❖医学物理与工程学

Combining multiple MR sequences transfer learning networks for automatic grading of glioma

LI Yang, *SONG Yue*, *ZHANG Shuli**, *MU Weibin*, *LIANG Minghui*
(*School of Medical Technology*, *Qiqihar Medical University*, *Qiqihar* 161003, *China*)

[**Abstract**] **Objective** To propose an automatic grading algorithm with combining multiple MR sequences transfer learning networks for low grade glioma (LGG) and high grade glioma (HGG), and to evaluate its efficacy. **Methods** Axial MR T1WI, T2WI and fluid attenuated inversion recovery (FLAIR) images of 76 LGG and 259 HGG patients were extracted from public database, each consisted 20 adjacent images of the largest tumor layer on the axial image, including 6 700 images which were divided into training set (n=4 690), validation set (n=1 005) or test set (n=1 005) at the ratio of 7∶1.5∶1.5 with the same random sequence. GoogLeNet pretrained network was used as transferred source of the parameters of glioma grading model, and the output module was redesigned. The single sequence models, *i.e.* T1WI, T2WI and FLAIR model, were trained respectively, and the convergence of the models were judged according to the accuracy curve and loss value curve during training process. The combined multi-sequence model voting mechanism was introduced to reduce the impact of single sequence model on misclassification, then the performances of single sequence model and combined multi-sequence model were evaluated based on the test set. **Results** The accuracy curves of glioma grading of each single sequence model in both training set and validation set showed steady upward trend, and the loss value curves showed steady downward trend, and then gradually converged. The area under the curve (AUC) of T1WI, T2WI, FLAIR model and combined multi-sequence model for grading of glioma in test set was 0.951 3, 0.934 2, 0.961 4 and 0.995 0, and the accuracy was 97.01%, 97.01%, 98.11% and 99.00%, respectively. **Conclusion** Combining multiple MR sequences transfer learning networks for automatic grading of glioma was concise and highly efficient.

[**Keywords**] glioma; neoplasm grading; magnetic resonance imaging; deep learning
DOI:10.13929/j.issn.1003-3289.2022.11.027

联合多 MR 序列迁移学习网络用于自动分级胶质瘤

李 阳,宋 悦,张淑丽*,穆伟斌,梁明辉
(齐齐哈尔医学院医学技术学院,黑龙江 齐齐哈尔 161003)

[**摘 要**] **目的** 提出一种联合多 MR 序列迁移学习网络算法用于自动分级低级别(LGG)与高级别(HGG)胶质瘤,并评估其效能。**方法** 于开源数据库中提取 76 例 LGG 和 259 例 HGG 患者头部 MR 轴位 T1WI、T2WI 及液体衰减反转恢复(FLAIR)序列图像,均包含与轴位图像所见肿瘤最大层面相邻的 20 个层面图像,共 6 700 幅图像;采用相同的随机数列按 7∶1.5∶1.5 比例将各序列图像分为训练集(n=4 690)、验证集(n=1 005)及测试集(n=1 005)。以 GoogLeNet 预训练网络为胶质瘤分级模型的参数迁移源,重新设计输出模块,分别训练 T1WI、T2WI 及 FLAIR 单一模型,根据训练过程中的准确率和损失值曲线判断其收敛性;引入联合多序列模型投票机制,以降低单一序列模型对误分类的影响,利用测试集数据评价单一序列模型及联合多序列模型的效能。**结果** 各单一序列模型对训练集和验证集胶质瘤分级的准确率曲线均呈稳步上升趋势,损失值曲线均呈稳步下降趋势,之后均逐渐收敛。单一 T1WI、T2WI、FLAIR 模型及联合多序列模型

[**第一作者**] 李阳(1998—),男,四川达州人,在读硕士。研究方向:计算机视觉在 MRI 中的应用。E-mail: 1913630863@qq.com
[**通信作者**] 张淑丽,齐齐哈尔医学院医学技术学院,161003。E-mail: zhangshuli63105@163.com
[**收稿日期**] 2022-07-17 [**修回日期**] 2022-09-26

对测试集胶质瘤分级的曲线下面积(AUC)分别为 0.951 3、0.934 2、0.961 4 及 0.995 0,准确率分别为 97.01%、97.01%、98.11%及 99.00%。**结论**　以联合多 MR 序列迁移学习网络进行胶质瘤自动分级过程简洁、效能高。

[**关键词**]　胶质瘤;肿瘤分级;磁共振成像;深度学习
[**中图分类号**]　R739.41;R445.2　[**文献标识码**]　A　　[**文章编号**]　1003-3289(2022)11-1715-05

胶质瘤是起源于大脑或脊髓的常见肿瘤,约占中枢神经系统肿瘤的 30%;WHO 分级系统根据病理特征和遗传学特征将其分为Ⅰ～Ⅳ级,其中Ⅰ～Ⅱ级为低级别胶质瘤(low-grade glioma, LGG),Ⅲ～Ⅳ级为高级别胶质瘤(high-grade glioma, HGG)。迁移学习网络利用大规模数据模型参数降低训练成本、解决数据量不足问题,近年来其在医学影像学中的应用成为研究热点。本研究提出一种联合多 MR 序列迁移学习网络胶质瘤自动分级算法,并评估其效能。

1　资料与方法

1.1　构建数据集　于开源数据集——2019 年脑部肿瘤分割竞赛(brain tumor segmentation challenge 2019, Bra TS2019)数据集[1]内来源于多机构的术前 MRI 中选取 76 例 LGG 和 259 例 HGG 患者的带标签头部平扫轴位 T1WI、T2WI、液体衰减反转恢复(fluid attenuated inversion recovery, FLAIR)序列数据用于训练模型,均含与轴位图像所示肿瘤最大层面相邻的 20 个层面图像,共 6 700 幅图像;将图像格式转化为[224,224,3],以相同的随机数列按 7∶1.5∶1.5 比例将各序列图像划分为训练集、验证集和测试集,分别包含 4 690、1 005 及 1 005 幅图像。

1.2　方法

1.2.1　数据增强　以离线增强和在线增强两种方式对训练集数据进行增强。离线增强:对原图进行水平翻转后,分别将原图和水平翻转图像旋转 5°和－5°,得到 6 倍增强效果,共 28 140 幅图像(图 1);在线增强则是在训练过程中随机对训练批量(batch)进行指定范围内的变换操作,即对每个 batch 中的图像随机进行 15 个像素的水平和垂直平移、以 50%的概率水平反射、随机于－45°～45°旋转及随机在 20°内进行水平和垂直错切变换,故每期(epoch)训练输入均不相同。

1.2.2　选择预训练网络　GoogLeNet 预训练网络的参数量和深度适宜且错误率较低(表 1),故以之作为迁移源。

1.2.3　改进 GoogLeNet 预训练网络　GoogLeNet 预训练网络的提取特征部分包括 5 个模块,前 2 个由卷积层和池化层、后 3 个模块由 Inception 模块(图 2)和池化层构成。重新设计其输出模块为节点数量分别为 1 024、512 和 2 的 3 个全连接层(图 3),内含大量参数,在 3 个全连接层间添加随机丢失概率为 0.5 的 Dropout 层,以减少神经元共同适应的复杂度、促进各神经元融合。

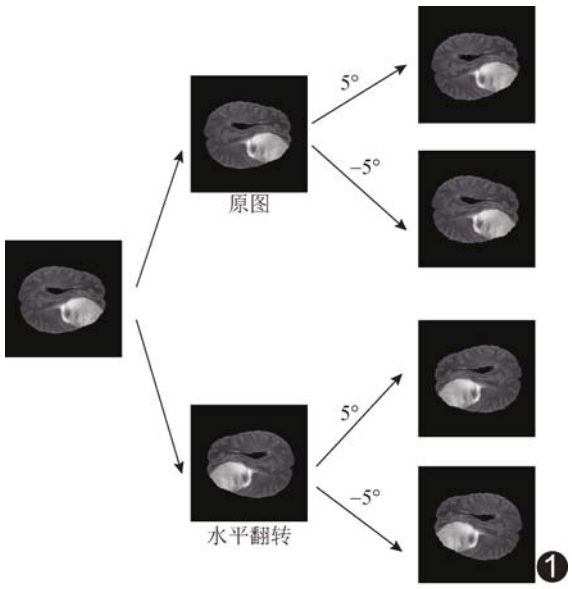

图 1　数据离线增强示意图

表 1　各常见预训练网络的参数量、深度及在 CIFAR100 数据集的错误率

预训练网络	参数量(M)	深度	Top-1 错误率(%)	Top-5 错误率(%)
MobileNet-v2	2.4	53	31.92	9.02
VGG-16	34.0	16	27.07	8.84
Xception	21.0	71	25.07	7.32
Inception-v3	22.3	48	22.81	6.39
GoogLeNet	6.2	22	21.97	5.94

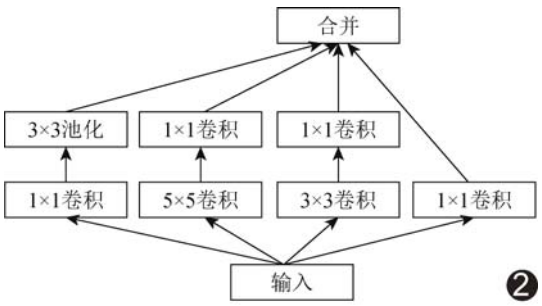

图 2　Inception 模块的结构图

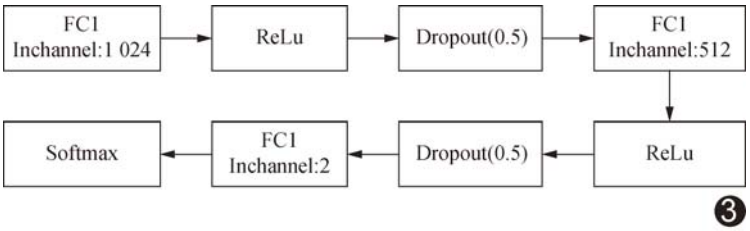

图 3　改进的 GoogLeNet 预训练网络输出模块图

1.2.4　选择损失函数及设置迁移学习模型训练参数　选取二值交叉熵（KL 散度）函数[2]作为各序列迁移学习模型的损失函数（公式 1）进行训练，以使损失值（损失函数的值）不断趋于 0。

$$D_{KL}(p \,||\, q) = \sum_{i=1}^{n} p(x_i)\log(p(x_i)) - \sum_{i=1}^{n} p(x_i)\log(q(x_i)) \quad (1)$$

其中，$p(x)$ 和 $q(x)$ 是同一随机变量 x 的 2 个单独的概率分布，$p(x)$ 为 MRI 胶质瘤级别标签的概率分布，$q(x)$ 为模型预测结果的概率分布。

将迁移学习网络输出模块的全局学习率设置为 0.000 1，输出模块学习率为 0.001；将批量尺寸（batchsize）设置为 128。对迁移学习模型进行 400 期训练后，发现其在 100 期时收敛、之后维持低损失值，故将训练的最大期数设置为 100。

1.2.5　联合多序列迁移学习网络投票机制　完成训练后，分别以 3 种单一序列模型预测测试集中各样本的胶质瘤分级，以结果中类别票数≥2 者为联合多序列模型的最终结果（图 4）。

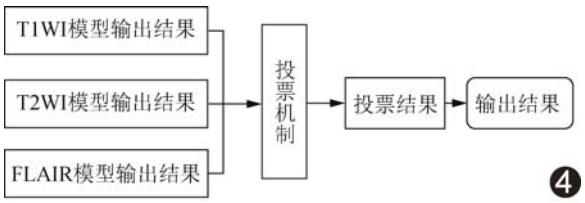

图 4　联合多序列模型的投票机制

1.2.6　评价指标　以准确率、敏感度、特异度、F1 分数（F1-score）和受试者工作特征（receiver operating characteristic，ROC）曲线评价单一序列及联合多序列模型进行胶质瘤分级的效能，曲线下面积（area under the curve，AUC）越接近 1 则分类效能越好。计算公式见 2～5。

$$\text{准确率}=\frac{TP+TN}{TP+FP+TN+FN} \quad (2)$$

$$\text{敏感度}=\frac{TP}{TP+FN} \quad (3)$$

$$\text{特异度}=\frac{TN}{FP+TN} \quad (4)$$

$$\text{F1 分数}=\frac{2\times\text{精确度}\times\text{召回率}}{\text{精确度}+\text{召回率}} \quad (5)$$

其中，$\text{精确度}=\frac{TP}{TP+FP}$，$\text{召回率}=\frac{TP}{TP+FN}$，TP、TN、FP 和 FN 分别为预测的真阳性数、真阴性数、假阳性数及假阴性数。

2　结果

各单一序列模型对训练集和验证集胶质瘤分级的准确率曲线（图 5）均呈稳步上升趋势，前 60 期上升较快，之后上升速度趋缓并逐渐收敛；损失值曲线（图 6）均呈稳步下降趋势，前 60 期下降较快，之后逐渐收敛。

T1WI、T2WI 及 FLAIR 模型对测试集胶质瘤分级的 AUC 分别为 0.951 3、0.934 2 及 0.961 4，其中 T2WI 模型的诊断敏感度最高，为 1.000 0，FLAIR 模型的其他各项指标均较高，准确率达 98.11%；联合多序列模型的 AUC 达 0.995 0，准确率为 99.00%。见表 2 及图 7。

3　讨论

HGG 具有高侵袭性，生长速度快且无规律，预后差；LGG 生长速度相对较慢，患者生存期可达 5～10 年。准确分级胶质瘤对制定治疗方案及预测预后具有重要意义。组织病理学检查为胶质瘤分级的金标准，但活检具有侵入性，且存在癫痫及感染等并发症。MRI 是常用非侵入性影像学诊断方法，可作为临床决策的重要依据；多项常规 MRI 定量特征可用于鉴别 HGG 与 LGG[3]，表明 MRI 可提供表征胶质瘤类别的信息。

机器学习方法现已广泛用于胶质瘤分级相关研究。机器学习模型的效能与手动提取的图像特征的质量密切相关，而胶质瘤具有高异质性，导致选定和提取特征的过程复杂，机器学习模型难以适应。TIAN 等[4]基于常规 MRI、弥散加权成像（diffusion weighted imaging，DWI）和动脉自旋标记（arterial spin labeling，ASL）图像提取特征，用于训练支持向量机（support vector machines，SVM）模型，其区分 HGG 与 LGG 的准确率为 96.8%。ZHANG 等[5]将基于 DWI 分割得到的胶质瘤区域覆盖于弥散张量成像（diffusion tensor imaging，DTI）平均弥散率（mean diffusivity，MD）和各向异性分数（fractional anisotropy，FA）图上，再提取

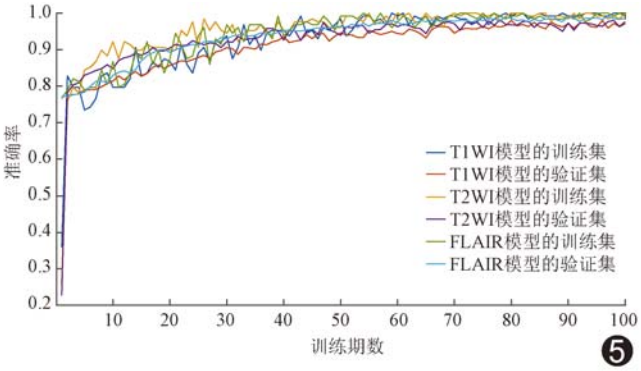

图 5　各单一序列模型对训练集和验证集胶质瘤分级的准确率曲线

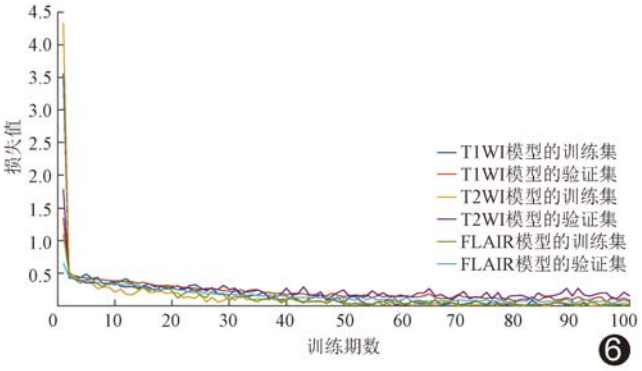

图 6　各单一序列模型对训练集和验证集胶质瘤分级的损失值曲线

表 2　联合多序列模型及单一序列模型对测试集胶质瘤分级的效能

模型	真阳性（幅）	假阳性（幅）	真阴性（幅）	假阴性（幅）	敏感度（%）	特异度（%）	准确率（%）	F1 分数	AUC
T1WI 模型	766	19	209	11	98.58	91.67	97.01	0.980 8	0.951 3
T2WI 模型	777	30	198	0	100	86.84	97.01	0.981 1	0.934 2
FLAIR 模型	775	17	211	2	99.74	92.54	98.11	0.987 9	0.961 4
联合多序列模型	777	10	218	0	100	95.61	99.00	0.993 6	0.995 0

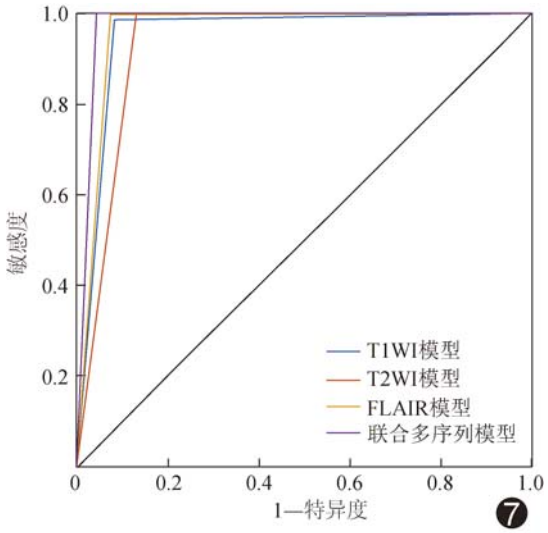

图 7　联合多序列模型及单一序列模型对测试集胶质瘤分级的 ROC 曲线

图像纹理及形状等特征以构建 SVM 分类模型，其分级胶质瘤的准确率达 98%。

深度学习是人工智能领域的主流算法，利用深度卷积神经网络（convolutional neural network，CNN）可自动提取深层图像特征，且性能好；经大量数据训练后，深度 CNN 具有良好效能。ZHUGE 等[6]采用常规 MRI 训练 2D Mask R-CNN 和 3D Mask R-CNN 以区分 HGG 与 LGG，其诊断准确率分别为 96.3% 和 97.1%。ÖZCAN 等[7]设计拥有 7 个线性结构卷积层的深度 CNN，以之训练模型区分 HGG 与 LGG，准确率为 97.1%。深度 CNN 用于医学影像学领域的主要问题在于难以获得带标签的训练数据，利用数据增强和迁移学习可在一定程度上加以解决。LUNDERVOLD 等[8]发现，通过迁移学习方法将以大量自然事物图像预先训练的深度学习模型用于特定领域，可在提高分类任务准确率的同时降低训练成本。本研究经过图像增强使用于训练单一序列模型的数据达 28 140 幅图像，可有效保证模型的收敛性；作为迁移源的 GoogLeNet 预训练网络系经包含约 150 万幅图像的 ImageNet 数据集训练而得，能精细理解自然事物的微小差异，只需微调即可适应胶质瘤分级任务。本研究对 GoogLeNet 预训练网络加以改进，在输出模块 3 个全连接层间添加随机丢失概率为 0.5 的 Dropout 层，以减少神经元共同适应的复杂度、促进各神经元融合；将批量尺寸（batchsize）设置为 128，以减少训练过程振荡和降低训练成本。

为使模型有针对性地提取胶质瘤的图像特征，有学者[5,7]在训练模型前手动提取胶质瘤 ROI，过程复杂，难以实现全流程自动化。刘征华[9]采用迁移学习方法基于头部增强 T1WI 分割的胶质瘤区域训练 Inception-v3 网络，分级胶质瘤准确率为 94.19%。本研究所用数据无须预分割胶质瘤 ROI，实现了全过程自动化；以平扫 MRI 作为数据来源使其适用性更佳，各单一序列模型分级测试集胶质瘤的准确率均超过 97.00%。

不同序列 MR 图像能提供不同的诊断信息，联合

多种信息能更精准地进行诊断。多数胶质瘤自动分级方法[10-11]仅采用单一 MR 序列图像，未充分利用 MR 多序列成像优势；即使增加机器学习特征数量或改进深度学习模型，亦难以提升其性能。本研究引入以多种 MR 序列模型分别进行胶质瘤自动分级的投票机制，可避免不同序列图像之间的干扰，有效利用多种序列图像包含的不同信息，降低单一序列模型对误分类的影响，最终使分级准确率达 99.00%，且操作简便，有助于临床医师诊断，并可为开发胶质瘤自动分级系统提供技术支持。

总之，联合多 MR 序列迁移学习网络胶质瘤自动分级算法过程简洁、效能高；向深度 CNN 特征提取部分中添加不同模块对胶质瘤分级模型性能的影响有待后续进一步观察。

［参考文献］

［1］ MENZE B H, JAKAB A, BAUER S, et al. The multimodal brain tumor image segmentation benchmark(BRATS)［J］. IEEE Trans Med Imaging, 2019, 34(10): 1993-2024.

［2］ BRUCH S, WANG X, BENDERSKY M, et al. An analysis of the softmax cross entropy loss for learning-to-rank with binary relevance［C］//Proceedings of the 2019 ACM SIGIR international conference, Paris, 2019. New York: Association for Computing Machinery, 2019: 75-78.

［3］ 贾颖，杜学松，陈君辉，等. 基于常规 MRI 的定量影像学特征用于胶质瘤分级诊断［J］. 中国医学影像技术，2018，34(8)：1137-1142.

［4］ TIAN Q, YAN L F, ZHANG X, et al. Radiomics strategy for glioma grading using texture features from multiparametric MRI［J］. J Magn Reson Imaging, 2018, 48(6): 1518-1528.

［5］ ZHANG Z, XIAO J, WU S, et al. Deep convolutional radiomic features on diffusion tensor images for classification of glioma grades［J］. J Digit Imaging, 2020, 33(4): 826-837.

［6］ ZHUGE Y, NING H, MATHEN P, et al. Automated glioma grading on conventional MRI images using deep convolutional neural networks［J］. Med Phys, 2020, 47(7): 3044-3053.

［7］ ÖZCAN H, EMIROĞLU B G, SABUNCUOĞLU H, et al. A comparative study for glioma classification using deep convolutional neural networks［J］. Math Biosci Eng, 2021, 18(2): 1550-1572.

［8］ LUNDERVOLD A S, LUNDERVOLD A. An overview of deep learning in medical imaging focusing on MRI［J］. Z Med Phys, 2019, 29(2): 102-127.

［9］ 刘征华. 基于深度学习卷积神经网络的胶质瘤级别分类研究［J］. 实用放射学杂志，2020，36(7)：1015-1018.

［10］ LI S, ZHENG Y, SUN W, et al. Glioma grading, molecular feature classification, and microstructural characterization using MR diffusional variance decomposition (DIVIDE) imaging［J］. Eur Radiol, 2021, 31(11): 8197-8207.

［11］ POGOSBEKIAN E L, PRONIN I N, ZAKHAROVA N E, et al. Feasibility of generalised diffusion kurtosis imaging approach for brain glioma grading［J］. Neuroradiology, 2021, 63(8): 1241-1251.

新书推荐：《肌骨系统疑难病例：影像诊断与鉴别诊断》

本书由武警浙江总队医院放射科张联合主任医师组织翻译。全书共 7 章，按照部位分别论述了肌肉骨骼系统疑难病例的影像诊断思路。本书内容全面，图文并茂，配有千余幅高清图像，可以指导读者如何阅片和进行分析、诊断及鉴别诊断，有助于将复杂疑难病例化繁为简，达到快速、准确诊断的目的。

天津科技翻译出版有限公司印刷，大 16 开，铜版纸，精装，定价 180 元。

联 系 人：姜晓婷
手 机 号：15022613568(微信号)
地　　址：天津市南开区白堤路 244 号
邮　　编：300192