

基于 YOLOv7-CA-BiFPN 的路面缺陷检测

高敏, 李元

(沈阳化工大学 信息工程学院, 沈阳 110142)

摘要: 路面坑洼是主要道路缺陷, 会损坏车辆, 影响驾驶员的安全驾驶, 严重时还会导致交通事故, 针对这个问题, 提出了改进 YOLOv7 的道路坑洼检测算法; 使用 Mosaic+Mixup 进行内置数据增强, 扩充小样本数据集, 增强模型泛化能力; 引入 CA 注意力机制, 将横纵位置信息编码, 保证计算量的同时又能关注大范围位置信息; 采用 BiFPN 双向特征金字塔网络, 通过特征融合多尺度语义特征提高检测效率; 将损失函数 SIoU 替换 CIoU, 有效解决回归中的样本不平衡问题; 实验结果表明, 改进之后的算法在坑洼数据集的平均精度均值和精确率达到了 89.42% 和 90.12%, 相比于原本的 YOLOv7 版本提高了 6.18% 和 1.96%, 更准确更快速地应用于道路维修。

关键词: 坑洼检测; YOLOv7; 注意力机制; 数据增强; BiFPN

Road Surface Pothole Detection Based on YOLOv7-CA-BiFPN

GAO Min, LI Yuan

(College of Information Engineering, Shenyang University of Chemical Technology, Shenyang 110142, China)

Abstract: Road potholes are the main road defects of roads, which can damage vehicles, affect driver safety, and even lead to traffic accidents in severe cases. To address this issue, an improved YOLOv7 road pothole detection algorithm is proposed. Mosaic+Mixup is used to carry out the built-in data augmentation, expand the small sample datasets, and enhance the model generalization ability; By introducing an coordinate attention (CA) attention mechanism, the horizontal and vertical position information is encoded to ensure computational complexity while paying attention to the large-scale position information; BiFPN bidirectional feature pyramid network is adopted to improve detection efficiency through the feature fusion of multi-scale semantic features; By replacing the loss function SIoU with the CIoU, the sample imbalance in regression is effectively solved. Experimental results show that the improved algorithm achieves the mean value and accuracy of 89.42% and 90.12% in pit datasets, which are 6.18% and 1.96% higher than that of the original YOLOv7 version. It can be more accurately and quickly applied to road maintenance.

Keywords: pit detection; YOLOv7; attention mechanism; data augmentation; BiFPN

0 引言

坑洼是由于车辆行驶和气候状况导致的道路上的逐步损坏而形成的一个凹陷。路面上的凹凸不平, 不仅会对汽车造成损害, 而且会对司机的行车造成不利的影

响, 甚至会引发交通事故。中国道路, 尤其是城乡乡镇, 坑洼道路不仅会影响市民行车体验, 出现颠簸、积水等情况, 而且下雨后会影响到视线, 行车会车出事故的概率很高, 由政府负担的成本也在快速增长。为此, 国家的各个部门必须对道路进行维护, 确保道路的健康与安全, 所以对沥青路面的裂缝、凹陷进行及时的监测非常重要^[1]。路面凹陷的检测是计算机视觉领域中的一个热点, 当前, 研究者们对凹陷的检测做了各种各样的尝试。上述方法价格昂贵, 检测范围小, 且不能完全实现非接触式检测; 而传统的人工巡查方式, 不但需要耗费大量的人力资源, 而且由于主观情绪的影响, 检测标准也不统一, 目前深度学习有望以低成本、高效率和高

准确率应用于坑洼检测。

目前, 基于深度学习的物体检测技术主要有单阶段和两阶段两种。其中单阶段检测算法可以直接获得物体的类别以及位置, 其检测的时间相对较短, 具有代表性的有 YOLO 系列、SSD 等^[2]。两阶段检测算法, 先生成很多可能存在目标位置的锚框, 再以这些框为基线, 来检测物体。这种检测与单阶段相比, 速度会比较快, 具有代表性的是 R-CNN、YOLO 等。胡均平等^[3]提出基于改进的 YOLOv3 模型的路面坑洼检测, 通过修改激活函数、输出尺度进行融合以及 K-Means 聚类方法等等改进, 使得算法在复杂环境下对坑洼的检测进行了提升。焦双健等^[4]人设计出基于深度学习的坑洼检测系统, 以 YOLOv5S 为基础, 通过对大量样本数据进行网络模型训练, 获取最优模型, 实现自动化巡检。以上算法均在坑洼检测方面表现良好, 但在诸如小目标、稠密和遮挡等复杂场景下, 仍需进一步改进。

本文提出一种改进 YOLOv7 的路面坑洼检测算法, 通

收稿日期: 2023-05-24; 修回日期: 2023-10-18。

作者简介: 高敏(1999-), 女, 硕士研究生。

李元(1964-), 女, 博士, 教授。

引用格式: 高敏, 李元. 基于 YOLOv7-CA-BiFPN 的路面缺陷检测[J]. 计算机测量与控制, 2024, 32(9): 9-14, 43.

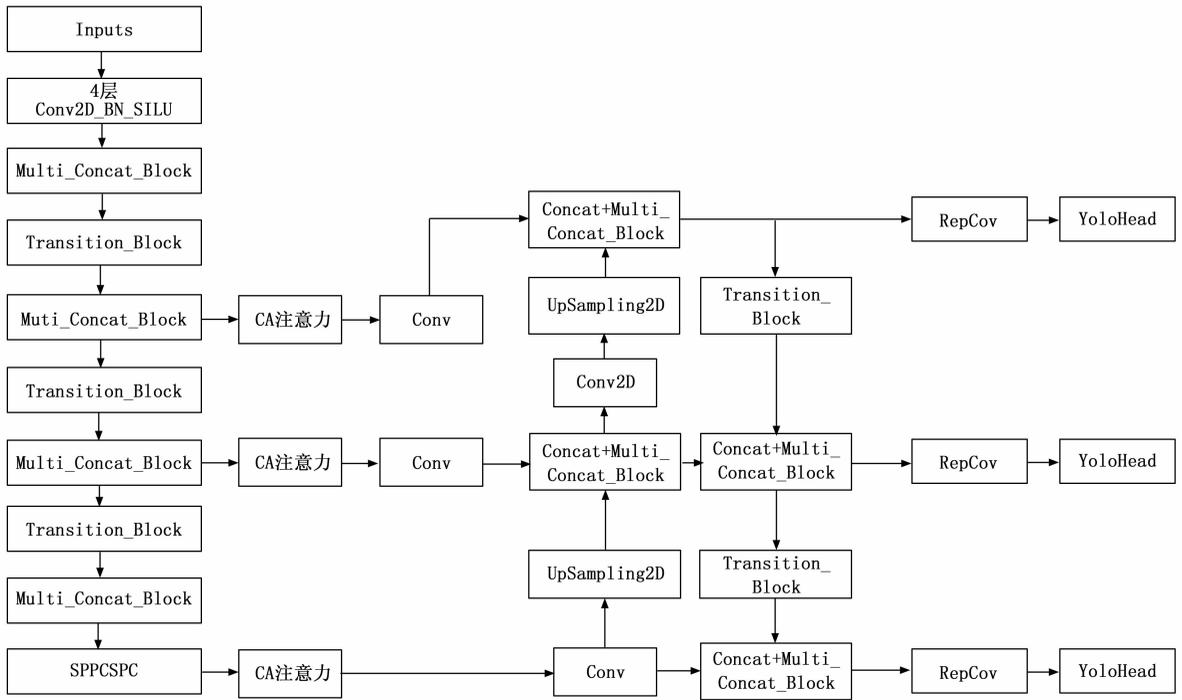


图 1 改进的 YOLOv7 算法网络框架图

过引入 CA (Coordinate Attention) 注意力机制，在通道注意力中，沿横向和纵向将坐标位置信息嵌入，从而使网络对坑洼的对象进行更多的关注；将 Siou 作为边界框回归损失函数，提高网络在训练过程中的收敛速度和回归精度。运用本文的改进方法可以显著提升模型训练速度和检测精度，相比于原始 YOLOv7 的平均精度均值提升了 6.18%，有效提高坑洼检测效率。

1 YOLOv7 网络模型

YOLOv7 主要分为以下几个部分：分别是 Backbone, FPN 以及 Yolo Head。

Backbone 是 YoloV7 的主干特征提取网络，输入的画面首先会在骨干网络内进行特征提取，提取到的特征被称为特征层，它是输入画面的特征集合。在主干网中，我们选择了 3 个层次的特征层，称为“有效特征层”^[5]。

FPN 是 YoloV7 的加强特征提取网络，将 3 个从主干网络中获取到的有效特征层，在此网络中进行特征层的融合，特征融合的目标在于将不同尺度上的特征信息进行综合。FPN 则是利用已有的有效特征层进行后续的特征抽取。在 YoloV7 中，仍然采用了 Panet 的构造，我们将通过上采样和下采样两种方式来特征融合。

Yolo Head 为 YoloV7 的分类器和回归器，通过前面两个部分我们得到了 3 个增强的有效的特征层。每一层特征层都有宽度、高度和通道数目，在这种情况下，可以把特征图看成是一组由 3 个先验框架组成的特征点，每个先验框架包含多个通道^[6]。Yolo Head 所完成的任务，其实就是判定特征点，并判定在特征点上的先验框是否存在与该框

相匹配的对象。与之前的 Yolo 版本类似，在 YoloV7 中使用了解耦头，也就是说，将分类与回归整合到一个 1x1 卷积中。

2 改进的 YOLOv7 坑洼检测算法

YOLOv7 在多状态、遮挡和小样本等的很多场景下的路面检测没有达到很好的效果，以上本文进行对应的改进，也有一定的提高。在网络内部提出 Masica + Mixup 数据增强算法^[7]。mosaic 是 yoloV4 发明的一种 Tricks，它的基本思想是把 4 幅图像随机剪切成一幅图像，然后拼合在一起，形成一幅图像。这样做的优点在于丰富了图片的背景，而且四张图片拼接在一起还变相地提高了 batch_size，在进行 batch normalization 时也会计算四张图片，因此对本身 batch_size 不是很依赖，单块 GPU 就可以训练 YOLOV4。

Backbone 网络和 FPN 之间加入 CA 注意力模块；在颈部加入双向特征金字塔网络 BIFPN，通过特征融合多尺度语义特征提高检测效率；使用 EioU 作为边界框回归损失函数改进之后的网络结构图如图 1 所示。

2.1 Mosaic + Mixup 数据增强

Mosaic 的思路就是将四张图片进行随机剪裁，再拼接到一张图作为训练数据，充盈背景复杂度，且在计算 BN 的时候会计算四张图片的数据。

Mixup 主要应用于图像分类，它从训练样本中随机选取两个样本，展开简单的随机权重求和，与此同时，样本的标签也相应地得到了权重求和，接着，预测结果与权重求和之后的标签求损失，在逆向求导中对参数进行更新，公式的定义如公式 (1) ~ (2) 所示：

$$\tilde{x} = \lambda x_i + (1 - \lambda)x_j \quad (1)$$

$$\tilde{y} = \lambda y_i + (1 - \lambda)y_j \quad (2)$$

2.2 余弦退火学习率

余弦退火策略可以让学习率按照余弦函数进行周期变化, 并在每个周期的最大值时, 重新设定学习率, 以初始学习率为最大学习率, 当 epoch 增大时, 学习率会首先急剧下降, 随后突然上升, 如此反复。学习速率的突然改变, 可以使梯度下降不会停留在任何局部极小点上, 并可能以“跳出”到整体最小点的方法。

2.3 CA 注意力机制

在当前坑洼检测任务中, 图像的模糊和坑洼与地面颜色相似度都影响识别的精度, 引入 CA 模块之后可以更深层次地获取方向感知和位置感知信息, 提升对坑洼目标区域的定位和识别能力。

CA 注意力机制模块的目的是为了提高移动网络学习特征的表达能力, 该方法能够将张量 $X = [x_1, x_2, \dots, x_c]$ $R^{H \times W \times C}$ 进行剖析, 并将其转化为一个尺寸相等的张量 $Y = [y_1, y_2, \dots, y_c]$ $R^{H \times W \times C}$, CA 注意力机制的实现过程如图 2 所示。

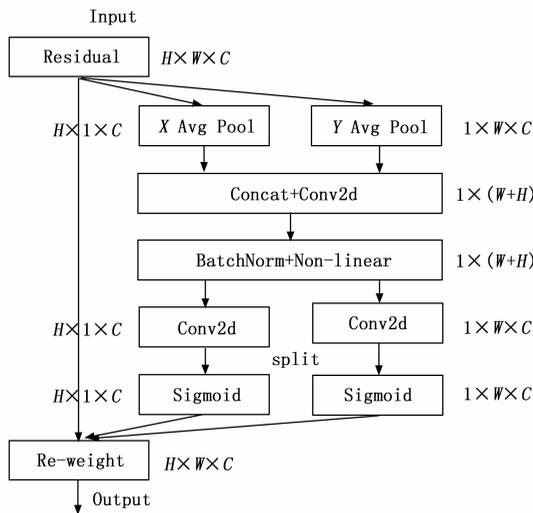


图 2 CA 注意力机制流程图

CA 为了获得图像宽度和高度上的注意力, 并对其准确位置信息展开编码, 首先, 把输入的特征图分成两个方向, 分别对其进行整体平均, 得到两个方向上的特征图。具体如下如公式 (3) ~ (4) 所示:

$$z^h(h) = \frac{1}{W} \sum_{0 \leq i < W} | x_c(h, i) \quad (3)$$

$$Z^w(w) = \frac{1}{H} \sum_{0 \leq j < H} x_c(j, w) \quad (4)$$

接下来将获得全局感受野的宽度和高度两个方向的特征图拼接在一起, 然后将它们送入共享的卷积核为 1×1 的卷积模块, 将其维度降低为原的 C/r , 再将经过批次归一化处理的特征图 F_1 送入 Sigmoid 激发函数, 得到形如 $1 \times (W + H) \times C/r$ 的特征图 f , 如公式 (5) 所示:

$$f = \delta\{F_1[(z^h, z^w)]\} \quad (5)$$

然后, 根据原先的高度和宽度, 将特征图 f 进行卷积, 为 1×1 的卷积, 从而获得与原先相同的特征图 F_h 和 F_w 。在通过 Sigmoid 激活函数之后, 分别获得了特征图在高度和宽度上的注意力权重 g^h 和在宽度方向的注意力权重 g^w 。如公式 (6) ~ (7) 所示:

$$g^h = \sigma[F_h(f^h)] \quad (6)$$

$$g^w = \sigma[F_w(f^w)] \quad (7)$$

通过以上的计算, 得出了在高度方向上的注意权重 g , 在宽度方向上的注意权重 g^w 。在此基础上, 再对原有的图像进行乘性加权处理, 就可以获得图像中的图像, 并给出了相应的图像特征值, 如公式 (8) 所示:

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j) \quad (8)$$

本文在 backbone 输出的 3 个结点之后分别引入 CA 注意力机制, 引导网络对图像中的坑洼目标区域进行了更多的关注, 从而减少了混淆背景对检测准确率的影响。改进之后的网络结构如图 1 所示。

2.4 特征金字塔

2.4.1 FPN

一般来说, 具有较大尺寸的特征图更适合于对小尺寸的目标对象进行检测, 具有较小尺寸的特征图则更适合对大尺寸的目标对象进行检测。如果把这两种方法相结合, 得到的特征金字塔可以更好地同时对大、小目标进行检测。最早期的特征金字塔网络 (FPN, feature pyramid network) 是一个自上而下的单个特征融合金字塔结构, 具体表现在图 3 中。特别地, 较低层次的特征对大样本的识别有帮助。然而, 从下层结构到上层特征之间的距离较远, 给获得精确的位置信息带来了困难。另外, 每一种算法都是在一个由启发性指派的特征层所形成的特征层基础上进行的。由于其它层次上的数据会对最后的预报有所帮助, 所以该过程是可以被更新的。在操作的过程中, FPN 网络的 p6-p3 输出层首先将上一次特征图像的上采样结果进行特征融合, 再通过与 p7 同样的操作流程, 最后生成 5 个不同尺寸的特征图像, 为后续的网络检测做准备^[8]。

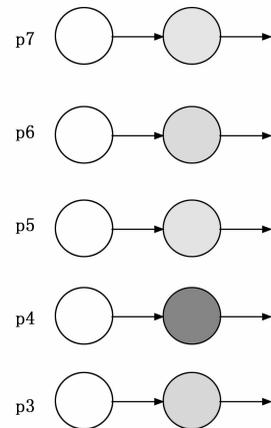


图 3 FPN 结构图

2.4.2 BiFPN

BiFPN 是基于 PAN 技术而得到的一种改进。双向性特征金字塔 (BiFPN) 采用双向性融合的思路, 重构自上而下、自下而上的路径, 实现多个维度的特征信息的融合, 利用上下抽样、上下抽样等方法实现多个维度的特征空间, 实现多个维度的特征空间的有效融合^[9-12]。BiFPN 的构成如下所示, 其中单输入的结点对特征融合的贡献不大, 因此删除了单输入的结点, 以使网络结构简单; 该算法在原输入结点与输出结点间添加一条边线, 目的是在不耗费任何代价的前提下实现更多特征的融合; 为了支持更高级别的特征融合, 将自上而下和自下而上的路径整合在一个模块中, 以便实现更高层次的特征融合。

与其它的特征融合网络相比, BiFPN 具有以下特点: 第一, 去除了单一输入端的节点; 结论是: 仅有一条输入端而无特征融合的结点, 对于以多个特征为目标的特征网来说, 其作用并不大。这样就形成了一个简单的双向网络; 其次, 当两个特征的层次相同时, 在原输入端和输出端各加入一条新的边缘, 使两个特征能够以最小的代价进行融合; 第三, 区别于单条自上而下、单条自下而上的传统 PANet, 我们将每条自上而下、自下而上的双向路径看作一个特征网络, 并对同一层进行多次重复操作, 实现高层特征的融合。

另外, 由于各输入特征的空间分辨率不同, 其对输出特征的影响往往是不均衡的。针对这一问题, 本文提出了对每一项输入进行加权处理, 使网络能够自动识别出每一项输入的重要度。

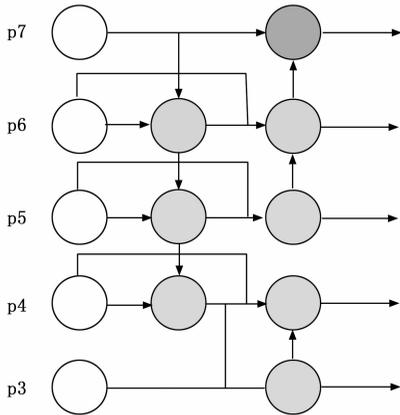


图 4 BiFPN 结构图

2.5 损失函数

损失函数作为一种新的度量方法, 被广泛应用于机器学习、深度学习等领域。这是一种数字评价方法, 可以通过对比模型的结果与真正的标记来衡量模型的表现。它的作用有: 1) 评价模型表现: 利用损耗函数来评价模型预测值与实际值的偏差。当损失值越小时, 说明该模型的预测结果越符合实际情况, 反之, 其误差越大。所以, 丢失函数可以用来测量模型的表现。2) 参数优化: 将损失函数作

为最优算法的目标函数, 用于对机器学习、深度学习等模型进行训练。利用最小化耗散函数的方法, 对所建模型进行了参数调节, 使得所建模型更接近于实际情况。3) Backstepping: 在深度学习中, 利用 Backstep 算法, 求出损失函数相对于模型参数的梯度。利用梯度对模型进行参数修正, 实现了模型的优化。在后向传递过程中, 损失函数起到了很大的作用, 它可以引导参数的调整方向。4) 模型的选取与对比: 针对不同的问题与模型, 采用不同的损失函数。在此基础上, 针对问题特点, 选取适当的损失函数, 使其达到最优。

在机器学习与深度学习中, 损失函数被认为是评价模型性能, 指导参数优化与模型选取的一个关键指标。这是一个非常重要的环节, 也是一个非常重要的环节。

YOLOv7 损失函数由定位损失函数 (Localization loss)、置信度损失函数 (Confidence loss) 和类别损失函数 (Classification) 组成^[13-14]。类别损失和置信度损失均采用 BCEWithLogitsLoss 函数进行计算, 定位损失采用 CIOU 计算, 计算公式如下:

$$L_{CIOU} = 1 - IOU + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \tag{9}$$

$$\alpha = \frac{v}{(1 - IOU) + v} \tag{10}$$

$$v = \frac{4}{\pi^2} \left[\left(\arctan \frac{w^{b^{gt}}}{h^{b^{gt}}} \right) - \arctan \frac{w}{h} \right]^2 \tag{11}$$

其中: b 和 b^{gt} 分别代表预测框和 GT 框的中心点, $\rho(b, b^{gt})$ 代表预测框与 GT 框中心点的欧氏距离, α 是平衡参数, c 是预测框和 GT 框的最小外接矩形对角线的长度, v 用来衡量长宽比是否一致^[15]。

CIOU Loss 对预测框与目标框的重叠面积、中心点距离以及长宽比等因素进行了分析, 却没有对实际框与预测框的角度不匹配问题进行分析, 所以我们将 SIOU Loss 作为 Bounding box 回归的定位损失函数。SIOU 损失函数如下所示。

$$SIOU_Loss = 1 - SIOU = 1 - IOU + \frac{\Delta + \Omega}{2} \tag{12}$$

$$\Delta = 1 - 2 * \sin^2 \left(\arcsin(x) - \frac{\pi}{4} \right) \tag{13}$$

$$\Delta = \sum_{t=x,y} (1 - e^{-(2-\Delta)\rho_t}) \tag{14}$$

$$\Omega = \sum_{t=w,h} (1 - e^{-\phi_t})^\rho \tag{15}$$

在角度成本增加的前提下, 一方面可以使损失函数更完整地表示出来, 另一方面可以减小惩罚项为 0 的可能性, 使损失函数的收敛性更好, 从而有效降低了损失的总自由度, 提高了推理的精度^[16-18]。

3 实验与分析

3.1 实验环境

PyTorch 是一种以 Python 类库为基础的开放源代码框架, 它拥有 GPU 的强大加速能力, 是当前最主要的机器学习框架之一。所以, 采用一种基于 PyTorch 的坑洼检测模

型。本文实验环境为 Windows10 操作系统, 12 G 内存, CPU: Intel (R) Xeon (R) Platinum 8255C CPU @ 2.50 GHz, GPU: RTX 3070 (8 GB), 在 Python 3.8、PyTorch 1.7.0、CUDA 11.0 和 PyCharm 2021.3 下实现模型的训练以及改进工作。

3.2 数据集

利用百度的街景地图来进行部分道路上的坑洼图片的采集, 利用互联网的爬虫技术, 从已有的公共数据集中收集到了一些坑洞的照片, 经过仔细的筛选, 然后利用 Python 库的“Labeling”, 对所得到的数据进行人工标记, 完成本文数据集的建立。数据集包括了沥青和混凝土路面, 包括了不同大小的坑洞、不同角度和距离的拍摄, 还包括了重叠、阴影和遮挡等情况, 一共有 1 200 张, 按 8: 1: 1 分为训练集、验证集和测试集。

3.3 评价标准

本文使用准确率 (P, Precision)、召回率 (R, Recall) 和平均精度均值 (mAP, mean average precision) 对模型的改进效果进行度量, 用每秒传输帧数 (FPS) 用作评估速度的指标^[19-21]。以上的评价指标计算公式如下所示:

$$P = \frac{TP}{TP + FP} \quad (16)$$

$$R = \frac{TP}{TP + FN} \quad (17)$$

$$AP = \int_0^1 p(R) dR \quad (18)$$

$$mAP = \frac{1}{n} \sum_{i=1}^n AP \quad (19)$$

TP 代表了被准确地分为正例的数量, 也就是, 样本是正的, 预测结果也是正的。FP 代表了一个被误判为正例的数量, 即一个负的样本和一个正的预测结果。FN 代表了一个被误判为负的样本的数目, 也就是说, 一个正的样本和一个负的预测。其中, 准确度是指在评价过程中, 正样本所占检测样本的比重, 而召回率则是指检测出的正样本占所有正样本的比重。

3.4 模型训练结果与分析

首先, 将收集整理的数据集分为训练集、验证集、测试集 3 种数据文件。在配置的参数下, warmuo_momentum 为 0.8, iou 为 0.2, box 为 0.05, cls 为 0.5, cls_pw 为 1.00, obj 为 1.00, anchor 为 4.1, weight_decay 为 0.000 6, warmup_epochs 为 4.0, NMS 为 5。将 train.py 在 Anaconda 建立的虚拟环境中运行, 经过 300 个循环后, 将最佳的训练模式保存下来。对 YOLOv7 模型展开训练, 首先, 将样本数据通过 Mosaic 输入端进行数据增强处理, 之后, 使用自适应锚框计算, 并将原始样本图片统一缩放到 640×640 尺寸, 之后, 进入主干网络进行切片操作。接下来, 利用改进的加强特征提取网络 BiFPN, 提取样本图像中深层次的特征信息。BiFPN 使用 Fast normalized fusion 来对带权特征进行融合。为了提高计算结果的稳定性, 我们采用了 ReLU 函数。为了进一步提高深度学习模型

的检测效率, BiFPN 采用可分离卷积融合特征, 每次经过卷积后添加批次归一化和激活。

本文将改进后的 YOLOv7 算法与原来的算法进行对比, 在这篇文章中, 训练输入图像的尺寸是 (640, 640), 迭代次数 300, 批处理大小是 4, 模型最大学习率为 1e-2, 最小学习率为最大学习率的 0.001, 优化器为 SGD, 动量系数为 0.937, 学习率下降方式为 cos。图 5 为训练过程中的损失对比曲线, 其中, 横坐标代表迭代次数, 纵坐标代表损失值。从图中可以看出前 100 个迭代损失值下降都较快, 第 101 轮次迭代的 YOLOv7 和本文算法的损失值为 0.138 9 和 0.121 5; 通过对 YOLOv7 的迭代结果进行分析, 得出了 YOLOv7 的迭代损失值稳定为 0.036 7, 本文算法稳定在 0.031 16。在此基础上, 采用 SIoU 损失函数对模型进行改进, 使模型的回归精度得到改善, 同时也加快了模型的收敛性, 取得了较好的训练结果。

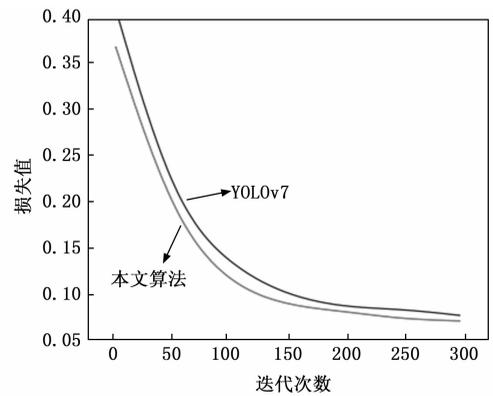


图 5 loss 对比曲线

结果表明, 本文提出的算法与 YOLOv7 在准确率、平均准确度等方面相比有明显优势。图 6 是精度对比曲线, 可以看出, 早期精度提高得很快, YOLOv7 与本算法的精度分别达到了 88.16% 和 94.07%。图 7 表示平均精度均值对比曲线, 同样在前期的 mAP 提升较快, 通过分析结果表明, 该算法的 map 分别为 83.24% 和 96.71%。结果表明, 该方法的准确度和 mAP 都得到了比较显著的提高。图 8 为对路面坑洼检测效果图, 图中可以看出, 明显的坑洼都被检测了出来, 具有良好的检测效果。

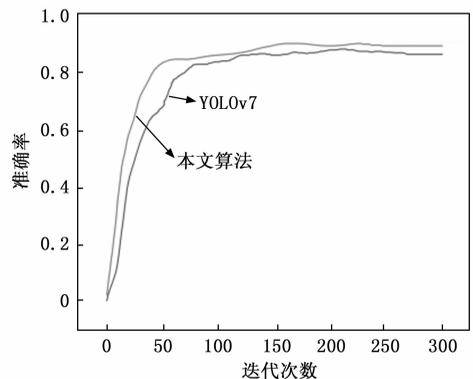


图 6 P 对比曲线

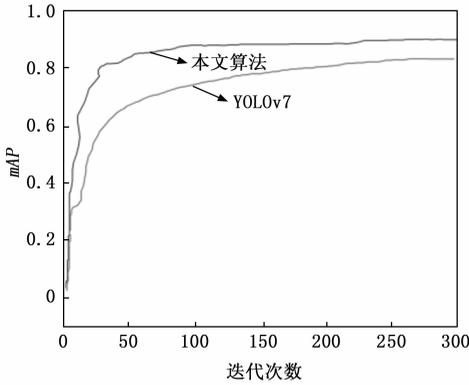


图 7 mAP 对比曲线

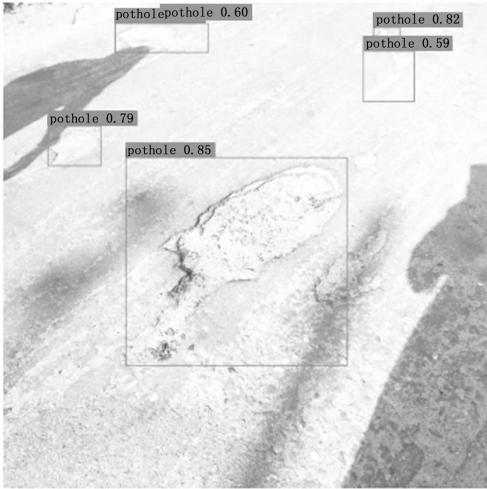


图 8 检测效果图

3.5 实验对比分析

本文将重点研究 YOLOv7 和不同改进点对模型性能的影响，并将该算法和已有的算法进行对比。检测性能对比如表 1 所示，改进 1 进入 CA 注意力机制，改进修改 SIOU 损失函数，改进 3 引入 BiFPN 加权特征金字塔，表中“√”表示使用该改进。

表格首先不添加任何改进使用原本模型进行实验，实验结果准确率为 88.16%，平均精度为 83.24%。改进 1 添加了 SIOU 损失函数，与原模型进行对比，结果精度提升了 1.22%，平均精度提升了 0.78%，参数量下降了 0.250 M，有一定的改进效果；改进 2 将 FPN 替换为 BiFPN，与 YOLOv7 进行对比发现，准确率提升很大，有 4.3% 的改进，平均精度也有所提高，参数量相比略高一点；改进 3 采用了 CA 注意力机制，并保持其他不变，与原模型以及其他改进进行对比发现，准确率相比较改进 1 和改进 2 有所下降，召回率有所提高，参数量也随之减少。通过把改进 1、改进 2、改进 3 一起融合之后进行实验验证表明，与原本的 YOLOv7 模型相比较，整体效果均有提升。

在表 1 中，通过 YOLOv7 算法、改进 1、改进 2、改进 3 相对比，以上都能提高目标检测的性能，BiFPN 会使参数

量增加 0.459 M 参数量，CA 注意力机制不会引入额外参数；与改进前比较，YOLOv7 算法的准确率、召回率和 mAP 分别提高 5.91%、20.45% 和 13.47%。试验结果显示，本算法可以有效地提高网络的检测性能。

表 1 消融实验

算法	损失函数	BiFPN	注意力机制	Precision / %	Recall / %	mAP / %	Parameters
YOLOv7	×	×	×	88.16	72.83	83.24	38.275
改进 1	√	×	×	89.38	72.83	84.02	38.025
改进 2	×	√	×	93.46	77.72	88.70	38.734
改进 3	×	×	√	88.82	77.72	88.12	38.078
本文算法	√	√	√	90.12	79.35	89.42	38.395

为了进一步验证本文算法的检测性能，将本文算法与其他目标检测算法 SSD、FasterR-CNN、YOLOv5、YOLOv7 进行对比，并保持所有原始数据不变，采用 mAP、FPS 为评价标准。在表 2 中列出了指数。

表 2 与现有算法对比

现有算法	输入尺寸	mAP / %	FPS
SSD	300×300	60.93	28.1
FasterR-CNN	600×600	65.56	6.1
YOLOv4	640×640	73.21	14.9
YOLOv5	640×640	78.73	38.7
YOLOv7	640×640	83.24	27.5
本文算法	640×640	89.42	25.3

从表 2 可以看出，本文算法相较于其他算法有着较好的检测速度以及较高的平均精度值，参数量也较小，便于程序的快速运行，综上，本文算法具有更准确快速的检测性能。

4 结束语

在路面坑洼检测方面，拟采用改进的 YOLOv7，采用 BiFPN 加权特征提取网络，提高路面坑洼检测的精度，并将 CA 注意力机制引入到骨干网络中，高感受野提升检测网络中的特征层次。替换 SIOU 损失函数，有效解决回归中的样本不平衡问题，提高了模型的泛化性。通过对多种检测指标的测试，YOLOv7 算法的准确率可达 90.12%，平均检测速率可达 25.3 帧，对路面缺陷的准确定位及辨识效果都有明显的提升。在未来工作计划中，进一步优化模型结构，并将其嵌入到嵌入式器件中，以期为更轻量、更效率的道路坑洼检测提供技术支持。

参考文献:

[1] 陈鹏, 应骏. 基于卷积神经网络的多场景道路坑洼图像检测 [J]. 上海师范大学学报 (自然科学版), 2020 (1): 96-101.
 [2] 邵延华, 张铎, 楚红雨, 张晓强, 等. 基于深度学习 YOLO 目标检测综述 [J]. 电子与信息学报, 2022, 44 (10): 3697-3708.

(下转第 43 页)