



基于改进Yolov8的敦煌壁画元素检测算法

周颜林, 邬开俊, 梅源, 田彬, 俞天秀

引用本文

周颜林, 邬开俊, 梅源, 田彬, 俞天秀. 基于改进Yolov8的敦煌壁画元素检测算法[J]. 计算机科学, 2024, 51(11A): 231000034-6.

ZHOU Yanlin, WU Kaijun, MEI Yuan, TIAN Bin, YU Tianxiu. [Dunhuang Mural Element Detection Algorithm Based on Improved Yolov8](#) [J]. Computer Science, 2024, 51(11A): 231000034-6.

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)

[基于多模态对比学习的场景图生成方法](#)

Multimodal Contrastive Learning Based Scene Graph Generation

计算机科学, 2024, 51(11A): 231200185-5. <https://doi.org/10.11896/jsjcx.231200185>

[基于双重标签分配的遥感有向目标检测方法](#)

Remote Sensing Oriented Object Detection Method Based on Dual-label Assignment

计算机科学, 2024, 51(11A): 240100058-9. <https://doi.org/10.11896/jsjcx.240100058>

[一种改进的基于YOLOv5s的轻量化航拍目标检测模型](#)

Improved Lightweight Aerial Photography Object Detection Model Based on YOLOv5s

计算机科学, 2024, 51(11A): 231100119-8. <https://doi.org/10.11896/jsjcx.231100119>

[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/jsjcx.240100003>

[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/jsjcx.230900037>

基于改进 Yolov8 的敦煌壁画元素检测算法

周颜林^{1,2} 邬开俊¹ 梅源¹ 田彬¹ 俞天秀²

1 兰州交通大学 兰州 730070

2 敦煌研究院 甘肃 敦煌 736200

摘要 敦煌壁画因其极高的艺术价值、历史价值、研究价值而备受关注。在壁画文创研发中,壁画元素检测扮演了一个十分重要的角色。但是,受到壁画脱落、颜料褪色、病虫害破坏、元素体量差异大等因素的影响,给壁画元素的检测工作带来了极大的困扰。为此,文中基于 Yolov8 算法进行了改进拓展工作并将其引入壁画元素的检测任务。具体来说,考虑到部分元素特征不明显的问题,设计了改进的 SPPCSPC 模块以增强模型的特征感知能力,扩大模型的感受野;考虑到元素体量差异巨大、元素风格多变的问题,在 C2f 模块末端引入 CoordAtt 注意力机制以增强网络对局部及非显著信息的关注能力。在敦煌壁画元素检测任务上,相比 5 项前沿检测算法,所提算法取得了先进的壁画原始检测性能。相比 Yolov8 基线算法取得了 2.2% @mAP 的性能提升,尤其是在 main_buddha 类别上提升了 12.2% @mAP 的检测性能。所提方法有效支撑了敦煌壁画的后续相关研究工作。

关键词: 敦煌壁画;改进的 Yolov8;目标检测;特征增强

中图分类号 TP399

Dunhuang Mural Element Detection Algorithm Based on Improved Yolov8

ZHOU Yanlin^{1,2}, WU Kaijun¹, MEI Yuan¹, TIAN Bin¹ and YU Tianxiu²

1 Lanzhou Jiaotong University, Lanzhou 730070, China

2 Dunhuang Academy, Dunhuang, Gansu 736200, China

Abstract The Dunhuang murals have garnered significant attention for their artistic, historical, and research value. In the research and development of cultural tourism surrounding frescoes, detecting elements within these frescoes is crucial. However, due to factors such as shedding, pigment fading, pest damage, and the significant discrepancies in elemental volume, detecting mural elements has become difficult. For this reason, this paper, which is based on the Yolov8 algorithm, continues the improvement and expansion work by introducing it into the fresco element detection task. Specifically, the design of an enhanced SPPCSPC module improves the feature-perception ability of the model and expands its sensory field. Additionally, the CoordAttention mechanism is introduced at the end of the C2f module to improve the network's ability to focus on local and non-significant information, which addresses the variability in volume and style of the elements. On the issue of detecting elements within Dunhuang murals, our algorithm outperforms five other cutting-edge detection algorithms in terms of mural detection accuracy. Compared to the Yolov8 baseline algorithm, it achieves a 2.2% improvement in mAP, particularly in the main_buddha category where we see a 12.2% improvement in detection accuracy. This accomplishment offers significant support for future research focused on Dunhuang murals analysis.

Keywords Dunhuang murals, Improved Yolov8, Target detection, Feature enhancement

1 引言

敦煌莫高窟是享誉世界的文化遗产地,距今已有 1600 多年的历史,现存壁画 4.5 万多平方米,壁画色彩丰富,内容题材形式多样,包含古代社会生活的方方面面,不仅具有极高的艺术价值、历史价值、科技价值、社会价值,而且具有十分重要的研究价值。由于自然环境变化以及人为因素破坏,敦煌壁画出现了不同程度的多种病害情况,20 世纪 90 年代以来,敦煌研究院与国内外保护机构共同合作、积极探索,通过计算机技术及摄影采集手段,展开了以敦煌壁画永久保存、永续利

用为目的的数字化保护工作^[1],目前已经积累海量的文物数字资源,为文创开发、考古、展示展览等应用提供了强有力的数据支持。

目标检测(Object Detection)作为计算机视觉领域主要的研究任务之一,具有广泛的工业应用基础(如遥感测绘、工业质检、疫情防控)并在敦煌壁画数字化应用中扮演着重要的角色。Shen 等^[2]将 Yolov4 引入石窟壁画的病害情况进行检测和分析,取得了高效的研究进展;Dong^[3]对 Yolov4 网络以及 SLIC 超像素分割算法进行了改进,主要用于石窟寺壁画病害的自动高精度识别检测;Zhang 等^[4]提出了一种基于

基金项目:甘肃省自然科学基金(23JRRA913)

This work was supported by the Natural Science Foundation of Gansu Province(23JRRA913).

通信作者:周颜林(zhouyanlin@dha.ac.cn)

YOLOv5 改进的算法,能够对壁画破损病害进行高精度定位检测和分类;Li^[5]基于 ResNet50 网络架构提出了新的改进模型,主要用于敦煌壁画考古年代进行检测;Yu 等^[6]开展了人工智能技术在敦煌文化遗产保护中的探索,在目标检测任务方面将 YOLOv5^[7]网络应用在 5 类敦煌壁画元素的检测任务中,取得了较好的检测性能。为了进一步拓展壁画检测工作,本文拓展了壁画数据量并结合先进的 YOLOv8^[8]算法开展了一些改进及应用迁移工作。具体来说,本文的主要贡献如下:

- 1)首次将 YOLOv8 网络应用到敦煌壁画元素检测任务中,拓展了壁画元素检测的最新研究进展;
- 2)设计了一个改进的 SPPCSPC 模块,通过对输入的特

征图进行多尺度空间金字塔池化处理,在提高模型感受野的同时加强了模型特征表达能力;

3)引入了 CoordAtt 注意力机制并进一步设计了一个 C2fCA 模块去替换原有的 C2f 模块,增强了网络对局部及非显著信息的关注能力。

2 相关工作

2.1 目标检测算法

目标检测一直以来都是计算机视觉的重要任务之一,如图 1 所示,自 2014 年 RCNN^[9]被提出以来,基于神经网络的目标检测算法因其较高的检测性能备受青睐。

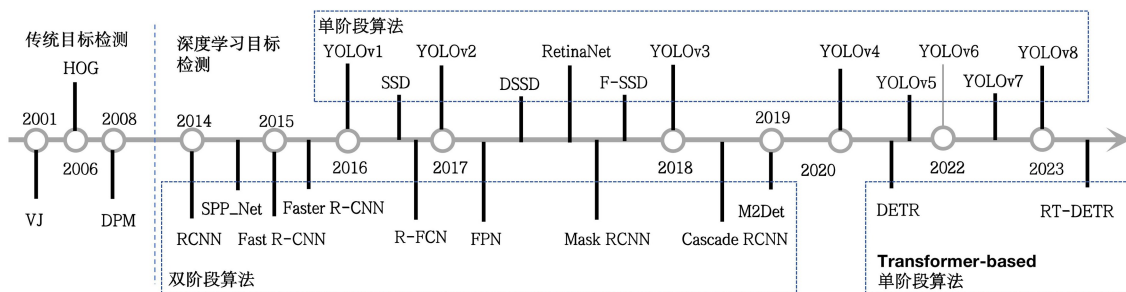


图 1 目标检测的发展历程

Fig. 1 Development history of target detection

双阶段目标检测算法主要有 RCNN, Fast-RCNN^[10], Faster-RCNN^[11]等。2016 年 Redmon 等提出了 Yolo 系列的第一个网络 YOLOv1^[12], Yolo 系列算法成为了单阶段目标检测算法的典型代表之一。Yolo^[7-8, 12-17]系列算法经过不断地优化和改进,在追求精度和速度的平衡方面有了巨大的发展。YOLOv8 是 Yolo 系列目标检测算法中最新的版本,整合多种数据增强技术,在检测精度和检测速度方面均达到了该系列算法的顶点。DETR 算法^[18]的提出,通过摒弃单阶段目标检测算法中冗余重复的后处理过程(如非极大值抑制、锚框筛选),进一步提升了目标检测算法的性能,但是其对训练数据量的需求也远超过 Yolo 系列算法。本文选择 YOLOv8 网络作为首选网络,除了考虑到网络优秀的性能,还考虑到了该模型在部署应用方面的高效便捷性。

2.2 注意力机制

注意力机制在计算机视觉和自然语言文本处理中都具有广泛的应用价值。顾名思义,其核心作用是使得网络将关注重点从全局信息转移到重点信息上,注意力机制具有较少的参数和较快的速度,从而能够进一步提升网络的性能。卷积神经网络中较常使用的是通道注意力机制,代表性的通道注意力机制是 SENet 模型^[19],通过自适应地计算每个通道的重要程度,弥补了卷积层和池化层提取图像特征过程中对特征通道之间关注不足的问题,通过为每个通道赋予通道权重的方式,抑制贡献相对较小的通道,从而提升神经网络的特征表达和性能;空间注意力机制^[20]在其作用上与通道注意力机制具有相似的作用,空间注意力机制将关注点从通道放到空间信息上,使模型具备能够自适应学习区域权重的能力,通过对重要区域的关注来提升网络性能,文献^[21]提出了一种增强的空间注意力模块 ESA,使得残差特征能够关注到关键空间信息;CBAM^[22]以串行的方式结合了通道注意力和空间注意

力,同时在通道和空间两个维度上增强了卷积神经网络对图像信息的关注度;文献^[23]提出了一种全局注意力机制 GAM,从放大通道、空间宽度和空间高度 3 个维度交互作用的角度出发,实现了注意力机制能够在 3 个维度上获取关键特征;文献^[24]从协同工作的角度考虑,提出了一种统一权值的注意力模块 SimAM;考虑到在通道信息上通过全局池化引入空间信息只能捕获局部的信息而无法取得长足信息,CoordAtt^[25]将位置信息嵌入到通道注意力,这不仅能够捕获跨通道的信息还能捕获方向感知和位置感知的信息,增强模型精准定位能力和敏感位置关注能力。因此,本文将 CoordAtt 引入到 YOLOv8 网络中,进而设计出 C2fCA 模块,增强网络在特征提取方面的表现。

2.3 空间金字塔池化

He 等^[26]于 2015 年提出了空间金字塔池化结构 SPP,其目的是将不同大小的输入特征图转化为固定尺寸的特征向量,从而避免了输入图像强制改变大小时产生信息失真导致网络性能下降的问题,空间金字塔池化能够并行处理不同尺度的池化操作,从而使得网络能够更好地使用多尺度图像进行训练;YOLOv5^[7]基于 SPP 结构提出了 SPPF 模块,改进后的 SPPF 较大幅度地减少了模型的计算量;YOLOv6 进一步基于 SPPF 提出优化的 SimSPPF 模块,模型计算速度再次提升;文献^[27]结合空洞卷积^[28]的思想,提出了空洞空间卷积池化金字塔 ASPP 模块,以一种自适应均值池化的方式获取全局特征;文献^[29]从人类视觉感受野的角度出发,借鉴了 Inception^[30]和空洞卷积的思想,提出了 RFP 模块,有效增大了网络的感受野;YOLOv7 使用了 SPPCSPC 模块,该模块在空间金字塔池化的基础上参考了 CSPNet^[31]的思想,通过不同尺寸的池化层增大并获得不同感受野,还能在减少参数量提升速度的同时提高模型精度。本文使用 SPPCSPC 替换 SPPF 模块,增大了模型的感受野。

3 网络结构介绍

本文的网络结构如图 2 所示。本文沿用了 Yolov8 的网络架构并进行了进一步优化设计。具体来说,引入 CoordAtt 注意力设计了 C2fCA 模块,替换了原有结构使用的 C2f 模块,同时简化了原始 SPPCSPC 结构进行优化设计并将其替换为原有的 SPPF 结构。以输入为 $640 \times 640 \times 3$ 的图像为例,利用 backbone 提取 3 个特征图,分别是 P1($80 \times 80 \times 64$)、P2($40 \times 40 \times 128$)、P3($20 \times 20 \times 256$),采用金字塔特征融合,将 P3 与 P2 融合形成 P4,再与 P1 融合形成最终的特征 D1($80 \times 80 \times 64$),再将 D1 与 P4 融合形成 D2($40 \times 40 \times 128$),最后将 D2 与 P3 融合形成 D3($20 \times 20 \times 256$),将 D1、D2、D3 分别送入 3 个检测头。

1) ISPPCSPC 模块。在利用 backbone 提取多尺度图像特征的过程中,有助于提升网络的多层次感知能力,为此我们对 Yolov8 所采用的 SPPF 模块进行了替换。具体来说,SPPCSPC 模块移除了头部的 2 层卷积操作,并以残差的形式连接首位,这旨在降低输入特征的信息损失量。通过实验发现,采用该处理具有更好的检测表现。

2) C2fCA 模块。C2f 模块作为 Yolov8 网络的核心模块,在特征提取和特征融合计算中均起到了极为重要的作用,因此本文选择将 CoordAtt 注意力机制融合进 C2f 模块中。将卷积之后的特征输入到 CoordAtt 注意力机制模块,利用并行的方式进行全局平均池化,分别将其编码为垂直方向和水平方向的感知特征图,然后利用将其融合进输入特征图像来增强特征图的表达能力,具体结果如表 3 所列。

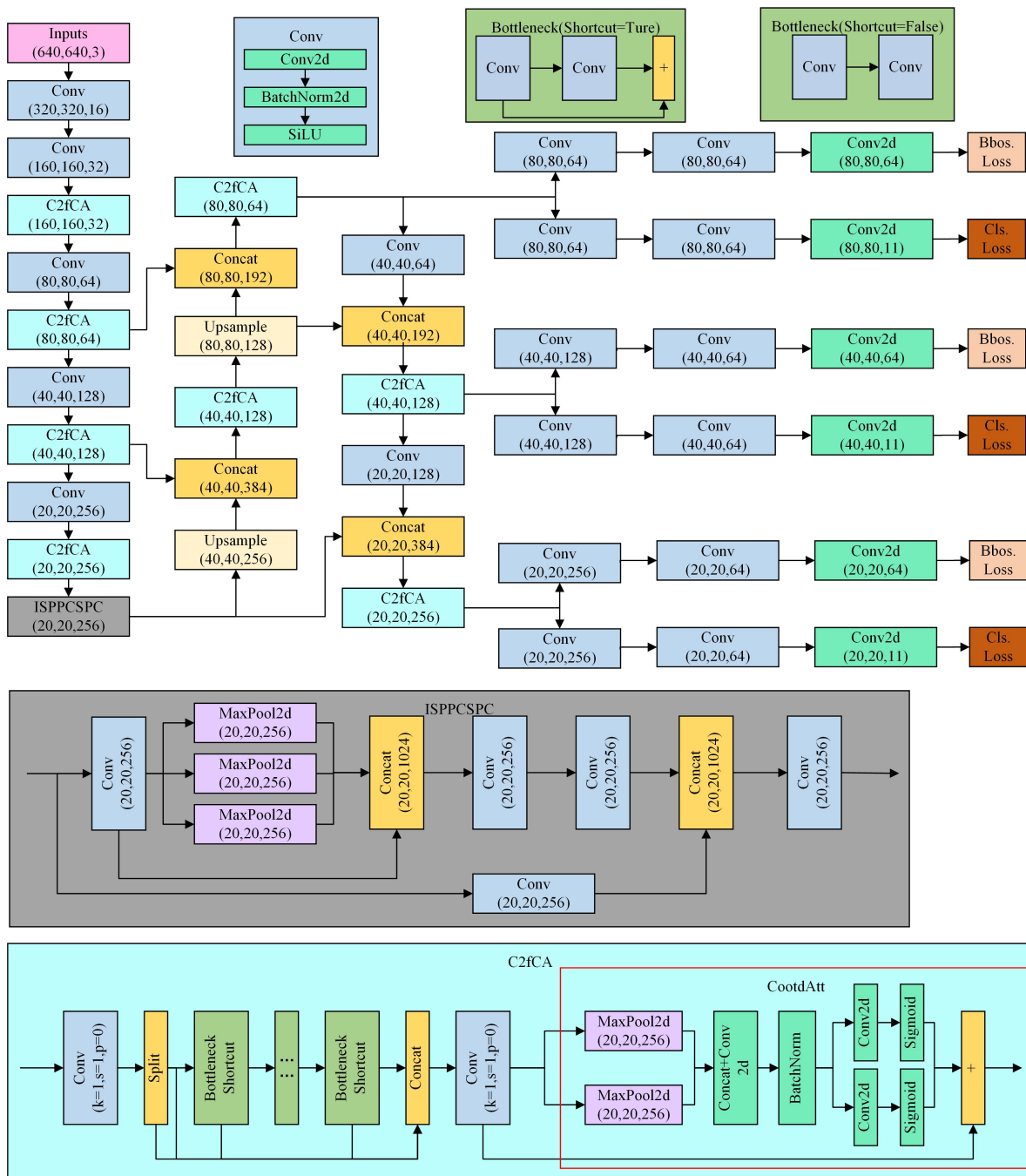


图 2 网络架构图

Fig. 2 Network architecture diagram

4 实验

4.1 实验环境及参数设置

本文实验采用 Pytorch 2.0.1+CUDA11.4 实现本文的改进模型并在 NVIDIA RTX3090 GPU 上完成了模型训练和测试。本文使用模型的 epochs 设置为 100 轮, batch size 设置为 12, 初始学习率设置为 10^{-2} , 学习率采用动态衰减策略, 最终学习率为 10^{-4} 。优化器采用 SGD 优化器, 数据集采用 MOSAIC 数据增强。

4.2 数据集介绍

本文使用自制的敦煌壁画元素数据集, 统计 11 类壁画元素的 6666 张图片中的 28059 个实例, 数据集制作过程综合考虑了各个类别的应用价值及频率, 图像和实例数量分布不均匀, 分布情况如图 3 所示。

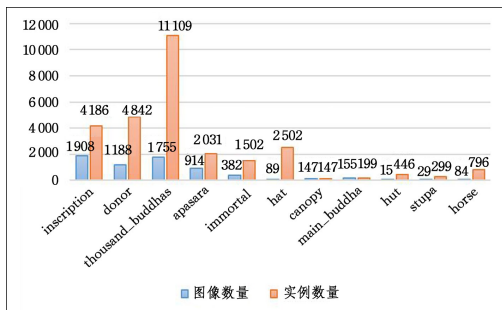


图 3 数据各类别及实例数量分布图

Fig. 3 Distribution of data categories and number of instances

敦煌壁画因病虫害侵袭及壁画材料物理和化学特性导致部分元素图像存在特征损失等不完整情况, 部分数据集展示如图 4 所示。从图中可以看出, 元素种类体量差异巨大, hat 等类别属于微小目标, main_buddha 等类别属于大样本目标, 同时 thousand_buddhas 目标较为密集, inscription 类别特征明显但是部分图像中存在 inscription 局部特征的情况。因此, 这些情况在实际预测时势必会带来检测精度的偏差。



图 4 壁画数据示例

Fig. 4 Example of mural data

从数据集的角度来看, “apasara”类别大多数的背景区分度明显低于其他类别, 背景造成了较大的干扰, 并且其目标范围相对较大; “thousand_buddhas”“donor”“immortal”等 3 个类

别数据密集型对象, 往往是多个对象密集分布在一起, 存在较为严重的自遮挡情况; “main_buddha”和“canopy”呈现一种伴生情况, 基本上都是一起出现; “inscription”和“hut”存在较多的被遮挡情况, 常常仅清晰地展示部分对象; “hat”类别属于微小目标的范畴。整个数据集本身还存在样本量分布不均匀、样本体量差异大、部分对象特征不完备的情况。在模型改进过程中, 注意力机制的引入和 SPPCSPC 模块的引入, 通过对重点区域关注和有效地扩大感受野, 均提高了模型抑制以上数据问题的能力。

4.3 目标检测结果

采用改进后的算法对敦煌壁画元素数据集进行检测, 检测结果如图 5 所示。正如前文的分析, 壁画褪色、脱落等因素造成的目标对象特征不完整、不明显, 导致部分对象检测结果较差。“main_buddha”和“canopy”图像清晰、特征完整, 检测结果较好。“apasara”对象因为背景图像的干扰, 往往存在将背景图像错误检测的结果。整体来看, 对于样本数量较多的“thousand_buddhas”等类别检测结果最好, 特别注意的是“main_buddha”虽然样本数量处于中低段分布, 但是获得了较好的检测结果。“hut”和“hat”对象属于目标样本相对较小的情况下, 检测结果也很好。但是, 模型在检测框的范围准确性上还有待提高, 特别是密集型对象分布的情况, 如“donor”对象, 检测框存在明显过大涵盖超过单个对象的情况。

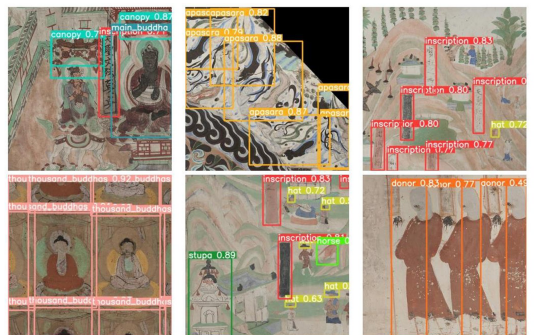


图 5 壁画检测结果

Fig. 5 Mural detection results

4.4 算法对比

为了对比 Yolo 系列模型在敦煌壁画元素数据集上的性能, 本文选择了 Yolo v5s, Yolo v6, Yolo v7 等比较经典的 Yolo 系列算法进行平均精度检测及在各种类别上的检测精度比较, 结果如表 1 所列。

表 1 各类算法的检测性能结果

Table 1 Detection performance results of various algorithms

方法	Params/MB	GFLOPs	mAP/%
DETR ^[18]	475.00	95.0	6.9
Yolov5-n ^[7]	3.90	4.2	82.3
Yolov6-n ^[16]	21.86	11.1	87.5
Yolov7 ^[17]	74.90	105.3	89.9
Yolov8-n ^[8]	7.30	8.1	89.6
ours	10.30	9.4	91.8

从表 1 可以看出, 本文算法在整体的平均精度及大部分类别的检测精度上均取得了先进的检测性能。如前文的预测, 在 stupa 和 horse 两个背景干扰较大的类别上, 改进后的算法对于局部及非显著信息的关注度和特征提取的能力明显

有所加强,对于 stupa 和 horse 的检测性能分别提升了 2.9% @mAP 和 1.7% @mAP。值得注意的是,本文算法对 main_buddha 类别的提升达到了 12.2% @mAP,这说明改进后的算法对大目标的检测性能有了极大的提升。但是对于微小目标而言,检测性能还有待进一步的提升。

4.5 消融实验

本文采用逐步堆叠的方式,验证各个模块对网络性能的提升情况。

4.5.1 ISPPCSPC 模块改进分析

从表 2 可以看出,使用简化后的 SPPCSPC 模块替换原有的 SPPF 模块之后,网络的整体性能提升了 0.6% @mAP,模型对于大多数目标的检测性能出现了不同程度的提升,但是也有 stupa 和 horse 两类目标的检测性能出现了小幅度的下降,出现这种情况的原因是,ISPPCSPC 通过不同程度池化加强模型特征提取的能力, stupa 和 horse 这两类目标因为与背景环境差异相对较小,反而造成了背景信息干扰的情况。移除 SPPCSPC 头部的 2 层卷积操作并以残差的形式连接首位,使得池化层能够直接学习输入数据而不是学习一个复杂的映射关系,从而更多地保留特征信息,同时能够减少模型复杂度与计算量。

ISPPCSPC 与 ASPP 相比,在整体性能相差不大的情况下,我们注意到基于 ISPPCSPC 的情况下特征不完备元素(如 horse)的检测表现更优,考虑到壁画退化和病害导致的对象特征不完备的情况在敦煌壁画中普遍存在,因此本文采用 ISPPCSPC 模块来提高模型对不完备特征的提取能力。

表 3 C2FCA 模块改进分析

Table 3 C2FCA module improvement analysis

策略	Yolov8n (+ISPPCSPC)	+CoordATT ^[25]	+CBAM ^[22]	+ESA ^[21]	+SimAM ^[24]	+GAM ^[23]
平均精度	90.2	91.8(↑)	89.6	90.8	91.3	87.6
inscription	96.3	96.5(↑)	95.9	96.1	96.5	94.7
thousand_buddhas	98.3	98.4(↑)	98.4	98.3	98.4	98.2
donor	98.4	98.5(↑)	98.6	98.4	98.4	98.4
apasara	96.8	97.2(↑)	96.9	97.2	96.5	95.6
hat	76.3	78.0(↑)	76.2	76.6	75.0	72.9
horse	83.7	85.1(↑)	83.4	83.4	82.7	79.6
immortal	93.8	93.7	93.3	93.9	94.0	93.1
hut	98.2	98.0	96.2	98.0	96.8	96.0
stupa	87.4	90.3(↑)	85.4	89.0	90.2	83.3
canopy	83.7	88.5(↑)	86.8	85.1	84.7	80.3
main_buddha	79.1	85.3(↑)	74.9	82.8	91.4	71.2
参数量/MB	10.3	10.3	10.4	10.3	10.2	19.7

表 4 C2FCA 模块其他组合结果

Table 4 C2FCA module other combination results

策略	+ISPPCSPC	+SPP ^[7]	+SimSPPF ^[16]	+ASPP ^[27]
平均精度	90.2(↑)	89.6	89.2	88.9
inscription	96.3	96.8	96.8	97.3
thousand_buddhas	98.3	99.3	99.3	99.1
donor	98.4	99.4	99.4	99.4
apasara	96.8(↑)	97.6	98.2	97.7
hat	76.3(↑)	75.5	76.1	76.0
horse	83.7(↑)	83.4	81.1	84.3
immortal	93.8(↑)	93.8	94.8	94.7
hut	98.2(↑)	98.0	98.4	97.5
sstupa	87.4	88.4	89.4	85.3
canopy	83.7	90.2	85.4	82.0
main_buddha	79.1(↑)	63.8	61.8	64.7
参数量/MB	10.3	7.3	7.3	7.4

表 2 ISPPCSPC 模块改进分析

Table 2 ISPPCSPC module improvement analysis

策略	Yolov8n ^[8]	+ISPPCSPC	+SPP ^[7]	+SimSPPF ^[16]	+ASPP ^[27]
平均精度	89.6	90.2(↑)	88.9	89.9	90.3
inscription	96.5	96.3	96.2	96.3	96.1
thousand_buddhas	98.3	98.3	98.3	98.1	98.4
donor	98.5	98.4	98.3	98.4	98.5
apasara	96.7	96.8(↑)	96.4	96.7	97.3
hat	75.7	76.3(↑)	73.3	75.9	77.1
horse	83.4	83.7(↑)	80.6	81.1	82.0
immortal	94.4	93.8(↑)	93.2	93.7	92.6
hut	97.3	98.2(↑)	97.4	97.0	97.4
sstupa	87.4	87.4	88.6	87.6	90.0
canopy	84.7	83.7	79.8	84.2	86.2
main_buddha	73.1	79.1(↑)	75.7	80.3	78.0
参数量/MB	7.3	10.3	7.2	7.2	7.3

4.5.2 C2FCA 模块改进分析

由表 3 可以看出,本文引入 CoordAtt 重新设计了 C2FCA 模块来替换原有的 C2f 模块,模型的整体性能提升了 2.2% @mAP,尤其关注 stupa 和 horse 两类可以发现,CoordAtt 注意力的引入使模型能够更加关注目标对象的信息,模型从相似背景中分离对象的能力进一步得到优化。值得注意的是,CoordAtt 的引入,虽然对部分类别的检测性能产生了极微小干扰,但是对大目标(main_buddha)的提升较为明显,达到了 12.2% @mAP,相对于 main_buddha 类并不是很多的训练图像和样本实例,改进后模型对大目标的检测性能提升效果较好。C2FCA 模块其他组合结果如表 4 所列。

结束语

本文首次将 Yolov8 网络应用在敦煌壁画元素检测任务中,同时本文提出的改进方法进一步提升了 Yolov8 在敦煌壁画元素检测任务上的性能,良好的检测性能表明 Yolov8 网络在敦煌壁画元素检测中的应用前景广泛,在支持敦煌学研究、敦煌壁画保护、敦煌壁画数字文化资源开放共享中表现优异。本文设计了 ISPPCSPC 模块和 C2FCA 模块,进一步提升了网络对多尺度图像特征的提取能力,尤其是在大目标(main_buddha 等其他类别)上的检测表现尤为显著。本文将有力地推动目标检测算法等深度神经网络在敦煌壁画保护、研究、利用等方面的应用。

参考文献

- [1] WU J. Digital Development of mural cultural heritage—A case

- study of Dunhuang Mogao Grottoes[J]. *China Cultural Heritage*, 2016(2): 34-38.
- [2] SHEN J X, BAO M Y. Detection and Analysis of Grotto Mural Diseases Based on Improved YOLOv4[J]. *Journal of Shanxi Datong University(Natural Science Edition)*, 2023, 39(2): 15-17, 22.
- [3] DONG Y X. Research on Automatic and High Precision Detection of Mural Diseases In grottoes[D]. Beijing: Beijing University of Civil Engineering and Architecture, 2022.
- [4] ZHANG Y E, WU L G. The Detection of Cave Mural Damage Based on Deep Learning[J]. *Yungang Research*, 2022, 2(1): 85-90.
- [5] LI X Y. Chronological Classification And damage Detection of Dunhuang Murals based on Deep Learning[D]. Harbin: Harbin Institute of Technology, 2019.
- [6] YU T, LIN C, ZHANG S, et al. Artificial Intelligence for Dunhuang Cultural Heritage Protection; The Project and the Dataset [J/OL]. *International Journal of Computer Vision*, 2022. <https://doi.org/10.1007/s11263-022-01665-x>.
- [7] <https://github.com/ultralytics/Yolov5>.
- [8] <https://github.com/ultralytics/ultralytics>.
- [9] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation[J]. *IEEE Computer Society*, 2014.
- [10] GIRSHICK R. Fast R-CNN[J]. *Computer Science*, 2015.
- [11] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks [C]//NIPS. 2016.
- [12] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-Time Object Detection[C]// *Computer Vision & Pattern Recognition*. IEEE, 2016.
- [13] REDMON J, FARHADI A. Yolo9000: Better, faster, stronger [C]//CVPR. 2017.
- [14] REDMON J, FARHADI A. Yolo3: An incremental improvement[J]. *arXiv*: 1804. 02767, 2018.
- [15] BOCHKOVSKIY A, WANG C Y, LIAO M. Yolo4: Optimal speed and accuracy of object detection[J]. *arXiv*: 2004. 10934, 2020.
- [16] LI C, LI L, JIANG H, et al. YOLOv6: A single-stage object detection framework for industrial applications[J]. *arXiv*: 2209. 02976, 2022.
- [17] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. Yolo7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[J]. *arXiv e-prints*, 2022.
- [18] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers[C]// *Computer Vision-ECCV 2020*; 16th European Conference, Glasgow, UK. Springer International Publishing, 2020: 213-229.
- [19] HU J, SHEN L, SUN G. Squeeze-and-Excitation Networks [C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). IEEE, 2018.
- [20] JADERBERG M, SIMONYAN K, ZISSERMAN A, et al. Spatial Transformer Networks[J]. *arXiv*: 1506. 02025, 2015.
- [21] LIU J, ZHANG W, TANG Y, et al. Residual feature aggregation network for image super-resolution [C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020: 2359-2368.
- [22] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module[C]// *Proceedings of the European Conference on Computer Vision(ECCV)*. 2018: 3-19.
- [23] LIU Y, SHAO Z, HOFFMANN N. Global attention mechanism: Retain information to enhance channel-spatial interactions [J]. *arXiv*: 2112. 05561, 2021.
- [24] HOU Q, ZHOU D, FENG J. Coordinate Attention for Efficient Mobile Network Design[J]. 2021. DOI: 10. 48550/arXiv. 2103. 02907.
- [25] YANG L, ZHANG R Y, LI L, et al. Simam: A simple, parameter-free attention module for convolutional neural networks [C]// *International Conference on Machine Learning*. PMLR, 2021: 11863-11874.
- [26] HE K, ZHANG X, REN S, et al. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2014, 37(9): 1904-1916.
- [27] CHEN L C, PAPANDEOU G, KOKKINOS I, et al. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 40(4): 834-848.
- [28] LIU J, LI C, LIANG F, et al. Inception convolution with efficient dilation search[C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021: 11486-11495.
- [29] LIU S, HUANG D. Receptive field block net for accurate and fast object detection[C]// *Proceedings of the European Conference on Computer Vision(ECCV)*. 2018: 385-400.
- [30] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015: 1-9.
- [31] WANG C Y, LIAO H Y M, WU Y H, et al. CSPNet: A new backbone that can enhance learning capability of CNN[C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. 2020: 390-391.



ZHOU Yanlin, born in 1989, bachelor degree, librarian. His main research interests include cultural relics digitization, and deep learning and computer vision.