

# 基于改进 YOLOv5 的桥梁裂缝模型研究

郭佳佳<sup>1,2</sup>, 董增寿<sup>3</sup>, 常春波<sup>3</sup>

1. 太原科技大学 车辆与交通工程学院, 太原 030024;
2. 湖北交投智能检测股份有限公司, 武汉 430051;
3. 太原科技大学 电子信息工程学院, 太原 030024

**摘要:** 桥梁裂缝人工检测耗时费力、安全性不高, 为了高效、准确、无接触地对桥梁裂缝进行识别检测, 提出一种基于改进 YOLOv5 的桥梁裂缝检测模型 YOLOv5-SA; 该方法在 YOLOv5s 模型的基础上, 首先对收集的数据集利用几何变换、光学变换等操作进行数据增强; 其次将融合视觉注意力机制 (SKNet) 添加到 Head 部分来提高模型对裂缝特征的代表能力; 最后在金字塔特征表示法 (FPN) 的基础上利用自适应空间特征融合 (ASFF) 模块加强网络特征融合能力, 增加对桥梁裂缝小目标的检测; 结果表明: 改进后的模型相对于 YOLOv5s 模型能更好地抑制非关键信息, 减少背景中的无效信息干扰, 提高桥梁裂缝目标检测精准度; 改进后的 YOLOv5-SA 模型准确率达到 88.1%, 与原 YOLOv5s 模型相比提高了 1.6%; 平均精度均值  $mAP$  0.5 和  $mAP$  0.5~0.95 分别达到 90.0%、62.1%, 相比而言分别提高了 2.2%、2.4%; 与其他桥梁裂缝检测相关方法 (Faster-RCNN、YOLOv4tiny) 相比, 提出的 YOLOv5-SA 模型也具有相当或更好的检测性能; 由此可见改进后的模型能更高效地检测复杂环境下的桥梁裂缝, 可以为工业检测提供一部分思路。

**关键词:** 桥梁裂缝; 目标检测; 注意力机制; YOLOv5; 特征融合; 图像处理

## Research on Bridge Crack Model Based on Improved YOLOv5

GUO Jiajia<sup>1,2</sup>, DONG Zengshou<sup>3</sup>, CHANG Chunbo<sup>3</sup>

1. School of Vehicle and Traffic Engineering, Taiyuan University of Science and Technology, Taiyuan 030024, China;
2. Hubei Communications Investment Intelligent Detection Co., Ltd., Wuhan 430051, China;
3. School of Electronic Information Engineering, Taiyuan University of Science and Technology, Taiyuan 030024, China)

**Abstract:** There are the characteristics of time-consuming and laborious, and the safety in bridge crack manual detections. In order to identify and detect bridge cracks efficiently, accurately and without contact, a bridge crack detection model YOLOv5-SA based on improved YOLOv5 is proposed. Based on the YOLOv5s model, firstly, the collected data set is enhanced by the geometric transformation and optical transformation. Secondly, selective Kernel networks (SKNet) were added to the Head part to improve the representation ability of crack features. Finally, on the basis of pyramid feature notation (FPN), adaptively spatial feature fusion module was used to strengthen the network feature fusion ability, and to increase the detection of small targets for bridge cracks. The results show that compared with the YOLOv5s model, the improved model can suppress non-critical information better, reduce the invalid information interference in the background, and improve the accuracy of bridge crack target detection. The accuracy of the improved YOLOv5-SA model reaches 88.1%, which is 1.6% higher than that of the original YOLOv5s model. The average accuracy of  $mAP$  0.5 and  $MAP$  0.5~0.95 reached 90.0% and 62.1%, respectively, which increased by 2.2% and 2.4%, respectively. Compared with other methods related to the bridge crack detection (Faster-RCNN, YOLOv4tiny), the proposed YOLOv5-SA model also has a comparable or better detection performance, therefore, the improved model can detect bridge cracks in complex environments more efficiently, which can provide some ideas for industrial detections.

**Keywords:** crack of bridge; object detection; mechanism of attention; YOLOv5; fusion of features; image processing

## 0 引言

混凝土常用于道路、桥梁、隧道、港口和大坝等工业工程。工业工程建造及使用过程中, 必然会出现一系列病

害<sup>[1]</sup>。裂缝是混凝土桥梁在自然过程中不可避免的问题, 它会引起整座桥梁的严重破坏。裂缝成因复杂, 危害性大, 形状、走向无规律性, 一般呈现多种形态和大小<sup>[2]</sup>。裂缝缺陷检测和修复不及时, 可能会造成桥梁重大经济损失以

收稿日期: 2023-02-03; 修回日期: 2023-03-10。

基金项目: 山西省基础研究计划(自由探索类)面上项目(202303021211205)。

作者简介: 郭佳佳(1998-), 女, 硕士研究生。

常春波(1979-), 男, 硕士, 副教授。

通讯作者: 董增寿(1970-), 男, 博士, 教授。

引用格式: 郭佳佳, 董增寿, 常春波. 基于改进 YOLOv5 的桥梁裂缝模型研究[J]. 计算机测量与控制, 2023, 31(12): 188-194.

及威胁生命安全。因此, 早期发现和确定桥梁裂缝, 以便进行精确的养护, 是特别重要的。到目前为止, 桥梁裂缝评估多由经验丰富的检查员利用桥检车及超声波、电磁波等各种无损检测技术实施的, 但是该方法检测效率低, 人工成本高<sup>[3]</sup>。深度学习作为近年来的热点和有效的监督学习方法之一, 为桥梁裂缝检测带来了很大的提升, 深度学习具有很强的特征提取和泛化能力, 以及高鲁棒性和可靠性。

一般来说, 裂纹识别算法可以分为手工特征工程和自动特征工程。手工制作的特征工程方法通常包括特征检测算法、边缘检测、特征变换算法、形态学算法等。早期, Sobel 边缘检测、Canny 边缘检测、傅里叶变换、中值滤波、数学形态学和直方图分析等多项图像处理技术被应用于混凝土桥梁表面裂缝评估<sup>[4-5]</sup>。此外, 许多传统机器学习检测采用滑动窗算法确定潜在裂纹的位置, 并采用人工设计的提取器进行特征提取。在面对复杂环境时 CrackTree<sup>[6]</sup>和支持向量机 (SVM, support vector machine)<sup>[7]</sup>的检测方法很难有效构造出一个统一的描述子, 因此, 它们的适应性和重用性都存在一定的局限。相反, 参数量多、鲁棒性好的卷积神经网络 (CNN, convolutional neural network) 和 YOLO (you only look once) 系列模型在图像处理中更具优势。Torres 等人<sup>[8]</sup>比较了神经网络 (NN) 分类器、贝叶斯分类器和传统的支持向量机 (SVM) 分类器, 总结了 CNN 在桥梁桥面缺陷分类方面的潜力。CNN 不需要手工创建规则, 能够自动提取多级特征表示。CNN 从图像中提取空间视觉特征 (从低级特征到高级特征) 方面具有强大的能力, 从而提高了计算机视觉任务的速度和精度。Cha 等人<sup>[9]</sup>使用 CNNs 检测混凝土裂缝, 并使用滑动窗口技术对任意大小的裂缝图像进行扫描, 验证了 CNNs 在区分大量类别方面的可靠性能, 但是此方法耗时较长, 存储的数据会逐渐增大。Kim 等人<sup>[10]</sup>提出了一种基于 R-CNN 的裂纹检测识别方法, 利用一个方形标记物来测量被检测裂纹的大小, 但此方法对形状多变、走向无规律的裂纹识别检测存在一定的局限性。Liang 等人<sup>[11]</sup>采用了一种 Faster R-CNN 算法, 利用边界框来自动检测钢筋混凝土桥梁系统的结构构件, 然而此方法不利于小物体及多尺度的物体检测。Huang 等人<sup>[12]</sup>利用 Mask RCNNs 改进架构对盾构隧道衬砌图像进行裂缝检测, 采用架构中的形态闭合操作将不相交的裂缝连接起来, 并且使用 4 个训练好的模型来探讨网络设计和预期性能的影响, 以选择最优模型, 但模型较繁琐, 检测时间长, 不适用于工业检测。Liu 等人<sup>[13]</sup>将 U-Net 应用于混凝土裂缝的像素级分割, 通过与 FCN 的对比, 证明 U-Net 在裂缝分割方面具有较高的精度, 但是模型复杂冗余太大, 导致网络训练很慢。为了克服两阶段模型的繁琐步骤以及提高检测速度, Redmon 等人<sup>[14]</sup>2016 年提出了 YOLO 模型, 采用回归的概念, 去掉候选框阶段, 使用 Anchor-Free 概念实现一级检测, 通过网络立即输出类别和位置信息, 实现端到端学习, 大幅提高了检测速度, 使物体检测网络真正能够

实时检测。此外, 在 YOLO 的基础上, Redmon 等人<sup>[15]</sup>融合进 YOLO 提出 YOLOv2 模型, 使用了一些优化的方法来提高平均准确率和检出率。Li 等人<sup>[16]</sup>结合深度可分离卷积、特征金字塔和 YOLOv3 减少参数, 有效提取了裂纹特征。Yu 等人<sup>[17]</sup>在 YOLOv4 的基础之上利用焦点损失优化损失函数, 提高了精度, 克服了复杂背景的挑战。YOLOv5 算法<sup>[18]</sup>不仅保持了前几代 YOLO 算法的检测速度, 而且灵活性更高, 鲁棒性和泛化能力更好, 大大提高了精度, 在小物体的检测上也有更大的优势。模型的速度比两级模型快得多, 更适合工业检测。此外, 端到端网络结构可以在一定程度上减少裂纹特征信息的损失。

上述方法对裂缝检测具有重要意义, 但仍存在问题, 由于池化层的原因, 卷积神经网络在增加上层卷积核的接收野和聚合图像背景时, 丢失了部分信息。通过将浅层特征和深层特征相结合, 提取出更有代表性的特征, 但会忽略掉一部分更小的特征。缺乏对复杂背景图像下裂缝关键特征的筛选能力, 容易出现过拟合现象, 导致模型识别效率较低。不能充分利用不同尺度特征来达到更好的特征融合, 出现误检漏检现象, 精度和速度之间不能很好的达到平衡。此外, 大多数研究侧重于单调背景中的裂缝检测, 例如实验室中的混凝土构件或没有杂物的桥面, 对复杂背景中的裂缝缺乏良好的检测能力, 难以在实际项目中使用。

为了解决这些问题并实现自然场景中桥梁裂缝的实时准确检测, 提高模型的检测精度, 提出了一种基于改进 YOLOv5 的桥梁裂缝检测模型, 在 YOLOv5s 模型的基础上, 增加注意机制来有效提取检测目标的关键特征, 并增加新的特征融合层来捕捉小物体的信息。改进后的模型结合 SKNet (selective Kernel networks) 和自适应空间特征融合模块, 突出识别对象的局部特征, 过滤关键特征, 增加特征融合, 提高识别精度和识别效率。该实验模型可实现复杂背景下桥梁裂缝的快速检测, 为裂缝的识别与检测提供理论依据, 为后续裂缝的养护奠定基础。

## 1 YOLOv5 模型

YOLO 作为一种用于多目标检测的深度学习框架, 模型体积小, 计算速度快, 将检测问题转化为回归问题, 整个图作为网络的输入。YOLOv5 网络是由 YOLOv4 和 YOLOv3 演化而来的一种典型的单级目标检测算法, 检测精度高, 实时性好。YOLOv5s 网络结构分为输入端 (Input)、骨干端 (Backbone)、颈部 (Neck) 和检测端 (Head)。网络模型如图 1 所示。

YOLOv5s 的输入端主要是对输入图像数据进行预处理, 包括马赛克数据增强以及自适应锚盒计算, 并将图像缩放到  $640 \times 640$  的网络输入大小。马赛克数据增强中算法随机抽取  $N$  个训练样本, 随机放大、缩小、旋转、拼接。数据增强会使训练样本更加多样化, 训练模型也更加稳健。自适应锚盒计算中网络在训练阶段输出预测框, 通过计算预测框与地面真值框的差值来反向更新和优化网络参数,

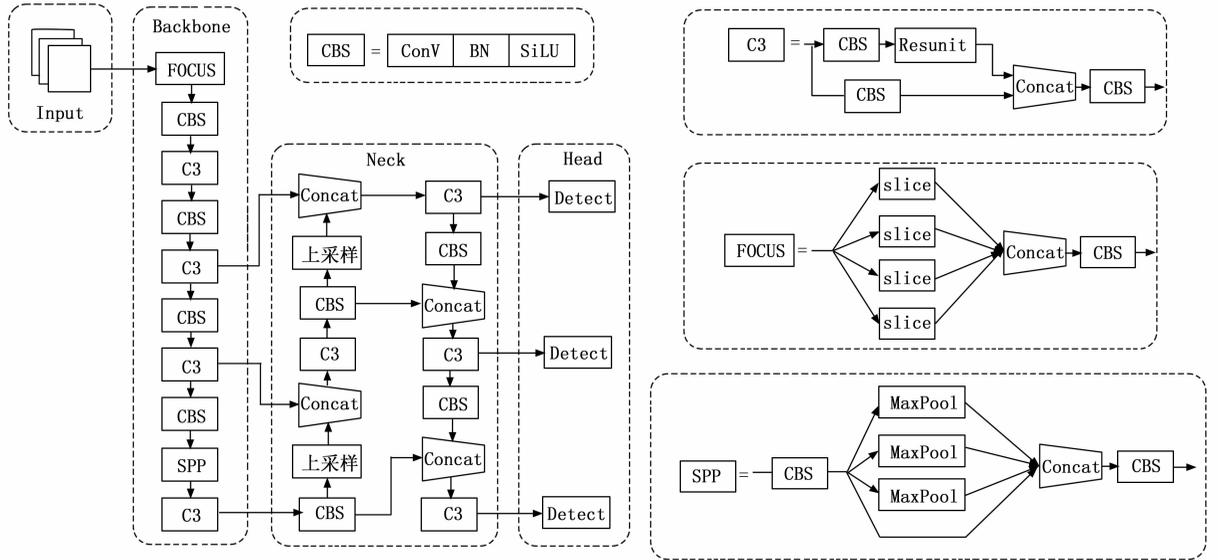


图 1 YOLOv5s 网络结构

显著提高了训练和预测速度。骨干部分负责特征提取，包括 Focus、C3 和空间金字塔池（SPP）结构，Focus 结构主要用于切片操作，使得特征提取足够充分，该方法减少了原始信息和计算量的损失，提高了训练速度。C3 块第一部分执行卷积操作，第二部分在卷积完成后直接与第一部分执行 concat，减少了计算量，提高了网络的学习能力。空间金字塔池<sup>[19]</sup>（SPP, spatial pyramid pooling）可以增加感知场，有助于解决锚点和特征图的对齐问题，可以显著提高特征提取效率，SPP 原理如图 1 所示。颈部网络采用了 FPN (feature pyramid networks)<sup>[20]</sup> 和 PAN (path aggregation network)<sup>[21]</sup> 特征金字塔结构，这两种结构共同增强了颈部网络的特征融合能力。检测端对不同尺度的特征信息进行融合，输出一个包含目标对象的类别概率、对象得分和该对象的边界框位置的向量。

## 2 YOLOv5-SA 模型

### 2.1 SKNet

在目标检测任务中，注意力机制主要作用于特征映射，通过先后生成通道和空间两个独立维度的注意映射，生成二维注意评分映射并应用到输入特征映射上，最后自适应特征细化。注意力机制<sup>[22]</sup>由 Borji 和 Itti 在 2013 年提出，此方法来源于对人类视觉的模仿。人眼擅长识别复杂图像或从复杂背景中提取典型特征。因此，随着深度学习的蓬勃发展，注意力机制可以应用到机器视觉中。根据可微性，该机制可以有硬注意或软注意。软注意是可微的，因此更有利于迭代计算注意权重，其相关研究也更为广泛。Wang 等人<sup>[23]</sup>在 2017 年提出了一种剩余注意网络，以掩码的方式为每个特征元素添加相应的注意权重，将注意力转移到图像中最重要的区域而忽略不相关部分，有助于更有效地提取关键特征。

由于桥梁所处环境比较复杂，数据集图像中部分非目

标因素与目标裂缝相似度较高且裂缝尺度多变。为了提高模型对裂缝特征的识别能力，将 SKNet 引入到 YOLOv5s 的 Head 部分。其中，SKNet 中引入的不同卷积核分支，会产生不同大小的感受野便于网络侧重桥梁裂缝图像多种尺度下的主要尺度特征，并显示出不同通道间的联系和每个通道的重要程度并赋予权重，以便于加强特征图中重要裂缝特征的表达能力，抑制不必要背景特征的表达能力，使卷积网络模型更加关注待检测裂缝的特征，让模型在定位和识别目标区域时更加准确。特色之处体现在通过网络自己学习来选择融合不同感受野的特征图信息，有利于不同尺度的裂缝识别。因此，所提取的特征在可见的裂缝实例中具有更大的信息量，提高了识别精度。

SKNet<sup>[24]</sup>是由 Li 等人首次提出的一种轻量级嵌入式模块，结构如图 2 所示，由 Split、Fuse、Select 三部分组成。Split 部分对维度为  $C \times H \times W$  特征图  $X$  进行  $3 \times 3$ 、 $5 \times 5$  的完整卷积操作，进而得到相同维度的不同特征图和。Fuse 部分对两者进行信息融合操作得到同维度，进而通过全局平均池化与全连接层获取特征图的注意力信息，从而创建一个紧凑的特征，以便为精确和自适应地调整感受野的大小提供指导。Fuse 部分计算公式如下所示：

$$U = U_1 + U_2 \quad (1)$$

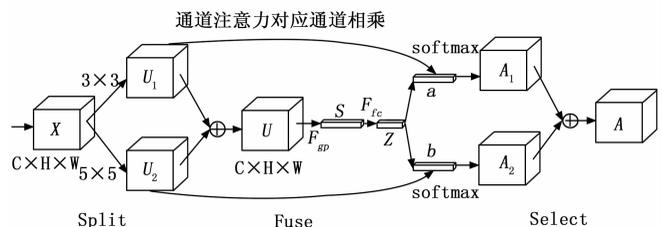


图 2 SKNet 结构

将两部分的特征图按元素求和得到  $U$ 。

$$s = F_{sp}(U) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W U(i, j) \quad (2)$$

维度为  $C \times H \times W$  (长  $\times$  高  $\times$  宽) 的特征图  $U$  通过空间维度的全局平均池化操作生成通道统计信息, 得到维度为  $C \times 1$  特征图  $S$ 。

$$z = F_{fc}(s) = \delta(B(W_s)) \quad (3)$$

经过先降维再升维的  $FC$  全连接层生成紧凑的特征  $Z$  (维度为  $d \times 1$ ,  $d < C$ ),  $\delta$  是 ReLU 激活函数,  $B$  表示批标准化 (BN),  $Z$  的维度为卷积核的个数,  $W$  维度为  $d \times C$ ,  $d$  代表全连接后的特征维度。

$$a_c = \frac{e^{A_c z}}{e^{A_c z} + e^{B_c z}}, b_c = \frac{e^{B_c z}}{e^{A_c z} + e^{B_c z}} \quad (4)$$

Select 部分中特征与不同权重卷积核进行卷积, 再经过 softmax 操作计算后得到新特征图。 $a$ 、 $b$  是 soft attention 向量, 卷积核  $a_c$  与  $b_c$  和为 1。 $A$  和  $B$  的维度均为  $C \times d$ ,  $A_c$ 、 $B_c$  为  $A$ 、 $B$  的第  $c$  行的数据。维度为  $1 \times d$  进而分别乘以对应的  $U_1$  和  $U_2$  特征图得到  $A_1$  和  $A_2$ , 最后的模块  $A$  同样由信息融合得到。经过一系列操作, 提高了卷积核的感受野, 得到了更多的信息, 增强网络对本质特征的注意力, 增强特征图的表现力, 有效提高网络的检测精度, 同时略微增加计算工作量和复杂度。

## 2.2 自适应空间特征融合 (ASFF)

不同特征尺度之间的不一致是单级检测器金字塔特征表示法 (FPN) 的主要限制, 但自适应空间特征融合 (ASFF, adaptively spatial feature fusion) 方法<sup>[25]</sup>可以在空间上过滤冲突信息以解决梯度反传不一致问题, 改善了特征比例不变性, 且推理时间也有所降低。

桥梁裂缝检测过程中, 图像的背景复杂, 这引入了相当大的背景噪音, 且裂缝形状和大小的规格并不固定, 目标规格差异, 给检测带来了一定的困难。裂缝检测模型需要一个强大的特征融合模块。然而, 原 YOLOv5s 模型中 FPN-PAN 结构存在多尺度特征融合不足的问题, 一般的特征融合方法虽可以丰富整体的特征信息, 但不同维度之间往往存在预测冲突, 因此在特征融合部分增加 ASFF 结构, 自主学习每个尺度的空间权重, 实现图像多尺度特征的充分融合。ASFF 结构如图 3 所示。

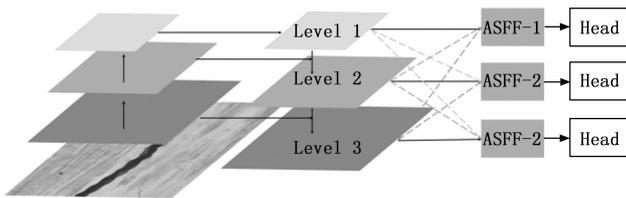


图 3 ASFF 结构

YOLOv5s 网络中的 Neck 部分输出得到 3 个特征层 Level11、Level12、Level13。权重参数  $\alpha$ 、 $\beta$  和  $\gamma$  与不同特征层相乘之后相加得到新的融合特征 ASFF-1、ASFF-2、ASFF-3。ASFF-1 为例, 公式如下所示:

$$y'_{ij} = \alpha'_{ij} \times x_{ij}^{1 \rightarrow l} + \beta'_{ij} \times x_{ij}^{2 \rightarrow l} + \gamma'_{ij} \times x_{ij}^{3 \rightarrow l} \quad (5)$$

$x_{ij}^l$  表示输出特征  $(i, j)$  向量在通道之间的映射。 $\alpha'_{ij}$ 、 $\beta'_{ij}$ 、 $\gamma'_{ij}$  为网络自适应学习的不同层次到 level  $l$  的空间重要度权重, 经相加再进行 softmax 操作, 可得范围区间为  $[0, 1]$  且和为 1。

$$\alpha'_{ij} + \beta'_{ij} + \gamma'_{ij} = 1 \quad (6)$$

$$\alpha'_{ij} = \frac{e^{\lambda'_a}}{e^{\lambda'_a} + e^{\lambda'_b} + e^{\lambda'_c}} \quad (7)$$

其中:  $\lambda'_a$ 、 $\lambda'_b$ 、 $\lambda'_c$  为控制参数, 上一步相加操作需保证通道数及特征大小相同, 因此利用  $1 \times 1$  卷积层从  $x^{1 \rightarrow l}$ 、 $x^{2 \rightarrow l}$ 、 $x^{3 \rightarrow l}$  计算  $\lambda'_a$ 、 $\lambda'_b$  和  $\lambda'_c$  保证通道数相同, 插值的方式保证特征大小一致, 进而得到  $\{y_1, y_2, y_3\}$ , 此外, 每个尺度上对各个层次的特征均进行自适应聚合。

## 3 实验和实验结果分析

### 3.1 数据集构建

由于桥梁裂缝检测没有公开可访问的数据集, 为了验证提出的目标检测模型, 利用 CCD 高清数码相机收集了太原市内中小型钢筋混凝土施工桥梁表面不同位置的裂缝图像数据。为了数据集的完整性, 数据图像包括不同光照、天气条件和遮挡阴影下的缺陷图像, 共计筛选出 1 500 张照片。经过数据相似度比较, 将收集到的桥梁裂缝图像与优秀论文<sup>[26]</sup>中部分质量较好裂缝图像合并, 共计 1 000 张。此外, 使用开源无损图像处理应用程序对图像进行规范化, 以辅助这项工作, 同时为了使目标检测模型更加有效, 通过几何变换、光学变换等数据增强技术丰富桥梁裂缝数据集, 完成裂缝检测数据集的构建, 最终共计得到 5 000 张桥梁裂缝图像, 按 7: 3 分训练集、验证集。

采用 python 版本的开源数据标注工具 Labelimg 制作 txt 格式的数据集标签文件, 每张图像都生成对应的标注信息文件。数据集标签的制作需要按照最小外接矩形、重点框取裂缝内容、单个框内的仅包含一条裂缝、框与框之间尽量避免重叠的原则。裂缝是研究对象, 在标注时需要框选出所有肉眼可见的裂缝, 包括一些细小裂缝, 命名桥梁裂缝标签为 crack。

### 3.2 实验环境

模型的训练与测试都在同一设备进行, 用于模型训练和测试实验的软硬件工具如下: 处理器 AMD Ryzen 9 3950X, 内存 64 GB, 系统 Windows10, 开发环境 torch1.7.0。根据服务器内存和显卡配置, 设迭代次数 (epoch) 200 来评估指标和可视化效果, 批次 (batch size) 为 16, 初始学习率为 0.01, 最小学习率为 0.001。

### 3.3 评价指标

在目标检测中, 评价模型性能的标准有很多。采用准确率  $P$  (precision)、召回率  $R$  (recall)、平均精度均值  $mAP$  (mean average precision) 和  $F_1$  共 4 个性能指标来评价模型。准确率用来衡量模型检测的性能, 召回率用于评估检测的全面性, 两者的综合为平均精度 (AP)。召回率计

算是所有被正确预测的目标的比例, (AP) 计算则指利用不同的 Precision 和 Recall 的点的组合, 得到的 P-R 曲线与  $x$  轴和  $y$  轴围成的面积,  $mAP$  是 AP 的均值, 用来衡量整个模型的性能。  $mAP_{0.5}$  为 IoU 设置为 0.5 时所有类别 AP 的平均值,  $mAP_{0.5 \sim 0.95}$  为不同 IoU 阈值下的平均值。由于  $P$  和  $R$  是相互矛盾的性能指标, 所以  $F_1$  是两者的调和平均值, 取值范围为  $(0, 1)$ , 同时使用两者评价模型的质量。各指标计算公式如所示:

$$Precision = \frac{TP}{TP + FP} \times 100\% \quad (8)$$

$$Recall = \frac{TP}{TP + FN} \times 100\% \quad (9)$$

$$F_1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (10)$$

$$AP = \int_0^1 P(R) dR \quad (11)$$

$$mAP = \frac{\sum_{i=1}^C AP(c)}{C} \% \quad (12)$$

其中:  $TP$  (true positives) 代表正确识别的正例样本,  $FN$  (false negatives) 为识别错误的正例样本,  $FP$  (false positives) 表示识别错误的负例样本,  $C$  代表目标检测对象的类别数, 模型检测仅涉及裂缝一种类别,  $C$  取值为 1。

### 3.4 实验结果分析

与传统检测模型相比, 单级物体检测模型具有较高的精度和速度, 在工业检测领域日益具有优势, 在本节中, 使用相同的数据集和训练参数对 YOLOv5-SA 模型和其他主流的目标检测模型 (YOLOv5s、YOLOv4tiny、Faster-RCNN) 进行训练, 给出了相应的讨论, 验证改进后的 YOLOv5-SA 模型的优势。各模型的训练损失曲线如图 4 所示, 模型的准确率、召回率、 $mAP_{0.5}$  以及  $mAP_{0.5 \sim 0.95}$  如图 5 所示。各检测模型的具体结果值如表 1 所示。

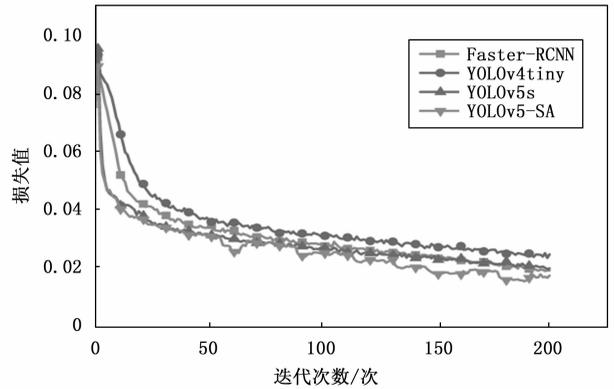


图 4 损失曲线

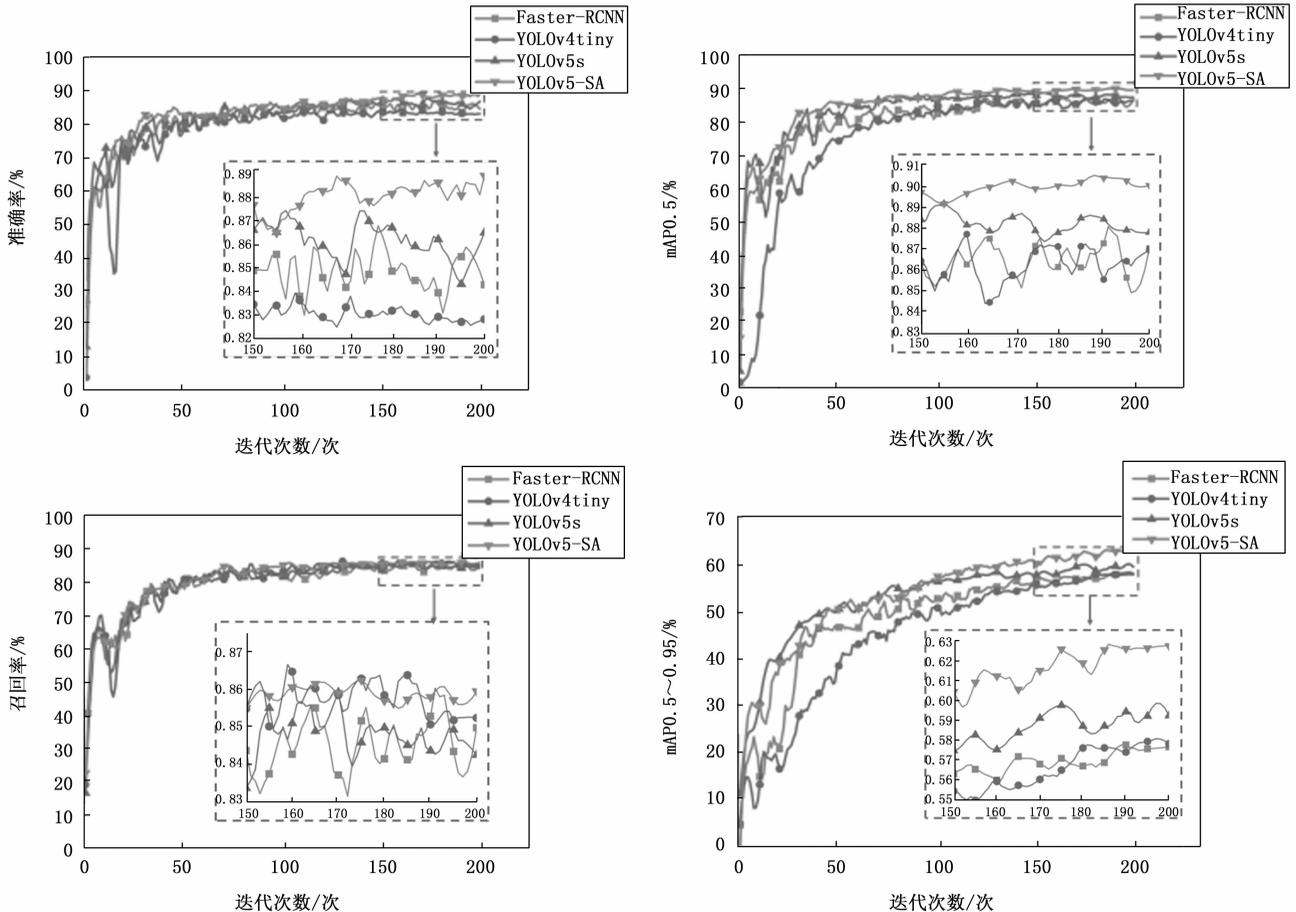


图 5 各模型对比图

从图 4 中可以看出, 在前 50 次迭代中, 各模型拟合较快, 损失值明显减小。改进后的模型最终稳定在 0.017 附近, 而原始的 YOLOv5s 稳定在约 0.021, 其他模型的稳定值均高于改进后的模型。改进后的 YOLOv5-SA 模型曲线平滑、无明显波动, 此外模型的损失函数值下降得更快, 损失值曲线低于原始 YOLOv5s 模型以及其他模型, 表明改进后的模型在特征提取方面更加高效, 收敛速度更快, 训练效率更高。

训练过程中准确率、召回率和  $mAP$  的变化趋势如图 5 中所示, 与其他 3 种模型相比, 可以观察到改进后的模型各检测指标正在稳步增加。在相同的数据集下 YOLOv5-SA 网络的收敛速度最快, 模型检验准确率高于原始模型及其他模型。对于  $mAP$  0.5 和  $mAP$  0.5~0.95 从图中可看出, 改进后 YOLOv5-SA 模型中此指标排名最高。总的来说, 注意力机制和自适应空间特征融合 (ASFF) 模块的引入有效提高了模型的检测精度。

表 1 中显示了 4 种模型的具体结果比较, 改进后的 YOLOv5-SA 模型  $mAP$  0.5~0.95 达到 62.1%, 与其他 3 种方法相比, 分别比 YOLOv5s、YOLOv4tiny 和 Faster-RCNN 模型高出 2.4%、4.0% 和 3.6%。 $mAP$  0.5 为 90.0%, 相对其他模型高出 2.2%、3.3% 和 3.0%。YOLOv5-SA 模型的召回率也是高于其他模型。YOLOv5-SA 模型的准确率为 88.1%, 在 4 种方法中最高, 与 YOLOv5s、YOLOv4tiny 和 Faster-RCNN 相比, 该算法分别提高了 1.6%、5.0% 和 2.7%。提出的 YOLOv5-SA 模型  $F_1$  评分为 87.2%, 分别比 YOLOv5s、YOLOv4tiny 和 Faster-RCNN 模型高出 1.5%、3.0% 和 2.2%。显而易见, 各模型在识别桥梁裂缝时的效果低于改进后的模型, 改进后的模型显著提高了检测性能。

表 1 模型结果 %

算法	Faster-RCNN	YOLOv4tiny	YOLOv5s	YOLOv5-SA
$P$	85.4	83.1	86.5	88.1
$R$	84.6	85.3	84.9	86.4
$mAP$ 0.5	87.0	86.7	87.8	90.0
$mAP$ 0.5~0.95	58.5	58.1	59.7	62.1
$F_1$	85.0	84.2	85.7	87.2

为了进一步对比测试改进后的 YOLOv5-SA 模型和原 YOLOv5s 模型对桥梁裂缝的识别能力, 同时也为了能够更直观地对比模型的预测表现以及验证模型的泛化能力。选取 16 张混凝土桥梁裂缝图像, 并且这些图像不属于测试集中的原有照片, 两种模型的测试对比结果如图 6 和图 7 所示, 由两组图的对比可知, 与 YOLOv5-SA 模型相比原有 YOLOv5s 模型中少量小目标裂缝未被检测到, 实验表明改进后模型的漏检现象减少, 模型在复杂场景下的表现有所提升。

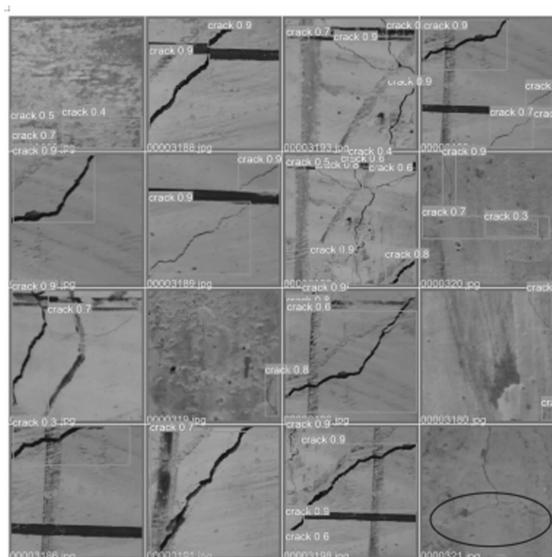


图 6 YOLOv5s 测试图

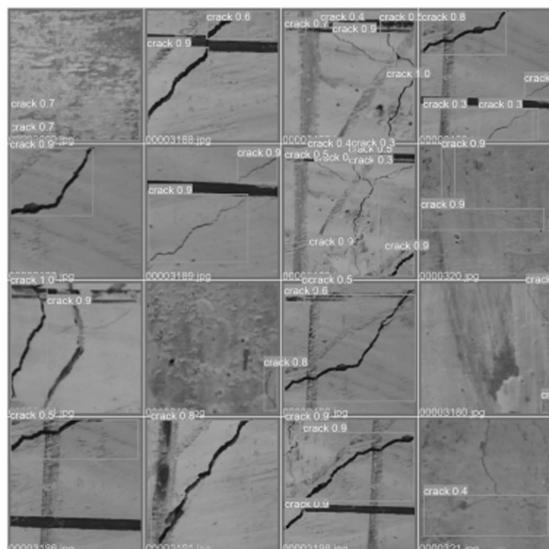


图 7 YOLOv5-SA 测试图

#### 4 结束语

针对桥梁裂缝检测难点, 提出改进的 YOLOv5 桥梁裂缝检测模型。借鉴人的视觉注意机制, 结合 SKNet 优化 YOLOv5s 网络模型, 有效提取检测目标的关键特征, 在特征融合部分结合自适应空间特征融合 (ASFF) 模块, 通过重组其检测端结构, 加强了网络特征融合能力, 拓宽了裂缝检测范围。构建复杂背景下的桥梁裂缝检测模型, 并采用对比实验对改进后的模型进行评估和分析。

结果表明, 改进后的 YOLOv5-SA 模型准确率达到 88.1%,  $mAP$  0.5 以及  $mAP$  0.5~0.95 分别达到 90.0% 和 62.1%, 与原有的 YOLOv5s 模型相比分别提高了 1.6%、2.2%、2.4%。为验证提出模型的优越性, 与两种常用的目标检测模型进行对比, 改进模型的  $mAP$  0.5 相较于 Fas-

ter-RCNN、YOLOv4tiny 分别提高了 3.0%、3.3%，*mAP* 0.5~0.95 分别提高了 3.6%、4.0%。改进后的模型能更有效的提取裂缝特征，有更好的收敛速度，提供了高精度、更可靠的模型识别方法。虽然此方法取得了较好的效果，但仍有一些需要改进的地方，在未来的工作中，将进一步完善数据集，构建自然场景下桥梁裂缝的检测方法，进一步增强模型的鲁棒性，减少环境噪声干扰的影响，并尝试将模型扩展到更多的应用领域。

#### 参考文献:

- [1] YUPENG R, JISHENG H, ZHIYOU H, et al. Image-based concrete crack detection in tunnels using deep fully convolutional networks [J]. *Construction and Building Materials*, 2020, 234: 117367.
- [2] XU Y, LI D, XIE Q, et al. Automatic defect detection and segmentation of tunnel surface using modified Mask R-CNN [J]. *Measurement*, 2021, 178 (4): 109316.
- [3] 王玮华. 基于图像处理技术的桥梁裂缝检测 [D]. 西安: 长安大学, 2013.
- [4] 冯 伍, 张俊兰, 苗秋瑾. 几种典型边缘检测算子的评估 [J]. *电子设计工程*, 2011, 19 (4): 131-133.
- [5] 张红霞, 王 灿, 刘 鑫, 等. 图像边缘检测算法研究新进展 [J]. *计算机工程与应用*, 2018, 54 (14): 11-18.
- [6] ZOU Q, CAO Y, LI Q, et al. Crack tree: automatic crack detection from pavement images [J]. *Pattern Recognition Letters*, 2012, 33 (3): 227-238.
- [7] 王 睿, 漆泰岳. 基于机器视觉检测的裂缝特征研究 [J]. *土木工程学报*, 2016, 49 (7): 123-128.
- [8] ALBERTO T B, ALAIZ C M, DORRONSORO J R. Faster SVM training via conjugate SMO [J]. *Pattern Recognition*, 2021, 111: 107644.
- [9] CHA Y J, CHOI W, BUYUKOZTURK O. Deep learning-based crack damage detection using convolutional neural networks [J]. *Computer-Aided Civil and Infrastructure Engineering*, 2017, 32 (5): 361-378.
- [10] IN-HO K, HAEMIN J, SEUNG-CHAN B, et al. Application of crack identification techniques for an aging concrete bridge inspection using an unmanned aerial vehicle [J]. *Sensors*, 2018, 18 (6): 1881.
- [11] LIANG X. Image-based post-disaster inspection of reinforced concrete bridge systems using deep learning with Bayesian optimization [J]. *Computer-Aided Civil and Infrastructure Engineering*, 2019, 34 (5): 415-430.
- [12] HUANG H, ZHAO S, ZHANG D, et al. Deep learning-based instance segmentation of cracks from shield tunnel lining images [J]. *Structure and Infrastructure Engineering*, 2020, 18 (1/2/3): 183-196.
- [13] LIU X, AI Y, SCHERER S. Robust image-based crack detection in concrete structure using multi-scale enhancement and visual features [C] // 2017 IEEE International Conference on Image Processing (ICIP), Beijing, 2017: 2304-2308.
- [14] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016: 779-788.
- [15] REDMON J, FARHADI A. YOLO9000: better, faster, stronger [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017: 6517-6525.
- [16] LI Y D, HAN Z Q, XU H Y, et al. YOLOv3-Lite: A lightweight crack detection network for aircraft structure based on depthwise separable convolutions [J]. *Applied Sciences*, 2019, 9: 3781.
- [17] YU Z, SHEN Y, SHEN C. A real-time detection approach for bridge cracks based on YOLOv4-FPM [J]. *Automation in Construction*, 2021, 122: 103514.
- [18] ZHU X K, LYU S C, WANG X, et al. TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios [C] // 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), Montreal, BC, 2021: 2778-2788.
- [19] KAIMING H, XIANGYU Z, SHAOQING R, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37 (9): 1904-1916.
- [20] LIN T Y, GIRSHICK B R, HE K, et al. Feature pyramid networks for object detection [C] // Honolulu, HI, USA: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 936-944.
- [21] LIU B, DING Z, TIAN L, et al. Grape leaf disease identification using improved deep convolutional neural networks [J]. *Frontiers in Plant Science*, 2020, 11: 1082.
- [22] BORJI A, ITTI L. State-of-the-art in visual attention modeling [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35 (1): 185-207.
- [23] WANG F, JIANG M Q, QIAN C, et al. Residual attention network for image classification [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017: 6450-6458.
- [24] LI X, WANG W H, HU X L, et al. Selective Kernel Networks [C] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 2019: 510-519.
- [25] LI D J, ZHANG Z L, WANG B G, et al. Detection method of timber defects based on target detection algorithm [J]. *Measurement*, 2022, 203: 103514.
- [26] ZHENWEI Y, TONGGANG S, CHENKAI S. A real-time detection approach for bridge cracks based on YOLOv4-FPM [J]. *Automation in Construction*, 2021, 122 (2): 103514.