

基于改进 YOLOv8 的脑肿瘤图像检测算法

郑泽毅¹, 曹嘉璇¹, 王家琪¹, 邹北骥², 郭纯³, 刘青萍^{1†}

(1. 湖南中医药大学信息科学与工程学院, 湖南 长沙 410208; 2. 中南大学计算机学院, 湖南 长沙 410083;

3. 湖南中医药大学第一附属医院, 湖南 长沙 410000)

摘 要:脑肿瘤是严重威胁人类健康的疾病, 早期检测对提高治疗效果至关重要。针对现有脑肿瘤检测算法在复杂背景或边缘模糊的情况下检测精度不足的问题, 本文提出了一种基于改进 YOLOv8 的脑肿瘤检测算法。该算法引入协调注意力机制, 强化肿瘤区域的特征聚焦; 在主干网络中采用深度可分离卷积, 降低计算复杂度并提升特征提取效率; 结合 Bottleneck Transformer 模块, 增强了全局信息建模能力。实验在 Brain Tumor Detection 数据集上进行, 结果显示检测精度达到 93%, 相较于原算法提升 1.1%, mAP0.5 和 mAP0.5 : 0.95 分别提升了 2.1% 和 1.7%。实验结果表明, 改进算法在脑肿瘤检测任务中表现出显著优势, 为医学影像辅助诊断提供了更加精准和高效的支持。

关键词:脑肿瘤检测; YOLOv8; 协调注意力机制; 深度可分离卷积; Bottleneck Transformer

中图分类号:TP391

文献标识码:A

Brain Tumor Image Detection Algorithm Based on Improved YOLOv8

ZHENG Zeyi¹, CAO Jiaxuan¹, WANG Jiaqi¹, ZOU Beiji², GUO Chun³, LIU Qingping^{1†}

(1. School of Informatics, Hunan University of Chinese Medicine, Changsha, Hunan 410208, China;

2. School of Computer Science and Engineering, Central South University, Changsha, Hunan 410083, China;

3. The First Hospital of Hunan University of Chinese Medicine, Changsha, Hunan 410000, China)

Abstract: Brain tumors are a severe threat to human health, and early detection is crucial for improving treatment outcomes. To address the issue of insufficient detection accuracy in existing brain tumor detection algorithms under complex backgrounds or with blurred tumor boundaries, this paper proposes an improved YOLOv8-based brain tumor detection algorithm. The algorithm introduces a coordinate attention mechanism to enhance feature focus on tumor regions; employs depthwise separable convolution in the backbone network to reduce computational complexity and improve feature extraction efficiency; and incorporates the bottleneck transformer module to strengthen global information modeling. Experiments were conducted on the brain tumor detection dataset, and the results show that the detection accuracy reached 93%, an improvement of 1.1% over the original algorithm. mAP0.5 and mAP0.5 : 0.95 were increased by 2.1% and 1.7%, respectively. The experimental results demonstrate that the improved algorithm exhibits significant advantages in brain tumor detection tasks, providing more accurate and efficient support for medical image-assisted diagnosis.

Key words: brain tumor detection; YOLOv8; coordinate attention; depthwise separable convolution; Bottleneck Transformer

收稿日期: 2024-12-16

基金项目: 国家重大科技专项项目(2018AAA0102100); 2023 年湖南省学位与研究生教学改革研究项目(2023JGYB151); 湖南中医药大学第一附属医院中医药传承创新专项重点项目(2024XYLH346); 湖南省普通高等学校教学改革研究项目(HNJG-20230527); 2024 年湖南中医药大学研究生科研创新项目(2024CX192)

作者简介: 郑泽毅(2001—), 男, 福建三明人, 硕士研究生, 研究方向: 医学图像处理, 深度学习。

† 通信联系人, E-mail: liuliu@hnuem.edu.cn

脑肿瘤是一种细胞在脑内或脑周围异常增殖的疾病,可能是良性,也可能是恶性^[1]。良性肿瘤如脑膜瘤通常生长缓慢且不侵袭周围组织,而恶性肿瘤如胶质母细胞瘤则生长迅速,具有极强的侵袭性。这些肿瘤会直接压迫脑功能,导致一系列症状,如颅内压升高、头痛、视力模糊以及认知和情绪变化等。持续存在或复发的肿瘤可能引发永久性神经损伤,甚至威胁生命。不同类型的脑肿瘤具有独特的生物学特性、治疗反应和并发症,有效预防和治疗脑肿瘤的关键在于深入了解其各个阶段。近年来,放射科医生利用多种成像技术对脑肿瘤进行检查,包括 X 射线、脑磁图(MEG)、计算机断层扫描(CT)、超声、脑电图(EEG)和磁共振成像(MRI)等技术,提升疾病诊断的准确性,并辅助制定适合的治疗方案。由于原发性脑肿瘤在大小、位置等方面差异巨大,早期诊断面临诸多挑战,通常会采用多种治疗策略组合,包括手术、放射治疗和化疗等方法,以尽可能切除肿瘤组织并抑制其生长。因此,早期检测和精确定位对脑肿瘤的诊断和治疗至关重要。

近年来,随着深度学习技术的发展,许多目标检测算法^[2]在医学图像中得到了广泛应用,为脑肿瘤的自动检测与定位提供了新的方法。由 Ren 等^[3]提出的 Faster-RCNN 模型是一种两阶段检测方法,首先会生成候选区域,然后基于这些区域进行精细的检测和分类。这种分阶段策略使得 Faster-RCNN 在精度方面具有优势,特别适用于需要高检测精度的场景。然而,由于其候选区域生成和精细分类的过程较为耗时,该模型在实时检测的应用中会受到一定限制。随后,Liu 等^[4]提出的单阶段检测网络(Single Shot MultiBox Detector, SSD),通过在不同尺度的特征层上同时进行检测,实现了多尺度的目标检测。SSD 的网络结构简化了检测流程,降低了计算复杂度。与两阶段模型相比,加快了检测速度,但在小目标检测或对精度要求较高的任务中表现受限。针对上述问题,由 Redmon 等^[5]提出的 You Only Look Once(YOLO)系列算法引入了单次前向计算直接得到检测结果的创新设计,通过将图像分割为网格并在每个网格中进行类别和边界框的预测,快速提升处理速度。YOLO 系列算法尤其适合医疗诊断的辅助系统这类需要实时反馈的应用场景。YOLOv8^[6]模型在前期版本的基础上进一步优化了特征提取网络和检测头,优化了处理速度与检测精度之间的权衡,实现了更高效的性能表现,使其在医学图像实

时诊断中具有较大优势。然而,对于高度重叠的肿瘤或边界模糊的病变区域,YOLOv8 仍有进一步优化的空间。

为了解决上述问题,提出了 BR-YOLOv8,该算法结合了 Coordinate Attention^[7](CA)机制、深度可分离卷积^[8](Depthwise Separable Convolution, DW 卷积)以及 Bottleneck Transformer^[9](BoTNet)开展以下工作:

(1)CA 通过将通道注意力分解为两个沿不同空间方向聚合特征的 1D 特征编码,能够在一个方向捕获长程依赖的同时在另一个方向保留精确的位置信息。这种方式有效增强了肿瘤区域的特征聚焦,使得模型能够更好地识别边界模糊或重叠的肿瘤,从而提升检测精度。

(2)在模型中引入 DW 卷积,优化了特征提取过程,提升了检测效率,同时保持了较高的性能。尤其在处理大规模医学图像时,这一改进显著降低了计算负担,使得模型在实时检测中更具优势。

(3)BoTNet 模块通过多层自注意力机制增强了全局信息建模能力,能够在不同尺度上捕捉图像的长程依赖关系。这一改进使得模型在复杂的医学图像中能够更准确地定位肿瘤区域,尤其在背景复杂或肿瘤小且边界不明显的情况下,表现出显著的性能提升。

1 脑肿瘤图像检测算法

1.1 BR-YOLOv8 网络结构

YOLOv8 是 Ultralytics 公司于 2023 年发布的一款算法。其网络架构分为主干网络(Backbone)、颈部网络(Neck)和输出头部(Head)三大关键部分。本文对 YOLOv8 算法进行改进,提出了 BR-YOLOv8 脑肿瘤图像检测算法,其网络结构如图 1 所示。相较于原始的 YOLOv8 模型,改进后的模型把原有的卷积替换成了 DW 卷积,并且在 backbone 网络中引入了 CA,以及在末端部分加入了 BoTNet。

1.2 Coordinate Attention 注意力机制

CA 是一种近年来提出的高效注意力机制,旨在克服传统通道注意力机制中对空间信息捕捉不足的问题。传统的注意力机制如 Squeeze-and-Excitation^[10](SE)网络缺失了精确的空间位置信息,这使得其在需要精确定位和捕捉空间关系的任务中,特别是目标检测和语义分割等视觉任务中,表现不佳。

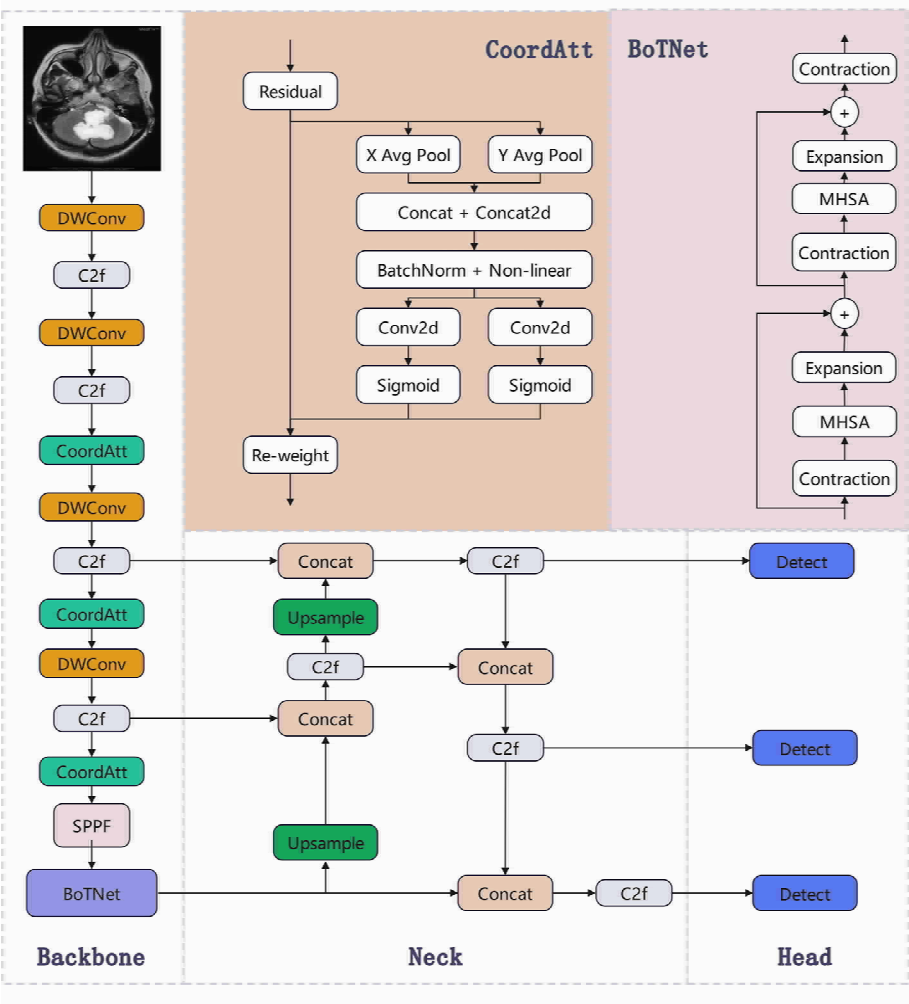


图 1 BR-YOLOv8 网络模型的架构

CA 有效地缓解了这一问题。将注意力机制分解为两个平行的一维特征图，分别沿着空间的两个方向进行聚合。这种设计不仅能够捕捉到通道间的依赖关系，还能显著加强空间方向上的感受野，从而避免了全局池化导致的位置信息丢失。每个方向的特征图被编码为独立的注意力图，分别表示输入特征图沿一个空间方向的长期依赖关系。这两个注意力图通过乘法操作与原始特征图进行融合，从而有效地增强了模型在空间维度上的表达能力，尤其是在捕捉远距离依赖关系时，能保持准确的定位。它的优势在于其灵活性和轻量性，能够轻松集成到传统网络模块中，且计算开销相对较小。通过这一机制，模型不仅能够更精确地捕捉重要的通道信息，还能在更高维度上处理空间信息，强化肿瘤区域的特征聚焦。

CA 通过自适应平均池化操作分别提取特征图在高度和宽度方向上的信息。具体来说，输入特征图 x 经过两个池化层，分别生成高度特征 x_h 和

宽度特征 x_w ，拼接过程如式(1)所示。

$$y = [x_h; x_w] \tag{1}$$

接下来，将拼接后的特征 y 输入一个 1×1 卷积层中，经过批归一化和非线性激活函数 y_swish ，生成压缩特征 y' 。这个过程不仅能减少通道数，还可以增强特征的表达能力。通过这种方式，能够更好地捕捉到通道之间的复杂关系。

随后，从压缩后的特征中生成注意力权重，分别用于高度和宽度方向的特征建模，具体公式如式(2)、式(3)所示。

$$a_h = \sigma(\text{Conv}(x_h)) \tag{2}$$

$$a_w = \sigma(\text{Conv}(x_w)) \tag{3}$$

式中， σ 表示 sigmoid 激活函数^[11]，用于将卷积输出映射到 $[0, 1]$ 区间，表示各位置的重要性。最后，输出特征通过式(4)计算得出。

$$\text{out} = \text{identity} \times a_h \times a_w \tag{4}$$

这个操作确保了每个通道的特征在空间上的重要性被有效地考虑。通过结合通道和空间注意

力,CA 能够显著增强模型对不同位置特征的响应,进一步提升图像分类、目标检测等视觉任务的性能。这种机制不仅保留了特征图的位置信息,还通过注意力机制引导模型聚焦于重要的空间区域,提高准确性和鲁棒性。

1.3 DW 卷积

DW 卷积是一种高效的卷积操作,通过将传统卷积分解为深度卷积 (Depthwise Convolution, DC) 和逐点卷积 (Pointwise Convolution, PC) 两部分,大幅度减少计算复杂度和模型参数,同时保持良好的特征提取能力。这种设计已广泛应用于轻量化神经网络,如 MobileNet^[12] 和 EfficientNet^[13],尤其适合计算资源受限的场景。DC 是 DW 卷积的第一步,它在每个输入通道上单独执行卷积操作,旨在提取每个通道的空间特征。假设输入特征图的通道数为 C_{in} ,卷积核大小为 $k \times k$,DC 的计算如式(5)所示。

$$Y_{out}^{(c)} = X_{in}^{(c)} * K_{DW}^{(c)}, c = 1, 2, \dots, C_{in} \quad (5)$$

其中 $X_{in}^{(c)}$ 是第 c 个输入通道的特征图, $K_{DW}^{(c)}$ 是深度卷积核。DC 会保留输入通道数 C_{in} 不变,仅在空间维度上提取特征。

PC 是 DW 卷积的第二步,它通过 1×1 的卷积核对 DC 的输出通道进行线性组合,负责融合跨通道的信息并将通道数从 C_{in} 转换为 C_{out} 。PC 的计算如式(6)所示。

$$Y_{out} = \sum_{c=1}^{C_{in}} Y_{out}^{(c)} * K_{PW}^{(c)} \quad (6)$$

其中 $K_{PW}^{(c)}$ 是 PC 的 1×1 卷积核, Y_{out} 是最终输出特征图。

DW 卷积相比传统卷积,其计算复杂度显著降低。传统卷积的计算复杂度如式(7)。

$$FLOPs_{Conv} = C_{in} \times C_{out} \times K^2 \times H \times W \quad (7)$$

而 DW 卷积的复杂度如式(8)。

$$FLOPs_{DSC} = C_{in} \times H \times W \times (K^2 + C_{out}) \quad (8)$$

通过这种分离操作,计算复杂度减少了约 $\frac{k^2}{k^2 + C_{out}}$,特别当 $k^2 \gg C_{out}$ 时,DW 卷积可以显著提高效率。

1.4 Bottleneck Transformer

YOLOv8 通过将输入图像划分为多个网格,并在每个网格上独立预测边界框与类别概率,能够实现高效且精准的物体检测。在此基础上,集成了 BoTNet 架构后,YOLOv8 进一步增强了全局信息建模能力。BoTNet 引入的多层自注意力机制使得模型能够在不同尺度和空间范围内,动态地调整

对不同图像区域的关注程度。自注意力机制通过捕捉远程依赖关系,有效地整合了图像的全局信息,从而提升了模型对复杂场景和长距离上下文的理解能力。通过这一机制,模型能够更加精准地识别并区分相互关联的目标,即使在密集背景或遮挡条件下,依然保持较高的检测性能。在结合过程中,输入图像首先通过 YOLOv8 的卷积层提取初步特征 X_{feat} ,其中 $X_{feat} = \text{Conv}(X)$ 。然后将这些特征输入 BoTNet 的编码器进行自注意力处理,生成全局特征表示 X_{global} ,即表示为式(9)。

$$X_{global} = X_{feat} + \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (9)$$

最终,这些处理过的特征被送入 YOLOv8 的检测头进行边界框和类别预测,输出为式(10)。

$$Y_{out} = \text{sigmoid}(X_{global}) \quad (10)$$

这种结合不仅提高了检测精度,还增强了模型对复杂场景和多尺度物体的适应能力,使得结合 BoTNet 的 YOLOv8 在目标检测任务中表现更加出色。

2 实验设计与分析

2.1 实验环境

本实验采用 Ubuntu 20.04 操作系统,搭载 RTX4090 显卡(24 GB 显存)的 GPU 和 18 核 AMD-EPYC-9754-128-Core 处理器,深度学习框架为 PyTorch 2.0.0,使用的 Python 版本为 3.8.10,CUDA 版本为 11.8,训练过程中,模型学习率设置为 0.01, batch_size 大小为 32,训练轮数为 200,优化器为 SGD,动量设置为 0.937。

2.2 评价指标

本文使用准确率(P)、召回率(R)、平均精度均值(mAP)、算法参数量(Params)和检测时间(Detect Time)作为脑肿瘤图像检测的评价指标。以下参数用于上述评估指标的公式中:TP(预测为阳性样本,实际上也是阳性样本)、FP(预测为阳性样本,但实际上是阴性样本)和 FN(预测为阴性样本,但实际上是阳性样本)。Intersection over Union^[14] (IoU)表示预测边界框(B_{Pred})和真实边界框(B_{gt})之间的交集和连接的比率。 AP_i 表示类别索引值为 i 的 AP 值, N 表示训练数据集中样本的类别数(N 设为 3)。 $T_{forward}$ 表示前向传播时间, T_{post} 表示后处理时间。mAP0.5表示当检测模型的 IOU 被设置为 0.5 时的平均精度,mAP0.5 : 0.95表示当检测模型的 IOU 被设置为从 0.5 到 0.95(以 0.5 的间隔取值)时的平均精

度,数值越高代表模型表现越好。通过参数量和检测时间用来评估模型的计算复杂度,这两个指标越小,表示模型所需要的计算能力就越小。各个评价指标的计算公式如下。

$$P=\frac{TP}{TP+FP} \tag{11}$$

$$R=\frac{TP}{TP+FN} \tag{12}$$

$$AP=\int_0^1 P(R)d(R) \tag{13}$$

$$mAP=\frac{1}{N}\sum_{i=0}^N AP_i \tag{14}$$

$$IoU=\frac{|B_{Pred} \cap B_{gt}|}{|B_{Pred} \cup B_{gt}|} \tag{15}$$

$$DetectTime=T_{forward}+T_{post} \tag{16}$$

2.3 数据集

本实验使用了 Brain Tumor Detection^[15]数据集来验证改进后方法的有效性,图 2 展示了该数据集的样本。这些图像经过资深专家亲自标注和核实,能突出显示与指定标签相对应的肿瘤区域,帮助准确检测三个不同类别的脑肿瘤。该数据集共有 9900 张图像,实验中按 7 : 2 : 1 随机划分为训练集、验证集和测试集,方便模型进行训练,本文将该数据集图像大小重采样到 640×640。

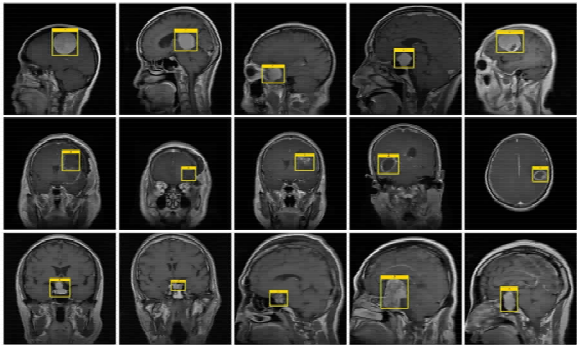


图 2 Brain Tumor Detection 数据集示例显示

Brain Tumor Detection 包含 3 种不同类型的对象。图 3 显示了与此数据集中对象的手动标记相关的信息。本文在图 3 中按照从左到右、从上到下的顺序显示子图。第一个子图显示数据集中每种类型的对象数量;第二子图显示了数据集中对象边界框的大小,所有对象框的中心坐标固定在一个点上,对象边界框的大小表明数据集包含大量小面积对象;第三子图是对象包围盒的中心点的坐标分布,可以看出,对象的中心点主要在图像数据区域的上方和下方;第四个子图是对象边界框的相应宽度和高度的散点图,最暗的颜色位于图的左下角。

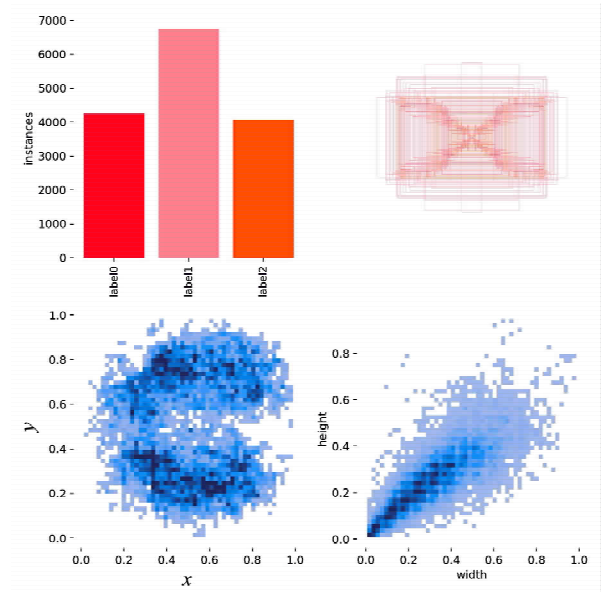


图 3 关于 Brain Tumor Detection 数据集中对象的信息

2.4 检测结果及分析

为了更好地验证模型性能,本文设计了 3 组实验。第 1 组实验旨在选出最合适本文研究的基准模型。第 2 组实验为消融实验,将原始 YOLOv8 模型与所有改进后的模型进行对比。第 3 组为对比实验,将改进后的模型与一些 YOLO 系列算法与其他优秀的检测模型进行比较。

2.4.1 基准模型选择实验

为了选择最适合脑肿瘤图像检测的基准模型,本文在 Brain Tumor Detection 数据集上将所提出的 YOLOv8 模型中一些不同尺寸,包括 YOLOv8n、YOLOv8s、YOLOv8m、YOLOv8l 进行了比较。结果如表 1 所示,兼顾以下各项指标综合分析,本文选取 YOLOv8l 作为基准模型,并在此基础上对其进行改进,使其对脑肿瘤图像检测的能力进一步提升。

2.4.2 消融实验

为了验证每一个改进部分对脑肿瘤图像检测模型的影响,本文在原 YOLOv8 的基础上设计消融实验。在网络模型中分别添加 CA、DW 卷积和 BoTNet,来对比各个模块对整个实验的影响,在精确率、召回率、mAP0.5、mAP0.5 : 0.95 这四个指标中来验证各个改进点的有效性。实验结果如表 2 所示。

由表 2 可知,加入 CA 后对模型的精度、召回率和 mAP0.5 : 0.95 的影响较为有限,表明 CA 在此任务中的贡献较为微小。引入 DW 卷积后,精度下降至 92.3%,但由于召回率的提高展现出了模型在小目标检测上表现出一定的优势。BoTNet 的引入使精度有所提升,并且在 mAP0.5 和召回

率方面也略有改善,表明该模块在捕获长程依赖和全局信息建模上起到了积极作用。

表 1 采用四种不同尺寸的 YOLOv8 对所提出模型的比较实验结果

模型	精度	召回率	mAP0.5	mAP0.5 : 0.95	参数量	检测时间/ms
YOLOv8n	89.3%	72.6%	80.6%	52.9%	0.31×10^7	3.48
YOLOv8s	91.6%	76.7%	83.5%	58.1%	1.1×10^7	4.05
YOLOv8m	91.5%	77.7%	84.2%	60.8%	2.5×10^7	5.45
YOLOv8l	92.6%	77.6%	84.5%	62.1%	4.3×10^7	6.2

表 2 消融实验结果

模型	CA	DW 卷积	BoTNet	精度	召回率	mAP0.5	mAP0.5 : 0.95
YOLOv8	—	—	—	92.6%	77.6%	84.5%	62.1%
	✓	—	—	92.7%	77.2%	84.2%	61.7%
	—	✓	—	92.3%	77.9%	84.3%	61.7%
	—	—	✓	92.7%	77.7%	84.6%	61.9%
	✓	✓	—	92.4%	77.4%	84.5%	61.7%
	✓	—	✓	91.8%	78.2%	84.2%	61.7%
	—	✓	✓	92.9%	77.3%	84.1%	61.4%
	✓	✓	✓	93%	78.7%	86.6%	63.8%

将 CA 和 DW 卷积结合后,精度下降至 92.4%,召回率和 mAP0.5 基本保持稳定,得出这两者的结合未能显著增强模型的表现。而当 CA 与 BoTNet 结合时,精度下降至 91.8%,mAP0.5 保持不变,但召回率显著提高,反映出该组合结合时对模型的全局信息建模能力有所补充。DW 卷积与 BoTNet 结合时,精度上升至 92.9%,召回率维持在 77.3%,但 mAP0.5 : 0.95 略微下降,说明两者组合在提升精度和召回率方面有所作用,但在处理复杂场景下的细节表现略逊色。当 CA、DW 卷积和 BoTNet 模块全部引入时,模型性能达到最优,精度、召回率、mAP0.5 和 mAP0.5 : 0.95 四个指标相比于基准模型分别提升了 0.4%、1.1%、2.1%和 1.7%,这说明三者的组合在脑肿瘤检测中表现出了显著优势。

2.4.3 对比实验

为了进一步验证模型性能,将所提出的模型与一些 YOLO 系列算法与其他优秀的检测模型进行

了比较。本文使用的 YOLOv5^[16] 采用了 Mosaic 数据增强^[17] 技术,这不仅提升了模型的训练速度,还显著提高了检测准确性。其独特的 Focus 结构^[18] 有效减少了计算量和参数量,从而在保持检测性能的同时加快了推理速度。YOLOv7^[19] 进一步优化了网络架构,引入了 ELAN 结构^[20],使得模型在性能与效率之间达到了更好的平衡。YOLOv9^[21] 在 YOLOv7 的基础上进行了优化,整合了最新的模型压缩技术,提高了实时检测能力。此外,SSD 是一种高效的目标检测算法,通过在不同尺度的特征图上进行检测,实现了较快的处理速度和良好的精度。其核心思想是同时预测边界框和类别概率,在一次前向传递中完成检测。Faster-RCNN 则结合了区域建议网络^[22] (RPN),确保了高检测精度。该方法在准确性上表现出色,但计算开销相对较大。通过比较分析,可以更全面地评估所提出算法的性能及其在目标检测领域的应用潜力。对比实验结果如表 3 所示。

表 3 对比试验结果

模型	精度	召回率	mAP0.5	mAP0.5 : 0.95	参数	检测时间/ms
YOLOv5	85%	76.1%	77.2%	46.1%	2.1×10^7	10.12
YOLOv7	90.9%	77.7%	83%	54.5%	3.7×10^7	9.52
YOLOv8	92.6%	77.6%	84.5%	62.1%	4.3×10^7	6.2
YOLOv9	89.9%	76%	82.7%	57%	5.1×10^7	23.99
SSD	91.6%	55.1%	70.5%	33.9%	13.8×10^7	156.56
Faster Rcn	71.9%	67.6%	69.34%	34.2%	2.35×10^7	25.3
BR-YOLOv8(Ours)	93%	78.7%	86.6%	63.8%	9.1×10^7	9.68

通过对比实验可以发现,本文所提出的 BR-YOLOv8 模型相较于传统的 Faster R-CNN、SSD 和 YOLOv5 模型提升效果显著。同时,与 YOLOv7、YOLOv8 和 YOLOv9 等较大模型相比,本文模型在精度上分别提升了 2.1%、0.4% 和 3.1%,召回率分别提升了 1.0%、1.1% 和 2.7%。mAP0.5 相较于 YOLOv5、YOLOv7 和 YOLOv9 分别提升了 9.4%、3.6% 和 3.9%,mAP0.5 : 0.95 相较于 YOLOv8 和 YOLOv9 分别提升了 1.7% 和 6.8%。尽管 BR-YOLOv8 的参数量和检测时间较 YOLOv8 有所增加,但在与精度、召回率、mAP0.5 以及 mAP0.5 : 0.95 这四个指标上的提升相比仍然具备优势。综上可知,本文提出的模型在保持较高检测精度的同时,达到了优秀的综合性能,在处理少数复杂背景或边缘模糊的脑肿瘤图像检测任务中表现出显著的优势。

图 4 为七种不同模型在脑肿瘤数据集上训练后的 mAP 曲线图。横坐标为迭代次数,纵坐标为 mAP0.5 值。由图 4 可知,从一开始经过大约 75 轮的训练,七种网络模型在 100 轮的训练后开始趋于平稳。除 YOLOv5、SSD、Faster-RCNN 这三种网络模型外,其余四种网络模型的 mAP0.5 值均达到了 80% 以上,且本文提出的 BR-YOLOv8 网络模型收敛速度与 YOLOv8 相近,但是最终的 mAP0.5 值达到最高,为 86.7%。与 YOLOv8 相比,该方法检测效果更好。

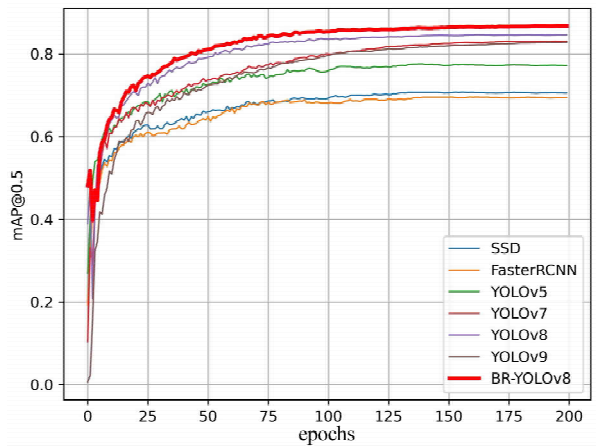


图 4 七种网络模型在 200 个 epochs 中的 mAP 值对比

2.4.4 可视化分析

深度学习模型的可解释性差,在某种程度上阻碍了深度学习的发展和应用。为了直观、方便地绘制模型的检测效果图,本文从混淆矩阵对模型的检测性能进行了对比实验分析,最后,为了验证该方法的可推广性,在自制数据上进行了推理实验。混

淆矩阵显示了分类模型进行预测时会对哪些部分产生混淆,显示分类模型现有的错误并预测未来可能发生的错误。对角线上的值代表 TP 值,对角线数据颜色越蓝,数字越大,则代表正确分类数据结果越好,分类性能越好。

为了直观地显示方法预测目标类别的能力,进一步绘制了 YOLOv8 和 BR-YOLOv8 的混淆矩阵,如图 5 所示。混淆矩阵的行和列分别表示真实类别和预测类别,对角线区域中的值表示正确预测类别的比例,而其他区域中的值表示错误预测类别的比例。从图 5 中可以看出,BR-YOLOv8 中混淆矩阵的对角线区域数值比 YOLOv8 更大,这表明本文模型正确预测物体类别的能力得到了增强。

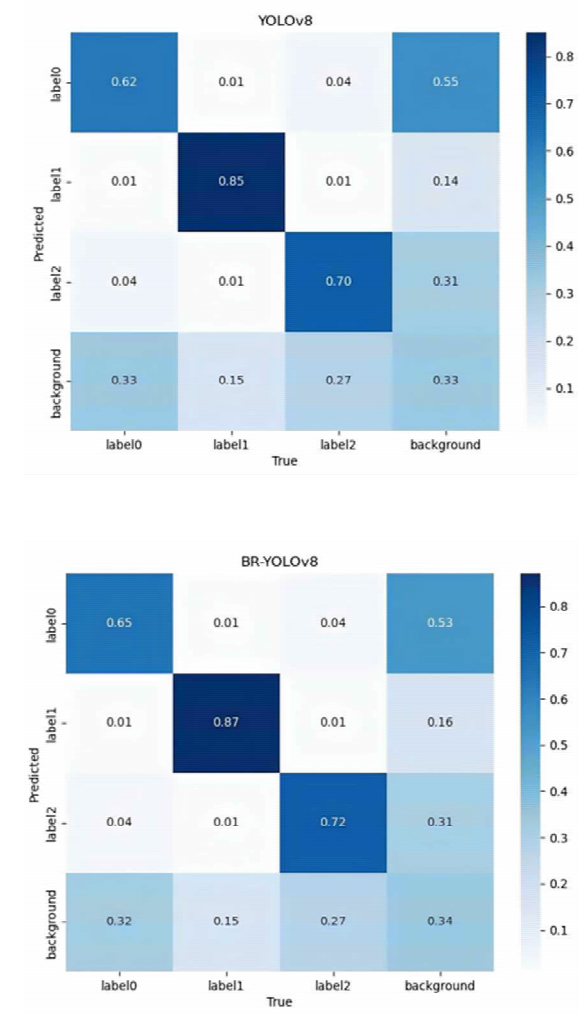


图 5 YOLOv8 和 BR-YOLOv8 的混淆矩阵对比

3 结 论

针对脑肿瘤图像检测中的精度不足等问题,提

出了 BR-YOLOv8 算法。该算法引入了 CA,以强化模型在处理脑肿瘤图像时对肿瘤区域的聚焦能力。同时,采用 DW 卷积优化了 Backbone 网络中的卷积操作,通过减少计算复杂度,使得模型在保持高效特征提取的同时,能够提升检测速度。最后结合 BoTNet,通过全局自注意力机制加强了对长程依赖的建模,使得模型能够更好地捕捉图像中的复杂信息。实验结果表明,BR-YOLOv8 在处理少数复杂背景或边缘模糊的脑肿瘤图像检测任务中表现优异,能够有效识别出细小且难以定位的脑肿瘤。然而,该模型相较于其他模型在检测时间指标上略显不足,尤其是在需要实时处理的场景下,可能面临一定的计算瓶颈。未来的工作可以着重探索轻量化网络设计和高效注意力机制,以进一步提升模型的计算效率,并增强其在少样本和跨域场景下的适应能力,同时确保在精度与速度之间取得更好的平衡。

参考文献

- [1] 姜林奇,宁春玉,余海涛. 基于多尺度特征与通道特征融合的脑肿瘤良恶性分类模型[J]. 中国光学(中英文),2022,15(6):1339—1349.
- [2] 周翔宇,曲喜悦,许杰,等. 改进 YOLOV5 目标检测模型的实时抽烟检测方法[J]. 计算技术与自动化,2023,42(4):81—84.
- [3] REN S. Faster r-cnn: towards real-time object detection with region proposal networks[J]. arXiv preprint arXiv:1506.01497, 2015.
- [4] LIU W, ANGUELOV D, ERHAN D, et al. Ssd: single shot multibox detector[C]//Computer Vision-ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11 - 14, 2016, Proceedings, Part I 14. Springer International Publishing, 2016: 21—37.
- [5] REDMON J, DIWALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2016: 779—788.
- [6] JOCHER G, CHAURASIA A, QIU J. Ultralytics YOLO (Version 8.0.0) [CP]. <https://github.com/ultralytics/ultralytics>.
- [7] 曹嘉璇,王家琪,邹北骥,等. 基于多分组协调注意力卷积的缺血性脑卒中病灶分割方法[J]. 智能计算机与应用,2025,15(4):9—16.
- [8] JIN X, XIE Y, WEI X S, et al. Delving deep into spatial pooling for squeeze-and-excitation networks[J]. Pattern Recognition, 2022, 121: 108159.
- [9] SRINIVAS A, LIN T Y, PARMAR N, et al. Bottleneck transformers for visual recognition[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 16519—16529.
- [10] JIN X, XIE Y, WEI X S, et al. Delving deep into spatial pooling for squeeze-and-excitation networks[J]. Pattern Recognition, 2022, 121: 108159.
- [11] 何从蓉,邢世威,吴晨健. 一种 Sigmoid 为激活函数的高精度 ANN 电路的设计[J]. 电子器件,2024,47(3):617—622.
- [12] HOWARD A G. Mobilenets: efficient convolutional neural networks for mobile vision applications[J]. arXiv preprint arXiv:1704.04861, 2017.
- [13] TAN M, LE Q. Efficientnet: rethinking model scaling for convolutional neural networks[C]//International Conference On Machine Learning. PMLR, 2019: 6105—6114.
- [14] 马建民,朱磊,骆友曾. 基于多目标跟踪算法的车站人流聚集监测研究[J]. 计算技术与自动化,2024,43(3):75—81.
- [15] DARABI P. Medical image dataSet; brain tumor detection[J/OL]. Roboflow, 2023. <http://doi.org/10.13140/RG.2.2.15973.20964>.
- [16] 方浩楠,李登鹏. 基于 YOLOv5 的行李运输系统轮对裂纹检测算法[J]. 计算技术与自动化,2024,43(2):77—81.
- [17] DADBOUD F, PATEL V, MEHTA V, et al. Single-stage uav detection and classification with yolov5: mosaic data augmentation and panet[C]//2021 17th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). IEEE, 2021: 1—8.
- [18] SONG A, ZHAO Z, XIONG Q, et al. Lightweight the focus module in yolov5 by dilated convolution[C]//2022 3rd International Conference on Computer Vision, Image and Deep Learning & International Conference on Computer Engineering and Applications (CVIDL & ICCEA). IEEE, 2022: 111—114.
- [19] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,2023: 7464—7475.
- [20] ROY A S, BAGADE P. Attentive-YOLO: on-site water pipeline inspection using efficient channel attention and reduced ELAN-based YOLOv7 [C]//VISIGRAPP (4): VISAPP,2024: 492—499.
- [21] WANG C Y, YEH I H, LIAO H Y M. YOLOv9: learning what you want to learn using programmable gradient information[J]. arXiv preprint arXiv:2402.13616,2024.
- [22] DONG N, ERICSSON L, YANG Y, et al. Label-efficient object detection via region proposal network pre-training[J]. Neurocomputing, 2024, 577: 127376.