例无强化标注,n=111

残差 U-Net 模型对一例颞叶脑肿瘤的分割结果见图 4。

基线 U-Net、注意力 U-Net、残差 U-Net、残差注 意力 U-Net 的 Dice 系数均值、标准差及秩均值见表 2。在分割肿瘤强化区域和核心区域时,4 种模型间的 Dice 系数差异无统计学意义。在分割全肿瘤时, Friedman 检验中 4 种模型间差异有统计学意义(P< 0.05),但在进一步的成对比较中,经 Bonferroni 校正 法调整显著性值后,两两间差异均无统计学意义。

讨论

本研究结果显示剔除 T1 增强序列会显著降低模型分割肿瘤强化区域及核心区域的准确性,而剔除 T2 FLAIR、尤其是同时剔除 T2 FLAIR 和 T2 会显著降低模型分割全肿瘤的准确度,以上结果可能与 BraTS

的标注原则有关。BraTS 中对肿 瘤强化区域的定义是相对于 T₁ 平扫和影像表现正常的白质,增强 后高信号的区域。肿瘤的核心区 域则包括肿瘤的强化区域和囊变、 坏死区。囊变、坏死区的典型表现 为T₁增强图像上的低信号。全 肿瘤的定义则进一步包括了肿瘤 周围的水肿及肿瘤浸润区域,通常 表现为 T₂ FLAIR 上异常的高信 号^[9]。T₁ 增强和 T₂ FLAIR 序列 可以提供较为明确的组织对比度 和肿瘤边缘信息,剔除这些序列 后,可能会导致模型对特征的提取 不充分,降低准确率的同时提高误 检率。且这两种序列可以提供不 同的肿瘤信息,去掉一种或两种, 则可能导致模型的鲁棒性降低,难 以处理相对复杂的分割任务。在 剔除 T₂ FLAIR 的基础上进一步 剔除 T₂后,模型对全肿瘤分割的 准确性进一步显著下降可能是因 为T₂序列同样对肿瘤的水肿或 浸润区较为敏感,图像存在较好的 对比度,将其剔除后,剩余序列 (T₁、增强)在部分病例相应区域的图像特征不足导致 的。本研究中4通道组和无T₁的3通道组在分割全 肿瘤时 Dice 系数显著高于其他组,结合分割其他区域 的结果,这两组具有较好的综合表现,且二者间差异无 统计学意义。该结果一方面提示我们训练和使用模型 时,保证信息的丰富度对模型分割的准确率具有重要 意义,另一方面也提示我们序列并非越多越好。本研 究中,相对于4通道组,剔除T₁平扫序列未对模型的 分割准确性产生显著性的影响,可能是由于T₁序列 在上述标注区域常无明显的对比度或清晰的边界,可 能所蕴含的相应图像特征较为稀少且不稳定,T₁的加 人可能无法提供足够提升模型表现的有用信息,反而 增加了噪声和冗余信息(针对标注区域),提高了模型 的复杂度和过拟合风险。影像组学的研究认为医学图 像包含大量人眼不易直接捕捉的图像特征^[10,11],且这



图 4 一个颞叶肿瘤的 T_1 、 T_1 增强、 T_2 、 T_2 FLAIR 图像,以及预先标注的 标签、4 种 U-Net 模型的分割结果。

表 2 各序列组合 Dice 系数均值、标准差及秩均值

| 序列组合 | ET(n=111) | TC(n=116) | WT(n=116) |
|-------------|------------------|------------------|-------------------------|
| 基线 U-Net | 0.876±0.113,2.39 | 0.893±0.158,2.48 | 0.911±0.077,2.32 |
| 注意力 U-Net | 0.875±0.126,2.29 | 0.895±0.152,2.32 | 0.909±0.071,2.31 |
| 残差 U-Net | 0.879±0.112,2.62 | 0.898±0.133,2.69 | 0.913±0.074,2.70 |
| 残差注意力 U-Net | 0.878±0.121,2.70 | 0.885±0.179,2.50 | $0.909 \pm 0.078, 2.67$ |

部分图像特征同样包含来源于组织的结构或功能信息,具有分析和诊断的价值。由于深度学习特征工程的自动化,本研究的结果可能提示,由 T₂ 及 T₂ FLAIR 图像特征所反映的脑肿瘤组织和边界信息,并 未被 T₁ 图像所完整和可靠地蕴含,或至少不能被本 研究所使用模型捕捉到。

本研究中对比的 4 种 U-Net 模型仅在分割全肿 瘤时表现出显著差异,且在随后的成对比较中,若经 Bonferroni 校正法调整显著性值后则两两间差异无统 计学意义。部分研究认为对 U-Net 的结构调整可能 不如非结构调整有效,如 Isensee 等提出的 nnU-Net ("no-new-Net"),仅围绕基础的 U-Net 模型调整,将 重点集中于数据处理、模型训练、推理、后处理上,其自 适应的特点令该框架具有良好的泛用性和优秀分割准 确性^[12]。本研究的结果可能提示部分对 U-Net 结构 的调整可以提高其在特定领域的分割效果,但也可能 在其他领域增加过拟合风险以及降低泛用性。本研究 中残差模型在分割全肿瘤时具有最大的 Dice 系数均 值和秩均值,而注意力模型则具有最小的 Dice 系数均 值和秩均值,结合 Friedman 检验结果,笔者推测在分 割全肿瘤时,残差网络可能较注意力门更有优势。

残差网络通过跳跃连接可以实现恒等映射的效 果,有效应对了深层网络的退化问题^[8]。在本研究中, 模型可能通过残差模块减少了信号的损失,保证了网 络中梯度信号的强度。此外,医学图像的像素之间是 存在局部相关的,则反向传播时的梯度也具有相关性, 但随着网络深度的增加,该梯度间的相关性会越来越 弱,难残差网络则极大的缓解了该相关性的衰减^[13], 保留了更多的梯度空间结构信息。

注意力门的加入会令模型逐步增加对图像兴趣区 的相对权重,理论上可以增加模型的敏感性和准确性。 而本研究中的注意力模型表现并未显著高于其他模型,笔者推测可能由于脑肿瘤形态、结构、大小、位置、 信号等的异质性较大,测试集中可能存在的一些与训 练集差异较大的病例,注意力门机制由于缺乏相应的 学习从而导致关注区域的错误和分类能力的下降。

综上所述,在脑肿瘤分割模型的构建与使用过程 中,输入图像的选择需要结合具体使用场景,在保证信 息丰富度的同时也要避免冗余。多余的序列无法提高 模型表现,反而可能引入噪声和冗余信息,提高模型的 复杂度和过拟合风险,此外还会增加数据的储存、传 输、运算的时间及硬件成本。在分割模型的选择方面, 残差网络可能减少了模型的信号损失,保留了更多的 原始图像特征信息,确保了梯度信号的强度,且减少了 梯度间相关性信息的损失,从而提高了模型的表现。

参考文献:

- [1] Sezgin M, Sankur BL. Survey over image thresholding techniques and quantitative performance evaluation[J]. J Electronic Imaging, 2004,13(1):146-168.
- [2] Adams R, Bischof L. Seeded region growing [J]. IEEE Transactions on pattern analysis and machine intelligence, 1994, 16(6): 641-647.
- [3] Asgari Taghanaki S, Abhishek K, Cohen JP, et al. Deep semantic segmentation of natural and medical images: a review[J]. Artif Intell Rev, 2021, 54:137-178.
- [4] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. Proceedings of the Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015:18th International Conference, Munich, Germany, October 5 ~9,2015, Proceedings, Part III 18, F, 2015[C]. Springer, 2015.
- [5] 谢婷婷,刘想,林子楹,等. 基于 U-Net 实现 CT 图像上肝段自动 分割和术前评估的初步研究[J]. 放射学实践,2023,38(1):47-51.
- [6] 李晓庆,王可欣,额·图娅,等.利用 U-net 算法在头 CT 平扫图像 上分割脑梗死的初步探究[J].放射学实践,2022,37(6):669-675.
- [7] Oktay O, Schlemper J, Folgoc LL, et al. Attention U-Net: Learning where to look for the pancreas [EB/OL]. arXiv: Computer Vision and Pattern Recognition, 2018. https://arxiv.org/abs/1804. 03999.
- [8] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[EB/OL]. arXiv: Computer Vision and Pattern Recognition, 2015. https://arxiv.org/abs/1512.03385.
- [9] Baid U,Ghodasara S,Mohan S, et al. The RSNA-ASNR-MICCAI BraTS 2021 benchmark on brain tumor segmentation and radiogenomic classification [EB/OL]. arXiv: Computer Vision and Pattern Recognition, 2021. https://arxiv.org/abs/2107.02314v2.
- [10] Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data[J]. Radiology, 2016, 278(2): 563-577.
- [11] Van Timmeren JE, Cester D, Tanadini-Lang S, et al. Radiomics in medical imaging-" how-to" guide and critical reflection [J]. Insights into imaging, 2020, 11(1):1-16.
- [12] Isensee F, Petersen J, Klein A, et al. nnU-net: self-adapting framework for u-net-based medical image segmentation [EB/ OL]. arXiv: Computer Vision and Pattern Recognition, 2018. https://arxiv.org/abs/1809.10486.
- [13] Balduzzi D, Frean M, Leary L, et al. The shattered gradients problem: If resnets are the answer, then what is the question? Proceedings of the International Conference on Machine Learning, F, 2017[C]. PMLR, 2017.

(收稿日期:2023-05-01 修回日期:2023-06-04)