BioMedInformatics：脑影像AI的“最后一公里”：如何让深度学习分割与分类更稳健、更通用？

作者：writer 来源：科学网

本文原地址：https://www.iikx.com/news/progress/43464.html

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

。论文标题： Strategies to Improve the Robustness and Generalizability of Deep Learning Segmentation and Classification in Neuroimaging

论文链接：https://www.mdpi.com/2673-7426/5/2/20

期刊名：BioMedInformatics

期刊主页：https://www.mdpi.com/journal/biomedinformatics

一、引言

机器学习与人工智能已革新医学影像工作流，应用于去噪、辐射剂量降低、图像重建、工作流优化，以及图像分析、计算机辅助诊断、报告生成和临床决策支持。医学影像中的主要下游任务为分割与分类。临床部署依赖模型在真实场景中的一致性能。稳健性指模型在医学影像环境变异下保持性能；泛化性指模型在全新、未见数据集上有效。变异来源包括扫描仪制造商与类型、采集协议、患者体位、伪影和噪声。缺乏稳健性会导致分类错误或分割边界不精确；泛化性不足则常表现为依赖训练数据中的虚假相关，即过拟合。为此，数据增强、对抗训练、迁移学习、域适应、逆监督学习等策略被用于提升模型表现。

二、方法

本文遵循结构化协议，检索 Google Scholar、PubMed 和 Scopus。关键词包括：brain images、neuroimaging、brain imaging；segmentation/classification 与 deep learning；robustness、generalizability。纳入同行评审、英文、脑成像原始研究；排除综述、重复研究及未直接涉及稳健性或泛化性的研究。研究应用 PRISMA
原则指导检索，最终纳入180篇。数据提取涵盖定义、训练使用和评估方法。

三、提升稳健性与泛化性的策略

共享方法包括：优化技术，如 Dice loss、加权交叉熵、Adam、L1/L2 正则化、Dropout、Batch Normalization、Early Stopping，以及 PCA、ICA、LASSO
等降维方法；数据增强，如几何变换、颜色空间调整、噪声注入、随机擦除、Mixup 和 CutMix；集成学习，如 Bagging、Boosting、Stacking 和 Voting；模型架构改进，如 U-Net、Attention U-Net、Vision Transformer 及 CNN-Transformer 混合模型。

稳健性提升方法包括对抗训练，如 FGSM、PGD、CW 攻击，以及 Min-Max 优化和 Wasserstein 鲁棒优化。泛化性提升方法包括域适应与不变学习、风格迁移、不变风险最小化、对比学习；训练策略包括迁移学习中的特征提取、微调、冻结层，联邦学习和自监督学习。

评估与监测方面，分割任务常用 IoU、DSC、Hausdorff 距离；分类任务常用 Accuracy、Precision、Recall、F1、AUC-ROC
和混淆矩阵。研究还强调统计显著性、95%置信区间、计算复杂度分析，以及 K 折、分层 K 折、留一、留一医院和嵌套交叉验证。表2总结了各策略优缺点，并指出复杂技术计算成本较高，可采用知识蒸馏、量化、剪枝和神经架构搜索等轻量化策略。

四、临床转化挑战

神经影像数据质量高度可变，伪影、运动、强度归一化和类别不平衡会影响模型表现。人群变异与跨站点泛化面临域偏移问题，联邦学习和迁移学习是潜在解决方案。分割任务中，小病灶或低对比病灶难以准确勾勒；分类任务中，模型可能依赖扫描仪特异性噪声等虚假相关。紧急场景下低质量扫描进一步降低可靠性。不确定性感知框架、贝叶斯神经网络和集成方法有助于提升临床可信度。

五、消融研究：非增强头CT颅内出血分割与分类

颅内出血危及生命，CT 是初始诊断金标准。研究评估 EDH、SDH、SAH、IVH、IPH 五类出血。数据增强、混合损失函数、正则化、5层分层交叉验证、集成学习、nnUNet 与注意力模型均提升分割和分类表现。对抗训练、域适应、迁移学习和自监督学习在该场景尚未充分应用，未来可能进一步提升稳健性。

六、讨论

稳健性与泛化性仍是神经影像深度学习的关键挑战。模型对扫描仪、噪声、伪影和协议差异敏感，域偏移、标签稀缺、类别不平衡和对抗攻击均影响临床可靠性。结合传统机器学习与领域知识、针对 MRI/CT/PET 等模态定制策略、探索 RAG、Hyperbolic CNN 和脑扫描不对称性等新方法值得关注。伦理问题包括隐私、知情同意、数据所有权、代表性、算法偏倚、标注完整性、透明性、公平性和可解释性。未来需发展透明、可靠、可适应的算法，推动多模态融合、联邦与分布式学习，并加强跨学科合作。

七、结论

稳健性与泛化性直接影响神经影像分割和分类的可信度、可靠性与临床适用性。本文综述了多种提升策略，但仍面临计算需求高、真实世界持续监测、医学图像不断演变等障碍。通过跨学科创新，神经影像人工智能有望更稳健、更通用，并产生广泛临床影响。

Tran, A.T.; Zeevi, T.; Payabvash, S. Strategies to Improve the Robustness and Generalizability of Deep Learning Segmentation and Classification in Neuroimaging. BioMedInformatics 2025, 5, 20. https://doi.org/10.3390/biomedinformatics5020020

期刊介绍

主编：Dr. Alexandre G. De Brevern, 法国巴黎大学红细胞生物学研究部

BioMedInformatics (ISSN 2673-7426) 是一个跨学科的科学开放获取期刊，其使命是推动面向生物学与医学计算科学所需的多学科研究与教育范式。本刊展示基于计算机科学及相关领域的方法与成果，旨在从生物医学数据中提取信息，并将这些信息转化为知识，从而促进数据驱动的科研范式。期刊重点关注针对生物医学问题的数据驱动型解决方案——包括创新性方法与计算工作流程的开发，其侧重点至少与利用这些方法获得的特定生物医学发现具有同等学术权重。

2025 Impact Factor：3.6

2025 CiteScore：7.1

Time to First Decision：20.8 Days

Acceptance to Publication：5.6 Days

特别声明：本文转载仅仅是出于传播信息的需要，并不意味着代表本网站观点或证实其内容的真实性；如其他媒体、网站或个人从本网站转载使用，须保留本网站注明的“来源”，并自负版权等法律责任；作者如果不希望被转载或者联系转载稿费等事宜，请与我们接洽。
来源：BioMedInformatics

更多 科学进展 请访问 https://www.iikx.com/news/progress/

本文版权归原作者所有，请勿用于商业用途，爱科学iikx.com转发