

基于YOLOv8改进的脑癌检测算法

王喆, 赵慧俊, 谭超, 李骏, 申冲

引用本文

王喆, 赵慧俊, 谭超, 李骏, 申冲. [基于YOLOv8改进的脑癌检测算法](#)[J]. 计算机科学, 2024, 51(11A): 231100122-7.

WANG Zhe, ZHAO Huijun, TAN Chao, LI Jun, SHEN Chong. [Enhanced Brain Cancer Detection Algorithm Based on YOLOv8](#) [J]. Computer Science, 2024, 51(11A): 231100122-7.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[PS-YOLOv8:增强电力线路检测中的小规模损坏检测](#)

PS YOLOv8:Enhancing Detection of Small-scale Damage in Power Lines Inspection
计算机科学, 2024, 51(11A): 240100003-6. <https://doi.org/10.11896/jsjx.240100003>

[基于改进Yolov8的敦煌壁画元素检测算法](#)

Dunhuang Mural Element Detection Algorithm Based on Improved Yolov8
计算机科学, 2024, 51(11A): 231000034-6. <https://doi.org/10.11896/jsjx.231000034>

[基于Light-YOLOv8的围棋棋谱识别](#)

Go Chessboard Recognition Based on Light-YOLOv8
计算机科学, 2024, 51(11A): 230900037-7. <https://doi.org/10.11896/jsjx.230900037>

[面向汉字点选验证码的轻量级高效识别方法](#)

Lightweight and Efficient Recognition Method for Chinese Character Click-based CAPTCHA
计算机科学, 2024, 51(11A): 240100031-9. <https://doi.org/10.11896/jsjx.240100031>

[基于可见光-红外跨域迁移的红外弱小目标检测](#)

Infrared Dim and Small Target Detection Based on Cross-domain Migration of Visible Light and Infrared
计算机科学, 2024, 51(10): 287-294. <https://doi.org/10.11896/jsjx.230800013>

基于 YOLOv8 改进的脑癌检测算法

王喆¹ 赵慧俊^{2,3} 谭超¹ 李骏¹ 申冲^{2,3}

1 中北大学机械工程学院 太原 030051

2 中北大学仪器与电子学院 太原 030051

3 山西省省部共建动态测试技术国家重点实验室 太原 030051

(635087094@qq.com)

摘要 自动检测脑部肿瘤在磁共振成像中的位置是一个复杂、繁重的任务,需要耗费大量时间和资源。传统识别方案经常出现误解、遗漏和误导的情况,从而影响患者的治疗进度,对患者的生命安全产生影响。为了进一步提高鉴定的效果,引入了4项关键改进措施。首先,采用了高效的多尺度注意力 EMA(Efficient Multi-scale Attention),这种方法既可以对全局信息进行编码,也可以对信息进行重新校准,同时通过并行的分支输出特征进行跨维度的交互,使信息进一步聚合。其次,引入了 BiFPN(Bidirectional Feature Pyramid Network)模块,并对其结构进行改进,以便缩短每一次检测所需要的时间,同时提升图像识别效果。然后采用 MDPIoU 损失函数和 Mish 激活函数进行改进,进一步提高检测的准确度。最后进行仿真实验,实验结果表明,改进的 YOLOv8 算法在脑癌检测中的精确率、召回率、平均精度均值均有提升,其中 Precision 提高了 4.48%,Recall 提高了 2.64%,mAP@0.5 提高了 2.6%,mAP@0.5:0.9 提高了 7.0%。

关键词: YOLOv8; 脑癌; Efficient Multi-Scale Attention 模块; Bidirectional Feature Pyramid Network 结构; Missed Softplus with Identity Shortcut 激活函数; Minimum Point Distance Intersection over Union 损失函数

中图分类号 TP391.4

Enhanced Brain Cancer Detection Algorithm Based on YOLOv8

WANG Zhe¹, ZHAO Huijun^{2,3}, TAN Chao¹, LI Jun¹ and SHEN Chong^{2,3}

1 School of Mechanical Engineering, North University of China, Taiyuan 030051, China

2 School of Instruments and Electronics, North University of China, Taiyuan 030051, China

3 State Key Laboratory of Dynamic Measurement Technology of Shanxi Province, Taiyuan 030051, China

Abstract Automatically detecting the location of brain tumors in magnetic resonance imaging is a complex and laborious task that requires a lot of time and resources. Traditional identification schemes often misunderstand, omit, and mislead, affecting the progress of patient treatment and posing a risk to patient safety. To further improve the identification and appraisal results, this paper introduces four key improvement measures. Firstly, an efficient multi-scale attention EMA is adopted, which can encode global information, recalibrate information, and further aggregate information through parallel branch output features for cross-dimensional interaction. Secondly, the BiFPN(Bidirectional Feature Pyramid Network) module is introduced to shorten the time required for each detection and improve image recognition results. Then, the MDPIoU loss function and Mish activation function are improved to further enhance detection accuracy. Finally, simulation experiments are conducted for verification. The experimental results show that the improved YOLOv8 algorithm has improved precision, recall, and mean average precision(mAP) in brain cancer detection. Among them, precision, recall, mAP@0.5 and mAP@0.5:0.9 increases by 4.48%, 2.64%, 2.6% and 17.0% respectively.

Keywords YOLOv8, Brain Tumor, Efficient Multi-scale attention model, Bidirectional Feature Pyramid Network, Missed Softplus with Identify Shortcut activation function, Minimum Point Distance Intersection over Union loss function

1 引言

脑癌是一种常见的中枢神经系统恶性肿瘤,可广泛分布于该系统的各个解剖区域,主要发生于大脑(约占 90%以上),其余可发生于脑膜、脊髓和颅神经等部位。脑癌以其高死亡率和导致患者严重致残的特点而著称,对卫生保健系统造成了不可忽视的负担。研究表明,脑癌患者的存活率与早

期发现脑癌的时间呈正相关。因此,及时发现脑癌的位置、大小和数量对于保障患者的生命安全至关重要。在医疗诊断脑癌时,常采用磁共振成像(MRI)的方法。磁共振成像能够准确显示脑癌的位置、数量和大小。然而,人工进行 MRI 影像识别可能存在漏识别或误识别,对患者的健康和安全构成潜在威胁。因此,将图像识别技术用于脑癌检测具有重要意义,可以有效预防误判、漏判和错判,提高诊断的准确性与可靠性。

基金项目:光电信息控制和安全技术重点实验室基金(2021JCJQLB055010)

This work was supported by the Key Laboratory of Optoelectronic Information Control and Security Technology Fund(2021JCJQLB055010).

通信作者:申冲(shenchong@nuc.edu.cn)

随着基于深度学习的目标检测技术迅速发展,在医学领域,很多图像检测都得到广泛的应用,极大地推动了医疗领域的发展,帮助了更多的患者康复。目前 MRI 脑瘤图像目标检测算法主要可分为两类:一类侧重于应用神经网络对图像进行分割以定位肿瘤,然而这种算法复杂繁琐,可能存在遗漏和误识别,并且检测时间较长;另一类研究采用 CNN 网络对脑癌患者的图像进行检测,这种方法能够在一定程度上提高检测速度和效率,缩短诊断时间,而且识别效果优于传统的分割算法。

目前许多学者对脑癌检测有一定的研究,且取得了不错的效果。其中 He 等^[1]在原始的 3D-Unet 的基础上,以由 IN 层、PreLU 层和 3 组膨胀率分别为 1, 2, 3 的膨胀卷积层共同构成的混合膨胀卷积模块代替常规卷积模块,用 Dice 损失和 Focal 损失结合的损失函数代替原先的 Dice 损失函数,从而提高检测的精度。

Moinul 等^[2]首先用 CNN 模型建构对模型进行模型训练和结果分析,之后采用效果最好的 3 个训练后的 CNN 模型的平均值来制作模型,随后应用联合学习的方法来检测脑癌的位置。Kader 等^[3]采用差分深度卷积神经网络模型,之后通过使用预定义的超活跃值和差分因子,在差分 CNN 网络中生成特征图,来提取大脑 MRI 图像中的更多细节,从而分类正常和异常的脑癌图像。Kang 等^[4]在 YOLOv7 中加入 VGG,通过重新参数化来加速推理时间。通过更改 SPP 模块在颈部网络中最大池化不同卷积核的大小,从而实现局部特征和全局特征的融合。Hossian 等^[5]使用 Leaky ReLU 激活函数代替原先的 ReLU 激活函数,即用(Conv 2D + Leaky ReLU)层组合替换掉原先的(Conv 2D + ReLU)层组合,实现了精度的提升。

然而,在实际应用中利用 CNN 模型对 MRI 脑瘤图片进行检测存在若干问题:1)获取脑部 MRI 的脑癌区域较整个图像小,而背景区域明显多于目标区域,这一类别不平衡问题增加了分割难度;2)传统的网络结构限制了对图像信息的充分

利用和特征提取能力,从而制约了训练模型分割精度的提升;3)随着网络深度增加,学习效率下降,进而影响了分割准确率的有效提升。尽管部分研究专注于各类卷积神经网络(CNN)的应用^[6-8],但使用 YOLO 神经网络模型进行脑癌检测的研究相对较少。此外,在 CNN 模型中,复杂的网络架构如 ResNet^[9],Inception^[10],DenseNet^[11],MobileNet^[12]等得到了广泛探索。然而,这些方案有些未达到预期的检测精度标准,或因检测速度过慢而受到限制。

相较于传统 CNN 网络模型,本文改进模型 YOLOv8-BT 在检测速度上得到显著的提升。相较于 Faster R-CNN,其在单张图片上的识别速度提升了 13.1 ms。相较于 YOLOv8s 网络,其检测精度也有显著的提升,在检测小目标脑癌时可以减少误识别,并且在 Precision, Recall, mAP@0.5, mAP@0.5:0.9 指标上都有着显著的提升。

本文研究主要包括以下方面:

- 1)使用 EMA 注意力模块^[13],提升主干网络特征提取能力,以便识别小面积的脑癌。
- 2)增加 BiFPN 结构^[14],并对其结构进行改造,从而提高了模型性能,改善了原网络提取深度图像特征的能力,使之适用于处理复杂机器学习和深度学习任务。
- 3)将定位损失函数改进为 MPDIoU^[15],提升算法的收敛性和训练效率。
- 4)将 YOLOv8 中卷积模块的激活函数改进为 Mish^[16],从而更好地捕获数据中存在的非线性关系,提升模型的准确度和泛化能力。

本文第 2 章介绍 YOLOv8 算法原理;第 3 章介绍改进算法的原理;第 4 章展示实验效果和性能评估结果;最后总结全文并展望未来。

2 YOLOv8 算法原理

YOLOv8 网络结构如图 1 所示,结构包含了 3 个部分,分别是:主干网络(Backbone),颈部(Neck),预测头(Head)^[17]。

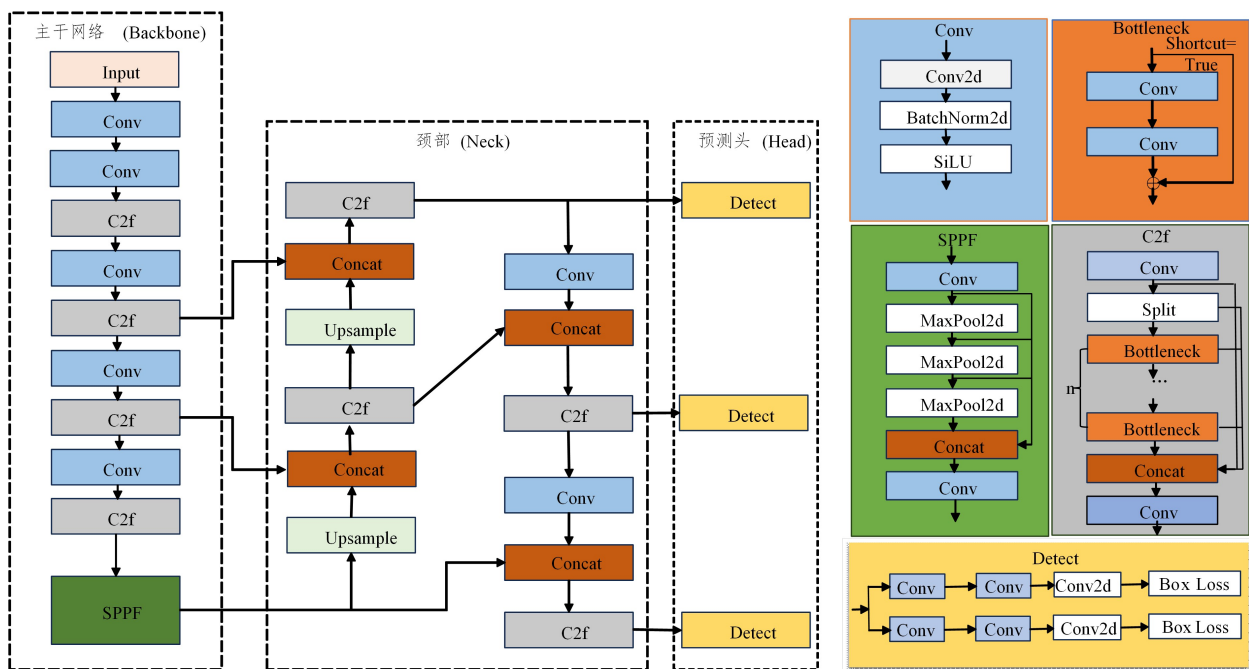


图 1 YOLOv8 网络结构示意图

Fig. 1 Schematic diagram of YOLOv8 network structure

主干网络的主要作用是对输入的图像进行一系列的特征提取,为后续的处理和分析做铺垫。YOLOv8 将主干网络中的第一层卷积从原先的 6×6 卷积核改为 3×3 卷积核,帮助降低模型的参数数量,从而降低计算成本,提高模型的训练速度。并且, 3×3 卷积核通常具有更好的非线性特性,有助于更好地捕捉数据中的复杂特征。此外, YOLOv8 借鉴了 YOLOv7 ELAN (Efficient and Light-weight Attention Network) 的设计思想,将传统的 C3 模块替换为 C2f 模块。C2f 模块采用了密集残差结构,增强了特征表达能力。通过拆分和拼接操作, C2f 模块还能够有效地改变通道数,这有助于降低计算成本、内存需求和模型复杂性^[18-19]。

颈部操作是对主干网络提取出的特征进行降维或调整,以便更好地应用于工程中。YOLOv8 用 C2f 结构取代 YOLOv5 的 C3 结构增加了跨层连接和 Split 操作,增强了算法的梯度流和特征融合,并且去掉了颈部模块中的两个卷积连接层,从而减少了参数数量和计算量。

3 YOLOV8 算法网络改进

3.1 注意力模块改进

与人眼识别物体时的过程相似,计算机视觉模型也使用类似的注意力机制。首先,模型对图像进行快速扫描,确定潜在关注区域。然后将更多的计算资源集中在这些区域,以便获得更多的目标细节信息,并有效地忽略掉不太有用的特征。这种方式有助于提高模型的检测效率和准确性,使其更好地处理复杂的视觉任务。

在进行脑瘤位置检测时,加入 EMA 注意力模块可以使网络不漏掉面积小的脑瘤,或将多个小的脑瘤识别到一起,从而影响医生的判断,进而耽误患者的治疗。因此本文将 EMA 嵌入到 YOLOv8 主干网络的末端,使网络进行有效的通道描述,在卷积运算中无需通道降维,并且使得高级特征图对微小信息的关注更佳,避免遗漏掉小面积的脑瘤信息。EMA 结构如图 2 所示,主要包含 Feature Grouping, Parallel Subnetworks, Cross-spatial learning 部分。

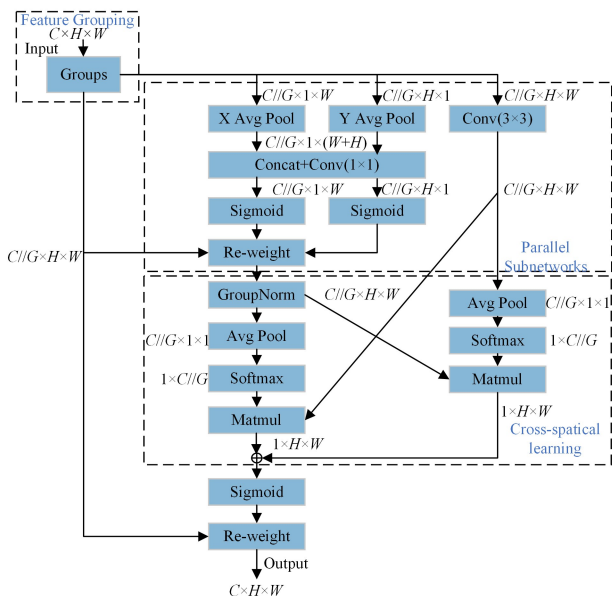


图 2 EMA 注意力模块结构图

Fig. 2 Structure diagram of EMA attention module

3.2 特征金字塔改进

为了避免在复杂场景下出现脑癌的误检问题,我们将改进的 Bi-FPN 添加到特征金字塔中。BiFPN (Bidirectional Feature Pyramid Network) 是一种加权双向特征金字塔,其主要功能在于双向跨尺度连接和加权特征图融合。双向跨尺度连接可以在没有大量计算资源消耗下,提取更多的特征。将 BiFPN 视作一个基本单元,即 BiFPN 中一对路径视为一个特征层,然后重复多次从而得到更多高层数据融合,其结构如图 3 所示。

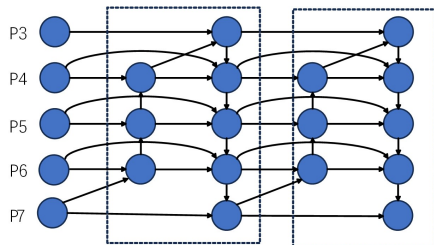


图 3 BiFPN 结构图

Fig. 3 Structures of BiFPN

原始的 BiFPN 有 5 次连续下采样操作。为提升检测效率且不影响检测效果,改进的 BiFPN 模型删除了 BiFPN 中的 P7 通道的输入和输出的残差链接,只对 P3, P4, P5, P6 这 4 个检测通道进行特征融合,其结构如图 4 所示。这种改进措施在加快检测速度的同时也保证了检测的效率,改进后的网络模型也有效地避免了漏检脑癌的问题。

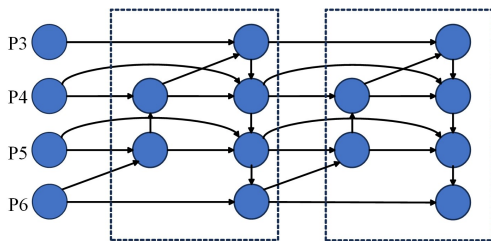


图 4 改进后的 BiFPN 结构图

Fig. 4 Improved BiFPN structure diagram

3.3 定位损失函数改进

损失函数的功能是度量神经网络预测信息之间的距离,它是决定网络效果的重要参数之一。YOLOv8 算法的损失函数主要由两个部分组成,分别是类别分类损失和定位损失。定位回归损失功能是衡量预测的边框和真实的边框之间的差异,其可以优化边框的位置和大小,使其更接近真实的目标。YOLOv8 定位损失函数采用 CIoU loss^[20] (Complete Intersection over Union Loss),表达式如式(1)所示:

$$L_{CIoU} = 1 - CIoU = 1 - (IoU - \mathcal{R}_{CIoU}) \quad (1)$$

其中, IoU 是交并比,用于衡量两个矩形框之间的相似度的指标。 \mathcal{R}_{CIoU} 是惩罚项,表达式如式(2)所示:

$$\mathcal{R}_{CIoU} = \frac{\rho^2}{c^2} + \alpha v \quad (2)$$

其中, ρ 是两个矩形框的中心点之间的欧氏距离, c 是两个矩形框的闭包区域的对角线的距离, v 是两个矩形框相对比例的一致性。 α 是权重系数,表达式如式(3)所示:

$$\alpha = \frac{v}{(1 - IoU) + v} \quad (3)$$

4.2 实验参数设置

本文模型训练时采取 Mosaic 数据增强来增加数据多样性和模型鲁棒性,以便处理复杂和多样性的数据。学习率衰减的策略为余弦退火,以防学习率波动过大,从而导致模型无法收敛。输入图像尺寸为 $320 \times 320 \times 3$; batch-size 设置为 32,即每次输入 32 张图像;epoch 设置为 300 轮;训练时学习率(learning rate)设置为 0.01,以达到稳定整个网络的效果。

4.3 数据集构建和模型训练

在脑癌检测研究中,构建高质量训练数据集和准确标注原始图像是模型训练的关键要素。数据集的质量和合理标注直接影响目标检测模型的性能。本文选择了 Kaggle 平台上的一个开源数据集,该数据集包含了 1 344 张脑部图像,以及研究院标注的脑癌位置信息。这些图像的尺寸大约为 248×338 。按照比例随机将这些图像划分为训练集和测试集,为后续模型训练奠定基础。

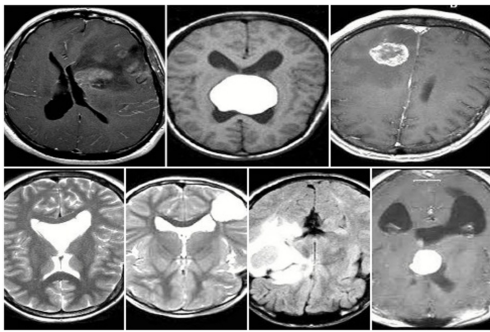


图 7 部分脑癌数据集图

Fig. 7 Partial brain cancer dataset images

随后使用 labelImg 对训练集图像的脑癌位置进行标注,如图 8 所示,并生成对应的 txt 文件,其显示标注脑癌的个数和位置信息。最后将标注好的 txt 文件放到 label 中,准备进行模型训练。



图 8 LabelImg 对数据集标注

Fig. 8 LabelImg for dataset annotation

4.4 评价指标

为验证改进后模型的精度和运行速度,本文选用精确率(Precision)、召回率(Recall)、平均精度均值(mean Average

Precision, mAP)、每张图片检测时间作为评估性能指标。mAP@0.5 表示 IoU 设为 0.5 的平均准确度,mAP@0.5:0.9 表示 IoU 从 0.5 到 0.9、步长为 0.05 时的平均精确度。

精确率和召回率的表达式如式(9)和式(10)所示:

$$Precision = \frac{N_{TP}}{N_{TP} + N_{FP}} \times 100\% \quad (9)$$

$$Recall = \frac{N_{TP}}{N_{TP} + N_{FN}} \times 100\% \quad (10)$$

其中, N_{TP} 表示检测模型将磁共振成像的脑癌目标检测为 brain_tumor 的数量, N_{FP} 表示误将正常区域或背景检测为 brain_tumor 的数量, N_{FN} 表示误将脑癌区域识别为 no_tumor 的数量。

平均精度 AP 可以通过 Precision 和 Recall 围成的曲线来进行判断。以 Precision 为纵轴,Recall 为横轴描绘成的曲线通常称为 P-R 曲线,该曲线与坐标轴围成的面积大小就是 AP 的值。对所有检测到的 AP 进行取平均值就可以得到平均精度均值 mAP,其表达式如式(11)和式(12)所示:

$$AP = \sum_{i=1}^N P(i) \Delta R(i) = \int_0^1 P(R) dR \quad (11)$$

$$mAP = \frac{\sum_{i=1}^N AP_i}{N} \quad (12)$$

其中, P 代表 Precision, R 代表 Recall, N 代表总类数。平均精度均值(mAP)越大,说明目标检测模型的性能越好,能够更准确地检测出目标的位置和类别。

4.5 消融实验

为验证 YOLOv8 各项改进对于网络整体性能的影响,本文以原始 YOLOv8 为基准进行消融实验。具体实验结果如表 1 所列,P-R 曲线如图 9 所示。

分析表 1 可知,相较于原始的 YOLOv8,添加 EMA 注意力模块后 Precision,Recall,mAP@0.5,mAP@0.5:0.9 这 4 个指标分别提高了 2.78%,1.65%,2.4%,6.5%,表明新增的注意力模块可以提高主干网络特征提取能力,以便识别小面积的脑癌。将 YOLOv8 特征金字塔替换为加权双向金字塔 BiFPN 后,Precision,Recall,mAP@0.5,mAP@0.5:0.9 分别提高了 1.90%,1.34%,1.3%,1.6%,同时检测时间也缩短了,每张图片检测时间 T 相较于传统 YOLOv8 模型减少了 7.2 ms。将定位损失函数改进为 MPDIoU,从而提升算法的收敛性和训练效率;将 YOLOv8 中卷积模块的激活函数改进为 Mish,从而更好地捕获数据中存在的非线性关系,提升模型的准确度和泛化能力,这两项改进也提高了算法检测的精度。实验结果表明,YOLOv8-BT 改进算法 Precision 提高了 4.48%,Recall 提高了 2.64%,mAP@0.5 提高了 2.6%,mAP@0.5:0.9 提高了 7.0%,同时由于 BiFPN 模块的作用,其在检测时间上相较于 YOLOv8+EMA 缩短了 17.2 ms,相较于 YOLOv8 依旧增多了 2.8 ms。

表 1 YOLOv8-BT 模型消融实验

Table 1 Ablation experiment of YOLOv8-BT model

YOLOv8-s	EMA	BiFPN	MPDIoU	Mish	P/%	R/%	mAP@0.5/%	mAP@0.5:0.9/%	T/ms
✓					89.64	90.12	92.1	66.0	37.3
✓	✓				92.42	91.77	94.5	72.5	57.3
✓		✓			91.54	91.46	93.4	67.6	30.1
✓			✓		90.24	90.23	94.1	69.5	40.5
✓				✓	89.51	90.05	92.3	67.6	41.2
✓	✓	✓	✓	✓	94.12	92.76	94.7	73.0	40.1

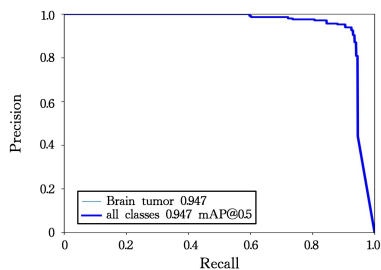


图9 YOLOv8-BT模型的P-R曲线图

Fig. 9 P-R curve of YOLOv8-BT model

4.6 检测结果对比

为验证 YOLOv8-BT 模型检测效果的提升,本文分别使用 YOLOv8 和 YOLOv8-BT 模型对磁共振成像的脑肿瘤进行检测,部分检测结果对比如图 10 所示。

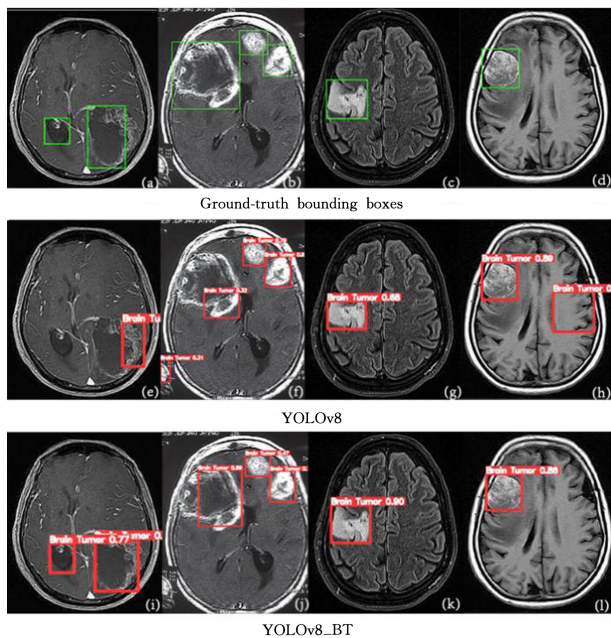


图 10 两种模型检测效果的对比

Fig. 10 Comparison of detection performance of two models

其中,(a)–(d)为 Ground-Truth Bounding Boxes 设置;(e)–(h) YOLOv8 检测效果;(i)–(l) YOLOv8-BT 检测效果。分别对图 10(e)、图 10(i)进行比较, YOLOv8 无法同时检测到两个大小不一的脑癌,改进后的 YOLOv8-BT 可以将两个面积大小不同的脑癌识别成 brain_tumor 类。对图 10(f)、图 10(j)进行比较,可见 YOLOv8 模型存在误检测,其将背景识别为 brain_tumor 类,并且没有将大面积脑癌完全识别出来。改进后的 YOLOv8-BT 没有误识别,并将 3 个脑癌全部识别出来。对图 10(g)、图 10(k)进行对比,可见对于中等面积的脑癌识别, YOLOv8-BT 模型提高了对 brain_tumor 类识别的置信度,同时 YOLOv8-BT 的置信度高于 YOLOv8 模型。对图 10(h)、图 10(l)进行对比, YOLOv8 模型将正常部分误识别为 brain_tumor 类, YOLOv8-BT 模型没有将大脑健康的区域识别为脑癌区域,且置信度大于 YOLOv8 模型。

4.7 深度学习目标检测算法对比实验

为了进一步验证 YOLOv8-BT 在脑癌检测方面的优势,本文将选用 4 种主流的目标检测模型的不同版本进行对比实验,分别是 Faster R-CNN, YOLOv5-s, YOLOv6-s, YOLOv7-

tiny, RCSVGG-CSP。在相同的实验环境下,使用相同的数据集、训练策略和训练参数,本文将 YOLOv8-BT 与 Faster R-CNN, YOLOv5-s, YOLOv6-s, YOLOv7-tiny, RCSVGG-CSP 目标检测算法进行比较,实验结果如表 2 所列。

表 2 不同方法实验结果对比

Table 2 Experimental results comparison of different methods

Model	P/%	R/%	mAP@0.5/%	mAP@0.5:0.9/%	T/ms
Faster R-CNN	67.31	75.31	75.4	42.2	53.2
YOLOv5-s	88.54	89.55	89.9	54.2	37.6
YOLOv6-s	91.32	90.22	91.2	55.5	38.4
YOLOv7-tiny	92.11	90.49	92.1	58.6	37.7
RCSVGG-CSP	91.2	91.5	93.3	68.9	45.5
YOLOv8-BT	94.12	92.76	94.7	73.0	40.1

从表 2 可以看出:在脑癌检测中, Faster R-CNN 算法在检测性能上表现欠佳,且每张图片平均检测时间较长。 YOLOv5-s 虽然在单个图片的检测时间上相较于 YOLOv8-BT 缩短了 2.5 ms,但检测性能降低了不少。与 Faster R-CNN, YOLOv5-s, YOLOv6-s, YOLOv7-tiny, RCSVGG-CSP^[5] 相比, YOLOv8-BT 在 Precision, Recall, mAP@0.5, mAP@0.5:0.9 这 4 个指标上都有着显著的提升。因此证明 YOLOv8-BT 可以提高检测脑癌的精度。

结束语 为防止磁共振成像的图片的脑癌误检测,或因面积太小而没有检测到的问题,本文构建了一种 YOLOv8-BT 目标检测框架。首先在主干网络上引入 EMA 注意力模块,从而在检测时不会漏掉小面积脑癌的识别,并且避免将多个相近的脑癌共同识别为一个脑癌。之后本文引入 BiFPN 结构对特征金字塔进行改进,以便提高识别准确率和缩短每次的识别时间。将定位损失函数改进为 MPDIoU,提升算法的收敛性和训练效率,防止在训练过程中出现过拟合,提高模型的稳定性。将 YOLOv8 中卷积模块的激活函数改进为 Mish,从而更好地捕获数据中存在的非线性关系,提升模型的准确度和泛化能力。在引入这些改进后,相较于原始的 YOLOv8 网络,改进后的 YOLOv8-BT 的 Precision 提高了 4.48%, Recall 提高了 2.64%, mAP@0.5 提高了 2.6%, mAP@0.5:0.9 提高了 7.0%,可见其在识别准确度上有着显著的提升,引入的 BiFPN 结构改进也将检测时间缩短了 2.8 ms。但从结果上显示,检测时间仍然相对较长,之后的主要研究方向是继续降低检测时间来提高效率。

参考文献

- [1] HE C E, XU H J, WANG Z, et al. Automatic Segmentation Algorithm for Multimodal Magnetic Resonance-Based Brain Tumor Images[J]. Acta Optica Sinica, 2020, 40(6): 0610001.
- [2] MOINUL I, TANZIM M R, MOHAMMED K, et al. Effectiveness of Federated Learning and CNN Ensemble Architectures for Identifying Brain Tumors Using MRI Images[J]. Neural processing letters, 2022, 55(4): 3379-3809.
- [3] Abd El KADER, ISSELMOU, et al. Differential deep convolutional neural network model for brain tumor classification[J]. Brain Sciences, 2021, 11(3): 352.
- [4] KANG M, TING C M, TING F F, et al. RCS-YOLO: A Fast and High-Accuracy Object Detector for Brain Tumor Detection[J].

arXiv:2307.16412,2023.

- [5] HOSSAIN T, SHISHIR F S, ASHRAF M, et al. Brain tumor detection using convolutional neural network[C]//1st international Conference on Advances in Science, Engineering and Robotics Technology(ICASERT 2019). IEEE, 2019.
- [6] VANKDOTHU R, HAMEED M A, FATIMA H. A brain tumor identification and classification using deep learning based on CNN-LSTM method[J]. Computers and Electrical Engineering, 2022, 101:107960.
- [7] TIWARI P, PANT B, ELARABAWY M M, et al. Cnn based multiclass brain tumor detection using medical imaging [J]. Computational Intelligence and Neuroscience, 2022.
- [8] MONTAHA S, AZAM S, RAFID A K M R H, et al. Timedistributed-cnn-lstm: A hybrid approach combining cnn and lstm to classify brain tumor on 3d mri scans performing ablation study [J]. IEEE Access, 2022, 10:60039-60059.
- [9] ARG S, ALMEIDA D, LYMAN K. Resnet in resnet: Generalizing residual architectures[J]. arXiv:1603.08029, 2016.
- [10] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016:2818-2826.
- [11] ZHU Y, NEWSAM S. Densenet for dense flow[C]// 2017 IEEE International Conference on Image Processing (ICIP). IEEE, 2017:790-794.
- [12] SINHA D, EL-SHARKAWY M. Thin mobilenet: An enhanced mobilenet architecture[C]// 2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON). IEEE, 2019:0280-0285.
- [13] OUYANG D, HE S, ZHANG G, et al. Efficient Multi-Scale Attention Module with Cross-Spatial Learning[C]// ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2023:1-5.
- [14] TAN M, PANG R, LE Q V. Efficientdet: Scalable and efficient object detection[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020:10781-10790.
- [15] SILIANG M, YONG X. MPDIoU: A Loss for Efficient and Accurate Bounding Box Regression[J]. arXiv:2307.07662, 2023.
- [16] MISRA D. Mish: A self regularized non-monotonic activation function[J]. arXiv:1908.08681, 2019.
- [17] ABOAH A, WANG B, BAGCI U, et al. Real-time multi-class helmet violation detection using few-shot data sampling technique and yolov8[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023:5349-5357.
- [18] ZHANG Y, GUO Z, WU J, et al. Real-time vehicle detection based on improved yolo v5 [J]. Sustainability, 2022, 14 (19): 12274.
- [19] WANG Y, WANG H, XIN Z. Efficient detection model of steel strip surface defects based on YOLO-V7 [J]. IEEE Access, 2022, 10:133936-133944.
- [20] XUE J, CHENG F, LI Y, et al. Detection of farmland obstacles based on an improved YOLOv5s algorithm by using Ciou and anchor box scale clustering[J]. Sensors, 2022, 22(5):1790.



WANG Zhe, born in 2003, undergraduate, is a member of CCF (No. Q2550G). His main research interests include machine learning and autopilot.



SHEN Chong, born in 1986, Ph.D supervisor, professor. His main research interests include artificial intelligence and bionic navigation.