

基于注意力机制的 YOLOv5 路面裂缝检测与识别

周双喜^{1,2}, 杨丹², 潘远², 丁建新¹, 丁杨³

(1. 广州航海学院土木与工程管理学院, 广东 广州 510765; 2. 华东交通大学土木与建筑学院, 江西 南昌 330013; 3. 浙大城市学院土木工程学院, 浙江 杭州 310015)

摘要: 针对传统路面裂缝检测实时性差、准确率低的问题, 利用深度学习网络在目标检测方面的优势, 提出了一种改进的 YOLOv5 算法, 本文称为 YOLOv5s-attention, 以实现路面裂缝自动化检测与识别。首先, 对采集到的裂缝图片用 LabelImg 标注软件进行手工标记, 然后通过改进 YOLOv5 网络训练得到网络模型参数。最后, 利用所建立的模型对裂缝进行验证和预测。除此之外, 采用综合评价指标(F1-measure, F1)和平均精度均值(mean Average Precision, mAP)这两个指标来比较原 YOLOv5s、YOLOv5s-attention 模型在路面裂缝上检测与识别的性能。经 YOLOv5s 与 YOLOv5s-attention 比较发现, YOLOv5s-attention 检测准确率(Precision)提高 1.0%, F1 提高 0.9%, 平均精度均值(mean Average Precision, mAP)提高了 1.8%。由此可知, 该网络在实现道路裂缝自动化识别上具有一定的现实意义。

关键词: 道路养护; 路面裂缝; 目标检测; YOLO; 注意力机制; 图像处理

中图分类号: TU528

文献标志码: A

Detection and recognition of YOLOv5 pavement cracks based on attention mechanism

Zhou Shuangxi^{1,2}, Yang Dan², Pan Yuan², Ding Jianxin¹, Ding Yang³

(1. School of Civil Engineering and Management, Guangzhou Maritime University, 510765, China, 2. School of Civil Engineering and Architecture, East China Jiaotong University, 330013, China, 3. School of Civil Engineering, Hangzhou City University, 310015, China)

Abstract: Aiming at the problem of poor real-time performance and low precision of traditional pavement crack detection, this paper uses the advantages of deep learning network in target detection, and proposes an improved yolov5 algorithm, which is called yolov5s-attention in this paper, to realize the automatic detection and recognition of pavement cracks. Firstly, the collected crack images are manually labeled with LabelImg annotation software, and then the network model parameters were obtained by improving the YOLOv5 network training. Finally, the model is used to verify and predict the cracks. In addition, F1 and mean Average Precision (mAP) are used to compare the performance of the original YOLOv5s and YOLOv5s-attention models in detecting and identifying pavement cracks. The comparison between YOLOv5s and YOLOv5s-attention showed that the accuracy of YOLOv5s attention increased by 1.0%, F1 increased by 0.9%, and mAP increased by 1.8%. It can be seen that the network has certain practical significance in realizing the automatic recognition of road cracks.

Key words: Road maintenance; Pavement cracks; Object detection; YOLO; Attention Mechanism; Image processing

收稿日期: 2023-05-16

基金项目: 国家自然科学基金资助项目(51968022); 江西省主要学科学术和技术带头人培养计划(20213BCJL22039)

我国幅员辽阔,在道路建设上尤其发达,随之带来的是一系列的道路病害问题,且大部分道路产生的问题主要集中在裂缝上。当前研究路面裂缝的自动化识别问题,有两类方法,第一类是基于数字图像处理;第二类是现在大热的机器视觉方法,即,基于深度学习,建立卷积神经网络(Convolutional Neural Network, CNN)并利用网络实现数据特征的自动化识别,以此使网络模型按一定的规律不断自我调整来实现输入数据输出等于或者无限接近于标签。

本文研究是基于深度学习建立卷积网络来实现路面裂缝自动化检测与识别。基于深度学习的目标检测的网络有 Faster-RCNN^[1]、单步多框目标检测(Single Shot MultiBox Detector, SSD)^[2]、你仅看一次(You Only Look Once, YOLO)^[3]系列等。

Faster-RCNN 目前是应用在路面裂缝中最多的一个两阶段目标检测法,徐等人^[4]对沥青路面因复杂背景难以检测提出了用 ResNet50 替换主干特征提取网络,并在 ResNet50 中引入了注意力机制,实验表明,改进后的模型裂缝检测精准度达到 85.64%。Hao 等人^[5]以机场跑道裂缝为研究背景,采用 Faster-RCNN 网络检测跑道的裂缝,在主干网络里面加入注意力机制,实验表明,改进后的 Faster-RCNN 网络准确率达提高了 6.4%。在 Faster-RCNN 网络中添加了注意力机制使得检测精度有所提高,但此类算法属于两阶段目标检测法,使得训练速度小于单步目标检测算法。

为了提高裂缝自动化识别的速度,单步目标检测算法随之引入裂缝检测中。YOLO 系列是最流行的单级检测方法之一,在一阶段检测模型中实现了快速检测速度和高检测精度之间的良好平衡,并且已经应用于交通运输,如交通标志、车流、路面坑洼、视觉裂缝等检测。黄等人^[8]通过对 YOLOv3 网络的多尺度预测模型进行改进,提高了预测小裂缝的精度。但是由于裂缝分布较为复杂,以至于对该类检测桥梁裂缝的精度还需要进一步的研究。周等人^[9]将 YOLOv3 算法和卷积块注意模块(Convolutional Block Attention Module, CBAM)相结合。此文献所

改进过后的 YOLOv3 不仅降低了网络检测的复杂性和参数,还提高了网络预测精度。文献^[10]中应用 YOLOv4 用于桥梁表面裂缝的检测,并采用一些轻量级网络代替 YOLOv4 中的特征提取网络,可减少参数数量和骨干网络层数,实验结果表明,该方法的精确度、召回率和 F1 值分别比其他先进方法高。Hu 等人^[11]采用 YOLOv5l 为基础模型,在 Neck 模块引入注意力机制 CBAM 和 C3 模块,经实验可知,改进后的 YOLOv5l 模型识别裂缝的 $F^{(0.8)}$ 为 65.9%,召回率为 74.2%。Yu 等人^[12]从两方面改进了 YOLOv5m,在 C3 模块和 SPPF 模块之间添加了 Bot-transformer,经实验表明,改进后的 YOLOv5m 模型识别裂缝的 F1 值提高了 1.3%,mAP 提高了 2.1%,推理速度提高了 1.4 倍。文献^[13]中分别采用了 YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x 这四种 YOLO 模型来检测沥青路面的裂缝,经过实验对比发现,YOLOv5s 的检测速度最快。

从上述可知,在 YOLO 系列算法中引入注意力机制会提高裂缝识别准确率,而 YOLOv5s 裂缝检测速度是 YOLOv5 系列中最快的,因此本文在 YOLOv5s 网络中引入注意力机制得到一个新网络,即,YOLOv5s-attention,且将 YOLOv5s-attention 应用到路面裂缝检测与识别上。文章有以下几部分内容:1 是介绍处理数据集等相关工作;2 是介绍关于 YOLOv5s 算法的基本原理,和引入的注意力机制的原理;3 是讨论改进 YOLOv5s 系列模型的训练结果;最后为关于裂缝检测与识别研究的结论。

1 初始准备工作

1.1 数据集处理

本次研究目的是能利用最新的目标检测算法来自动识别道路裂缝,因此需要采集比较多的裂缝。本文将文献^[14]中的裂缝图片和文献^[15]中的裂缝图片制作成一个新的数据集,新的数据集收集了 3022 个样本,将自建的数据集以 7:2:1 的比例随机分为训练集、验证集和测试集,得训练集 2115 张、测试集 605 张、验证集 302 张。像素密度分别为 300dpi 和 96dpi,数据集中图片的尺寸不一,但由于 YOLO 网络中输入照片尺寸一

般设置为 32 的倍数。因此，在搭建网络时通过程序将裂缝图片尺寸自动修改成 640×640。在输入裂缝图片之前，我们将数据集中所有的照片用标注软件 LabelImg 进行标注，类别为 HCrack(横向裂缝)、30°Crack、45°Crack、LCrack(纵向裂缝)、HoneyCrack(龟裂裂缝)，标注类型保存为 TXT 格式。标注步骤为先建立数据集存放的文件夹，一并在文件夹 datasets 下建立两个子文件夹，分别为 images、labels，且 images 和 labels 文件夹中又包括三个文件夹，分别为 train、test、val。其中 images 用于存放需要使用的裂缝图片，labels 中存放对应的标注文件和索引文件 classes。然后在 Pytorch-GPU 环境下打开 LabelImg 标注软件，对裂缝照片画矩形框进行标注。图 1 是这五种裂缝的标签其中 0 表示横向裂缝，1 表示 30° 裂缝，2 表示 45° 裂缝，3 表示纵向裂缝，4 表示龟裂裂缝。

1.2 设备与软件准备

用于搭建基于 YOLO v5 算法的深度学习道路裂缝自定义识别模型的硬件 CPU 为英特尔 Intel(R) Xeon(R) Platinum 8260，内存为 128G，显卡(GPU)为 NVIDIA GeForce RTX 2080 Ti，内存为 75G。

使用的语言为 Python，环境配置为 Pytorch-GPU。配置 Python 语言环境的步骤为首先安装 Anaconda，用 Anaconda 来管理不同的环境。紧接着安装 Cudnn 和 CUDA，因为本文使用的电脑显卡为 20 系列，那么我们就需要安装 CUDA10.2。然后通过 win+R 启动 cmd，在窗口中激活 Pytorch-GPU 环境，且对应的 Python 版本为 Python3.8。当这些完成之后，我们打开 Anaconda 软件，更改环境，在里面下载编辑器 VSCode。

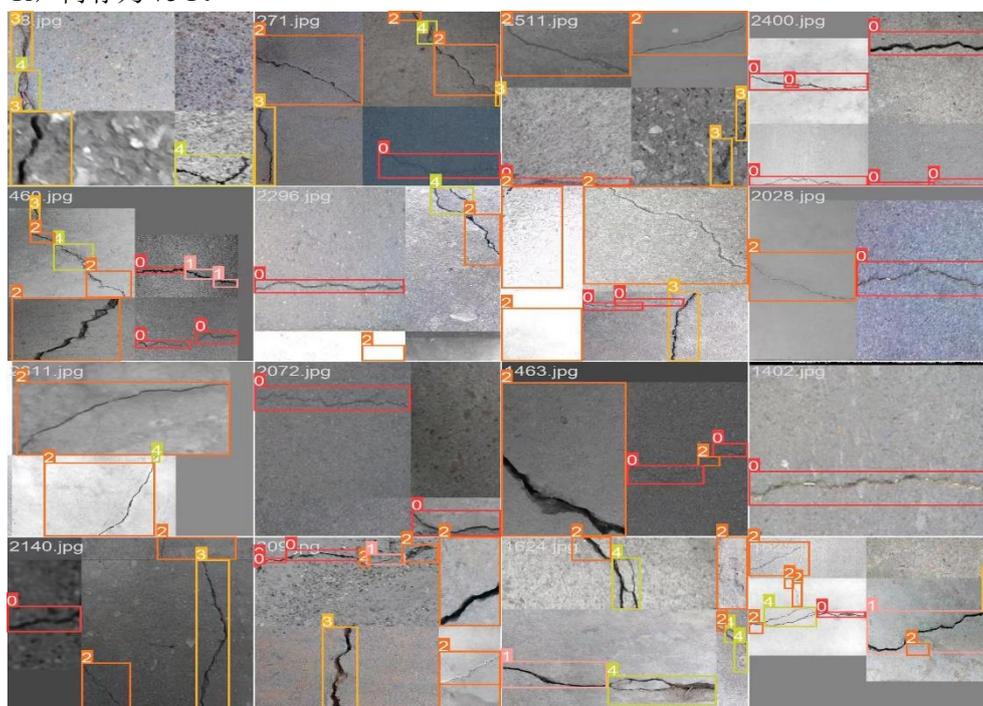


图 1 五种裂缝的标签

Fig. 1 Labels for five cracks

2 YOLOv5 算法介绍

2.1 YOLO 算法发展历程

从 1986 年开始, 进入了一个机器视觉的研究, Rumelhart 和 Hinton 等人首先提出了反向传播(Back Propagation, BP)算法^[16]。到 1998 年 Lecun 又利用 BP 算法训练 Lenet5 网络, 从此卷积神经网络进入人类视线^[17]。2006 年 Hinton 又在 Science Paper 中首次提出了 Deep Learning^[18]的概念。前期关于深度学习的概念已经有雏形, 但是又局限于当时的电脑硬件技术原因。直到 2012 年 Hinton 的学生 Alex Krizhevsky 用 GPU 研究出了一个 Deep Learning 网络模型^[19], 并且在视觉领域竞赛 ILSVRC 2012 中获得了第一名, Alex Krizhevsky 研究的 Deep Learning 网络模型在拥有几百万张的 ImageNet 数据集上的检测效果好于传统检测方法, 从传统检测准确率 70%多提升到了 80%多。到了 2015 年何恺明团队提出了残差网络 ResNet^[20], 且获得 2015 年 ImageNet 竞赛中分类任务、目标检测第一名, 和 COCO 数据集中目标检测和图像分割的第一名。该网络中具有超深的网络结构(突破 1000 层)、Residual 模块以及使用标准化(Batch Normalization, BN)加速训练(此网络丢弃了 dropout)这三个亮点。然而随着深度学习的快速发展, 单纯得对图像进行分类已经不能满足当下的工程

需求, 研究者们开始考虑如何抓取图像中的物体, 然后进行数次训练以实现自动识别功能, 随之发展起来的便是目标检测网络。目标检测模型主要分为两大类:

- (1) 第一类是 Two-stage, 首先通过专门的模块生成区域候选框 (Region Proposal Network, RPN), 寻找前景 (Foreground) 以及调整基于 anchors 的边界框, 然后基于之前生成的区域候选框 RPN 进一步分类以及调整基于 Proposals 的边界框。这一类主要的网络有 Faster-RCNN。
- (2) 第二类是 One-stage, 是直接基于 anchors 进行分类和同时调整边界框。主要的网络有 SSD、YOLO 系列。

这两类网络都有自己各自的优点, One-stage 检测速度快, Two-stage 检测更准确。本文是研究关于道路裂缝的自动识别, 在前人研究了许多神经网络模型的情况下, 对于道路识别精度已经达到比较好的一个精度了, 因此这次为了使得检测速度更加快一点, 选择将 YOLO 算法与道路裂缝结合在一起进行研究。

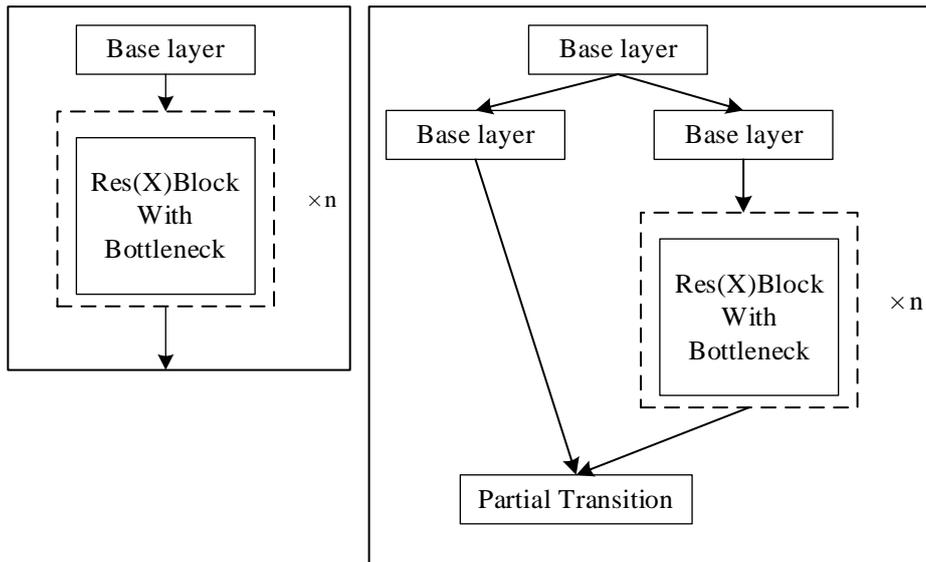


图 2 残差块

Fig. 2 Residual block

2.2 YOLOv5 网络算法

YOLOv5 算法有四个组成部分，分别是：输入(input)、主干特征提取网络(Backbone)、特征金字塔网络(Feature Pyramid Networks, FPN)和检测头(YOLOHead)。

Backbone 是 YOLOv5 的主干特征提取网络，在 YOLOv5 中 backbone 采用的主干特征提取网络是 CSP-Darknet53。CSP-Darknet53 中包含 Focus 模块、Conv_BN_SiLU 模块、C3 模块、以及 SPPF 模块，其中 Conv_BN_SiLU 模块包含卷积层、标准归一化和激活函数，C3 模块中引入残差网络，如图 2。将裂缝图片输入到 YOLOv5 网络中，然后图片在 CSP-Darknet53 中进行特征提取，从而得到输入图片的特征层。

在 Neck 中我们通过加强特征提取网络 FPN 把获取到的三个有效特征层将不同尺度的特征信息进行结合，即，特征融合。并且在 FPN 部分还需要继续特征提取，然后通过路径聚合网络 (Path Aggregation Network, PANet)进行上采样(UpSampling)特征融合和下采样(DownSampling)特征融合。

YOLOHead 是 YOLOv5 的分类器与回归器。详细信息如图 3 所示。

与广泛应用的 YOLOv3 相比，YOLOv5 在以下几个部位进行了改进：

- (1) 数据增强，其中包括 Mosaic 数据增强和 Mosaic 利用了四张图片进行拼接实现数据中增强。它最大的优点是可以丰富

检测物体的背景，且当进行标准归一化 BN 计算的时候可以直接计算出四张图片的数据。

- (2) 多正样本匹配。在之前的 YOLO 系列里面，在训练时每一个真实框对应一个正样本，即在训练时，每一个真实框仅由一个先验框负责预测。YOLOv5 中为了加快模型的训练效率，增加了正样本的数量，在训练时，每一个真实框可以由多个先验框负责预测。

2.3 注意力机制 Attention Mechanism

注意力机制是基于人类视觉研究出来的机器视觉，主要有 Soft attention 和 Hard attention 两类。选择与通道、空间有关的 Soft attention 中的 SENet(Squeeze-and-Excitation Networks)^[21]和 CBAM^[22]。

SENet 是通过 Squeeze 和 Excitation 两个模块实现对特征通道加入注意力机制，先用 Squeeze 对每个特征图进行全局池化，平均成一个实数值，然后通过 Excitation 操作，此时网络输出了 $1 \times 1 \times C$ 大小的特征图，此 Excitation 步骤提升了重要特征，抑制了不重要特征。如图 4。随之研究者又提出与 SENet 注意力机制类似的通过学习的方式自获取每个特征通道的重要程度的轻量的注意力模块 CBAM，且 CBAM 还具有通过类似的学习方式自获取每个特征空间的重要程度。CBAM 的通道注意原理和空间注意原理如图 5、6。

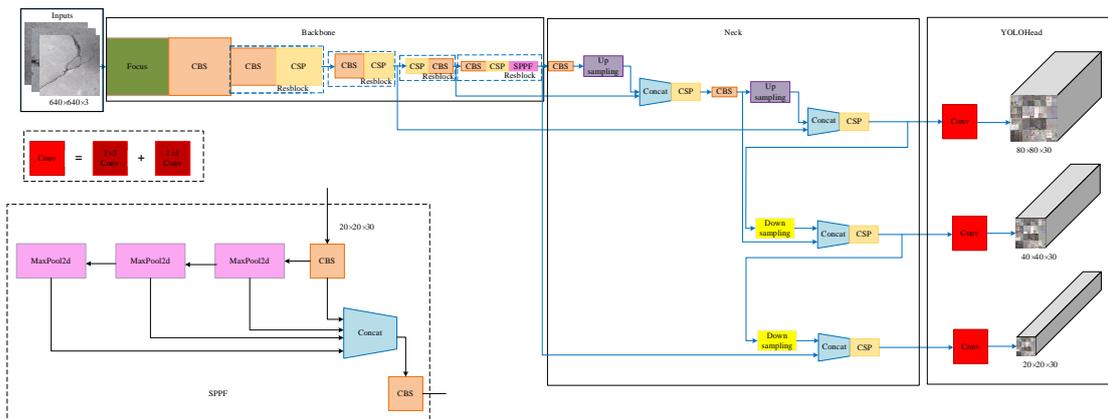


图 3 YOLO v5 网络结构

Fig. 3 YOLOv5 Network Structure

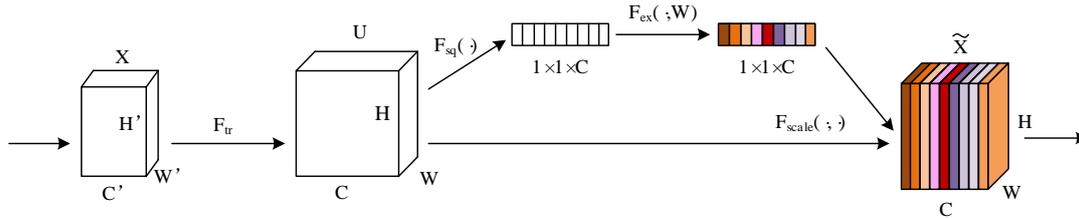


图 4 SENet 注意力模块

Fig. 4 SENet Attention Module

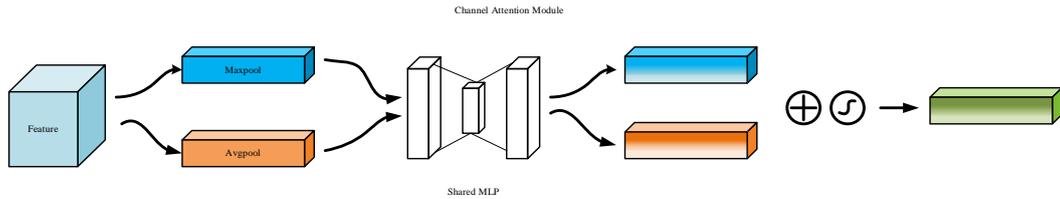


图 5 通道注意力模块

Fig. 5 Channel Attention Module

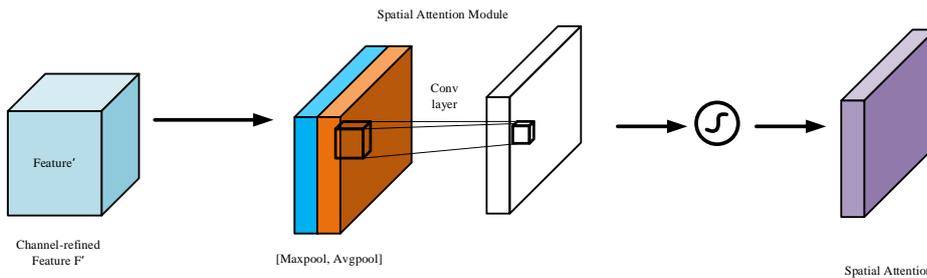


图 6 空间注意力模块

Fig. 6 Spatial Attention Module

2.4 改进主干特征提取网络

CSPLayer-Darknet53 是 YOLOv5s 的主干特征提取网络，其中使用到的 CSPLayer 模块记作 C3。C3 中使用了三个 Conv_BN_SiLU 模块，首先图片输入进主干网络，经过 Focus 层，Focus 层是在裂缝图片上每隔一个像素值取一个特征值，且原本的裂缝照片为 $4 \times 4 \times 3$ 的像素组成，现经过 Focus 层后，获得四个独立的特征层，并将这四个独立的特征层进行堆叠，此时关于高和宽的信息就集中到了通道上，使得通道数扩充了四倍，拼接起来的特征层从原来的三通道变为十二通道，也就是图片信息从原始的高宽集中到通道上，然后输入 C3 模块，在此模块中，首先通过两条通道进行卷积、标准归一化和激活函数处理，其中一条通道

还会经过 bottleneck 模块，然后将提取到的特征堆叠，最后将堆叠后的结果再一次卷积、标准归一化和激活函数处理，结合 C3 模块处理流程，可以在 bottleneck 模块中添加注意力机制。在 YOLOv5s 的主干结构中的 C3 层加入注意力机制 CBAM，不仅可以更好的从通道上提取裂缝的特征，还可以自学习的在空间维度上提取到裂缝的特征，得到一个新的 C3CBAM 层，可以更好的自学习裂缝特征。以及在主干网络第九层引入 SENet 注意力模块，可以将整个主干结构提取到的特征经过 SENet 模块进行自学习，这样可以提高召回率。图 7 是原 C3 模块和改进后的 C3CBAM 模块，图 8 是改进 YOLOv5s-attention 主干结构。

另外，还有其他三种改进 YOLOv5s 模

型，分别是：YOLOv5s-SE(1)、YOLOv5s-SE(4) 和 YOLOv5s-C3CBAM-SimSPPF。YOLOv5s-SE(1)是在 YOLOv5s 原主干结构第九层引入 SENet 注意力模块，作用是利用 SENet 将主干结构提取到的特征在通道上实现自学习，从而提高裂缝识别的准确率；YOLOv5s-SE(4)是在 YOLOv5s 原主干结构的每一层 C3 后面添加一层 SENet 注意力机制，可以实现裂缝特征每通过一层 C3 堆叠后都可以在通道方向上进行自学习，可以明显提高裂缝识别的准确率；YOLOv5s-C3CBAM-SimSPPF 是首先在 C3 内部添加了 CBAM 注意力机制，以实现裂缝识别准确率，以及引进 YOLOv6 中的 SimSPPF 结构^[23]，而 SimSPPF 是将原 SPPF 中的激活函数 SiLU 改为 ReLU，以实现训练时间的缩短。

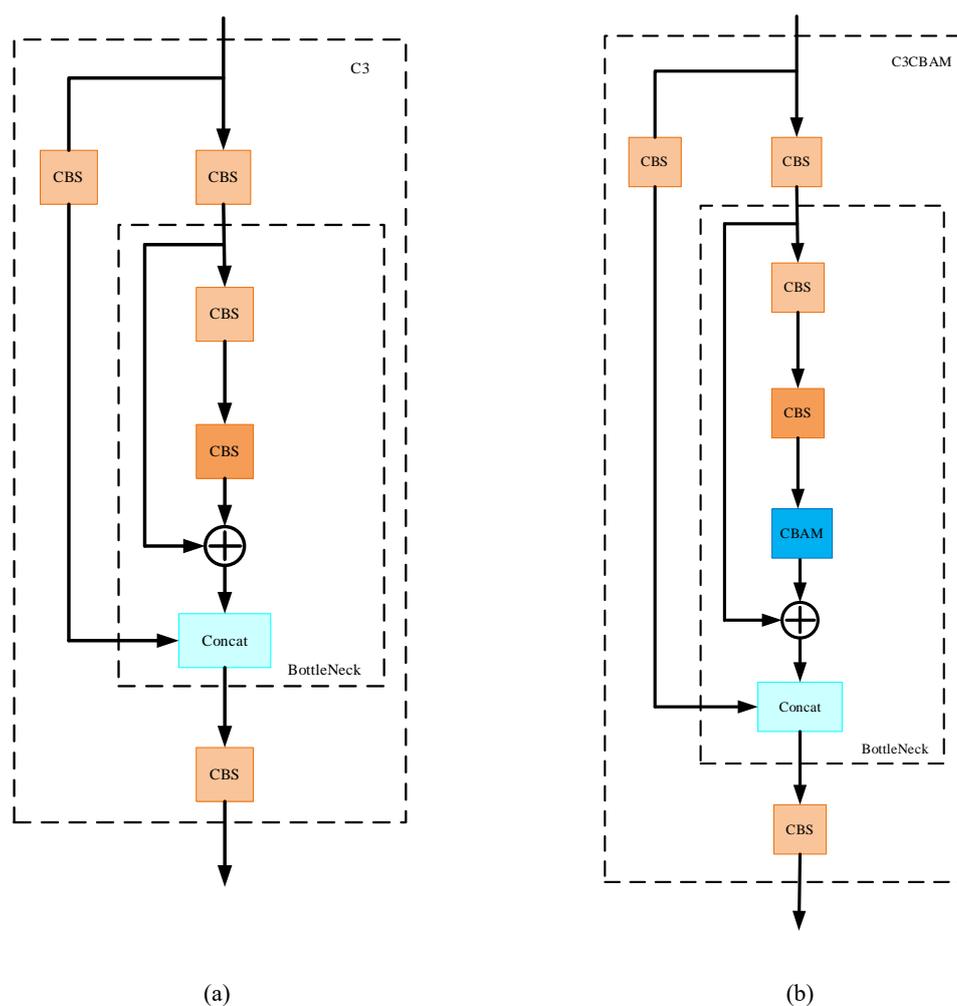


图 7 (a)是原 C3 模块，(b)是 C3CBAM 模块

Fig. 7 (a) is Original C3 block, (b) is C3CBAM block

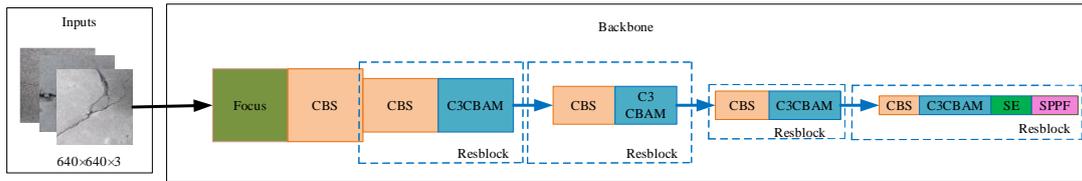


图 8 YOLOv5s-attention 主干结构

Fig. 8 YOLOv5s-attention backbone

2.5 基于改进 YOLOv5s 的路面裂缝检测

网络学习的具体操作为，首先将尺寸为 640×640 的裂缝图片输入到主干结构 C3CBAMDarknet 中，然后在主干结构中提取到的特征传入 Neck 结构中进行特征融合，以及再次提取传入的裂缝特征，最后传入 YOLOHead 中进行预测。图 9 是改进后的 YOLO v5s-attention 网络流程图。

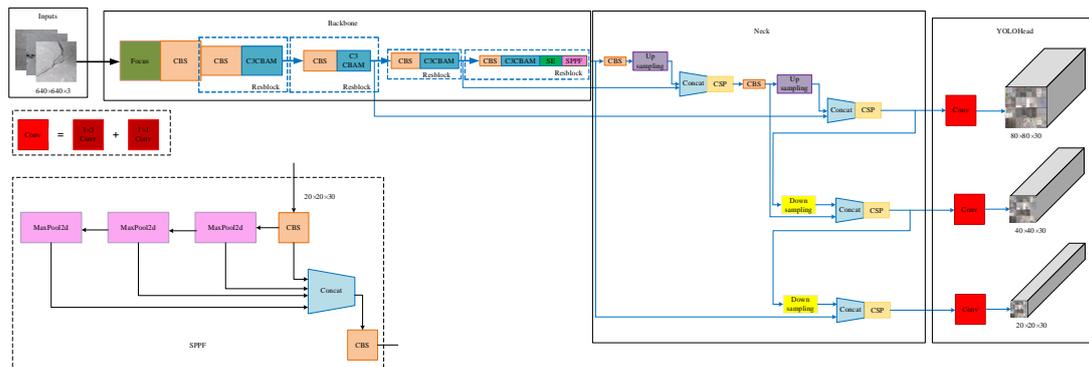


图 9 YOLOv5s-attention 网络流程

Fig. 9 YOLOv5s-attention network process

3 训练结果和评价

3.1 模型训练

首先将处理好的裂缝数据集放置在根目录中，此时裂缝数据集的形式是 TXT 格式。运行 train.py，开始网络训练，训练的数据集就是处理好之后的裂缝图片。训练得到的权重会自动放在 run 文件夹内。对 YOLOv3、原 YOLOv5s、YOLOv5s-attention 以及其他改进的 YOLOv5s 模型都不加 YOLOv5s 的预权重分别训练 300 epoch，然后得到相对应的值。

3.2 评价指标

综合评价指标 F1 是用来平衡准确率和召回率的调和平均值。准确率是检测出一类特征的数量与检测出所有特征数量之间的比率，是用来评价此模型是否能准确检测出

目标，即，查准率；召回率是检测出一类特征的数量和数据集中所有的该类特征数量的比率，衡量的是模型检测系统的查全率。准确率和召回率、综合评价指标 F1 的计算公式如下：

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (1)$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2)$$

其中 True Positive 是网络模型预测是裂缝的正样本，False Positive 是网络模型预测为裂缝的负样本，False Negative 是网络模型预测不是裂缝的正样本，即，误检。

$$F = \frac{2 \times P \times R}{\alpha^2 (P + R)} \quad (3)$$

其中, P 表示 Precision, R 表示 Recall, α 为权重因子。且当 $\alpha=1$ 时, 表示准确率和召回率的权重相同, 即, F 为 F1, 公式如下:

$$F1 = \frac{2 \times P \times R}{P + R} \quad (4)$$

一般情况下, F1 较高时更能说明模型更有效。

平均精度均值 mAP 是检测到所有目标类别平均精度(Average Precision, AP)的均值, 平均精度 AP 是用来衡量某一类检测效果的好坏, 在某一类裂缝检测中即为识别精度, 而平均精度均值 mAP 是衡量多个类别的检测效果, 即所有裂缝识别效果。

$$AP = \sum_{i=0}^{n-1} (r_{i+1} - r_i) P_i (r_i + 1) \quad (5)$$

其中 r_1, r_2, \dots, r_n 是按升序排列的 Precision 插值段第一个插值处对应的 Recall 值, P_i 表示第 i 个检测到的精度。

$$mAP = \frac{\sum AP}{N(\text{classes})} \quad (6)$$

表 1 YOLO 网络 Precision、Recall、F1 比较

Tab. 1 Comparison of YOLO networks Precision, Recall, and F1

Module	Evaluation	Precision(%)	Recall(%)	F1(%)
YOLOv3		43.0	50.7	46.5
YOLOv5s		55.3	59.0	57.1
YOLOv5s-SE(1)		56.9	57.1	57.0
YOLOv5s-SE(4)		57.3	54.7	56.0
YOLOv5s-C3CBAM-SimSPPF		58.0	57.0	57.5
YOLOv5s-attention(Ours)		56.4	59.8	58.1

3.3.2 平均精度均值 mAP

用平均精度均值 mAP(此时 IoU=0.5)衡量检测这五类裂缝的整体检测效果, 以用于评估 YOLOv5 系列模块的性能。下表是各个模型的 mAP 值。

从表 2 可知, YOLOv5s-SE 模型是在 YOLOv5s 的主干特征提取网络第九层添加了一个特征通道自学习的注意力机制 SENet,

3.3 训练结果评估

3.3.1 综合评价指标 F1

用综合评价指标(F1-measure, 此时 $\alpha=1$) 评估 YOLOv3、YOLOv5s、YOLOv5s-attention 以及其他改进的 YOLOv5s 模型是否能准确的识别出这五类裂缝。取 batch_size=1, 计算出 F1 值, 见表 1。

由表 1 可以看出, YOLOv5s 的准确率为 55.3%, 召回率为 59.0%, YOLOv5s-C3CBAM-SimSPPF 的准确率为 58.0%, 召回率为 57%, YOLOv5s-attention 的准确率为 56.4%, 召回率为 59.8%。由此可知, 经过改良的 YOLOv5s-C3CBAM-SimSPPF 模型精确度相对较高, 但是召回率比 YOLOv5s 模型低, 也就是在正确的检测出的结果中召回正确的个数少, 使得 F1 值小于 YOLO 模型, 在 YOLOv5s-C3CBAM-SimSPPF 模型引入了 YOLOv6 中的 SimSPPF 模块, 既不增加计算参数, 还减少了运算速度, 但是经训练表现出召回率不太高, 导致检错率上升, 因此去掉了 SimSPPF 模块, 将 SPPF 层前添加对特征通道自学习的注意力模块 SENet, 从而得到 YOLOv5s-attention 模型, 经验证可知, 此模型精确度比原 YOLOv5s 模型提高了 1.0%, 召回率提高了 0.7%。

mAP 值达到 56.9%, 但是此模型的召回率较低, 导致裂缝的检错率较高, 而 YOLOv5s-attention 模型各方面都表现得比 YOLOv5s 好, mAP 值提高了 1.8%。

表 2 YOLO 网络的 mAP 值

Tab. 1 The mAP value of the YOLO

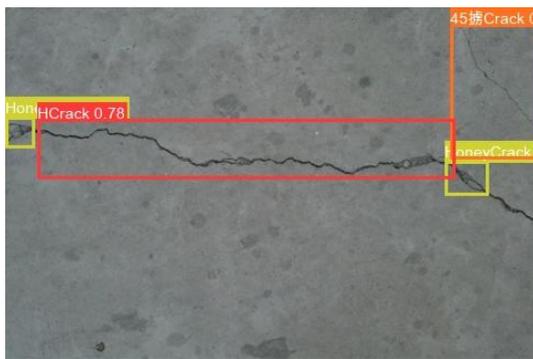
Module	Evaluation	mAP(%)
	YOLOv3	30.7
	YOLOv5s	53.7
	YOLOv5s-SE(1)	56.9
	YOLOv5s-SE(4)	54.1
	YOLOv5s-C3CBAM-SimSPPF	53.9
	YOLOv5s-attention(Ours)	55.7

3.4 图片预测结果分析

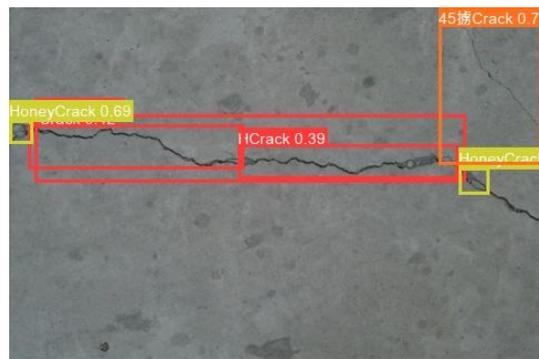
图 10 分别是是 YOLOv5s、YOLOv5s-C3CBAM-SimSPPF、YOLOv5s-attention 三个模型在自学习道路裂缝特征后裂缝预测结果。图(a)是使用 YOLOv5s 网络来识别的裂缝，此算法可以识别出龟裂裂缝、横向裂缝和 45° 裂缝，然而出现漏检裂缝的现象；图(b)是 YOLOv5s-C3CBAM-SimSPPF 网络识别裂缝结果，从图中可知虽然能够识别出三种裂缝，然而对横向裂缝识别的检测框不唯一，也就是不能识别出此条裂缝的类别，且仍存在漏检的裂缝；图(c)是使用 YOLOv5s-attention 网络识别裂缝的结果，可识别出龟裂裂缝、横向裂缝、45° 裂缝和 30° 裂缝，能别出其他两个网络未识别出的 30° 裂缝，且标签框能够准确识别出裂缝的类别，漏检的裂缝能够识别到的原因是由于在 YOLOv5s-attention 网络主干结构中添加了

关于通道的注意力机制 SENet，以及在 C3 模块中添加了 CBAM 模块，两者同时在通道和空间上作用，其召回率相对于 YOLOv5s 和 YOLOv5s-C3CBAM-SimSPPF 网络分别提高了 0.7%和 2.4%，因此能够识别出漏检的 30° 裂缝，另外 YOLOv5s 和 YOLOv5s-attention 网络使用了 K-means 聚类算法，使得标签框能够准确识别出裂缝的类别。

由图 10 中的(d)、(e)、(f)可以看出，原始 YOLOv5s 网络和改进后的 YOLOv5s-C3CBAM-SimSPPF 网络存在误检的现象，且不能检测到小裂缝目标，而 YOLOv5s-attention 网络即没有把划痕误以为是裂缝，也识别到图片中右下角的裂缝小目标。



(a)



(b)

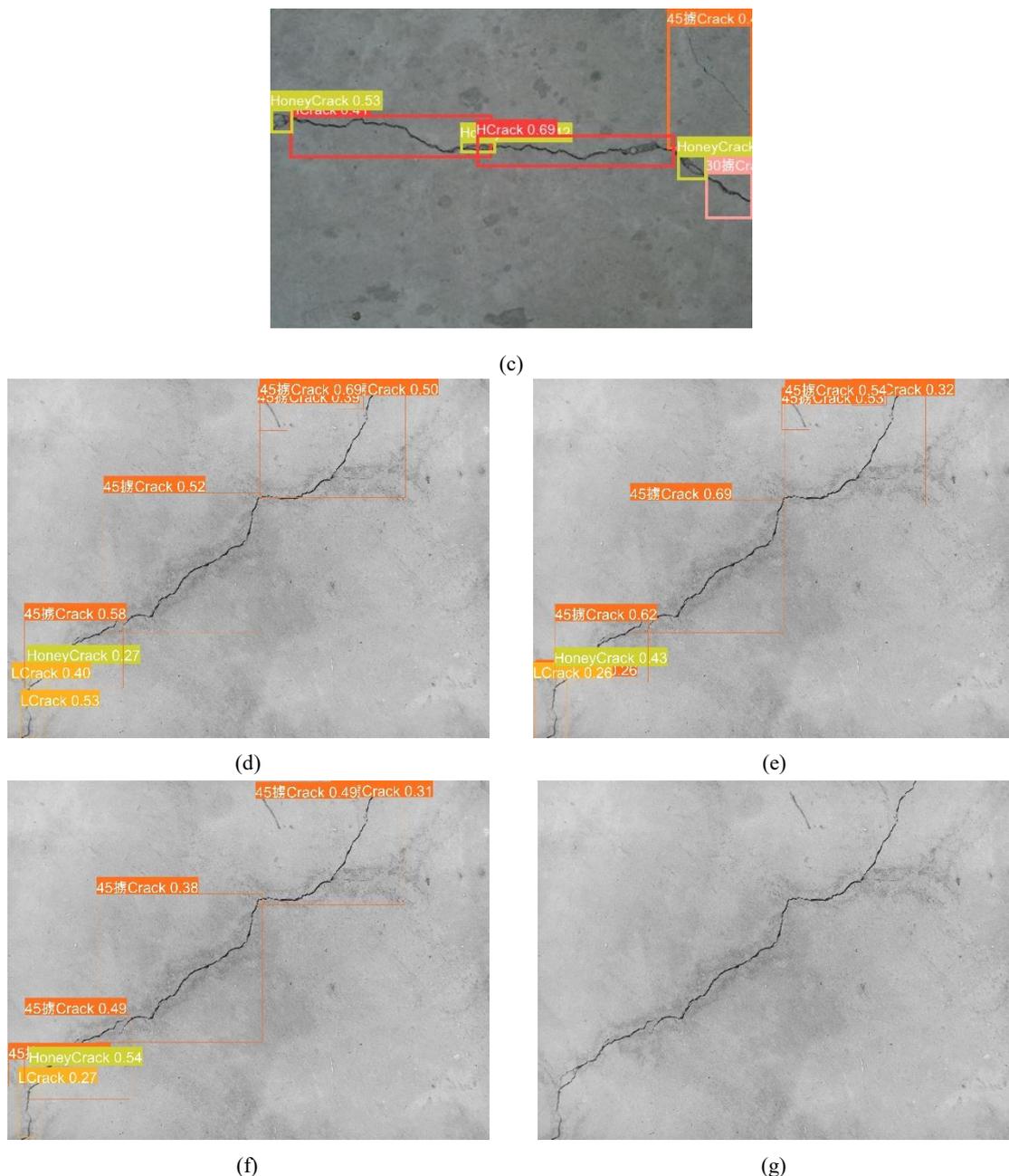


图 10 (a)、(d)图为 YOLOv5s 网络预测结果, (b)、(e)图为 YOLOv5s-C3CBAM-SimSPPF 网络预测结果, (c)、(f)图为 YOLOv5s-attention 网络预测结果, (g)为裂缝原图

Fig. 10 (a) The figure shows the YOLOv5s network prediction results, (b) the YOLOv5s-C3CBAM-SimSPPF network prediction results, (c) the YOLOv5s-attention network prediction results, (g) Original image of cracks

4 结论

针对路面裂缝问题提出了利用神经网络算法来自动化识别裂缝。讨论了 YOLOv3、YOLOv5s、YOLOv5s-attention 和一系列改良的 YOLOv5s 模型在裂缝识别上的速度和精度, 并采用综合评价指标 F1 和 mAP 来评价网络算法的优越性。通过网络训练对比,

YOLOv5s-attention 网络从目标检测网络中脱颖而出。YOLO v5s-attention 网络将主干结构 backbone 中的 C3 层添加注意力模块 CBAM, 以及在 SPPF 层和 C3 层之间添加注意力模块 SENet。由训练可得结论:

1) 与原 YOLOv5s 训练时长相比, 添加注意力机制后对训练时长影响不大, 平均三个小时训练 300epoch。

2) 新的 YOLOv5s-attention 网络对识别 30° 裂缝的准确率提高, 且可以预测出原 YOLOv5s 网络未识别到的 30° 裂缝。

3) 新的 YOLOv5s-attention 网络在识别裂缝能力上与原 YOLOv5s 网络相比可知, 其综合评价指标 F1 值上升 0.9%, 平均精度均值 mAP 提高了 1.8%。

接下来的工作就是进一步改良 YOLOv5 网络, 使得改良后的 YOLOv5 模型更加占用的内存更小, 即更加轻量化, 检测速度更快, 并且可实现安装在小型移动设备上, 进行实时检测路面裂缝, 随有随补, 避免道路裂缝开裂持续性。

参考文献

- [1] REN S Q, HE K M, Girshick R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137-1149.
- [2] WEI L, Dragomir A, Dumitru E, et al. SSD: Single Shot MultiBox Detector[J]. Springer, Cham, 2016.
- [3] Redmon J, Divvala S, Girshick R, et al. You Only Look Once: Unified, Real-Time Object Detection[C]//Computer Vision & Pattern Recognition. IEEE, 2016: 779-788.
- [4] 徐康, 马荣贵. 基于改进 Faster-RCNN 的沥青路面裂缝检测[J]. *计算机系统应用*, 2022, 31(07): 341-348.
XU K, MA R G. Asphalt pavement crack detection based on improved Faster-RCNN[J]. *Computer Systems & Applications*, 2022, 31(07): 341-348.
- [5] HAO S, SHAO L & WANG S. A Faster RCNN Airport Pavement Crack Detection Method Based on Attention Mechanism[J]. *Academic Journal of Science and Technology*, 4(2), 129-132.
- [6] HAN Z, CHEN H, LIU Y, et al. Vision-based crack detection of asphalt pavement using deep convolutional neural network[J]. *Iranian Journal of Science and Technology, Transactions of Civil Engineering*, 2021, 45: 2047-2055.
- [7] 李鹏程, 孙立双, 谢志伟, 等. 基于改进 MobileNet-SSD 的路面裂缝图像检测算法[J]. *激光杂志*, 2022(007): 43.
LI P C, SU L S, XIE Z W, et al. Pavement crack image detection algorithm based on improved mobilenet SSD[J]. *Laser Journal*, 2022(007): 43.
- [8] 黄开启, 刘小荣, 黄茂云. 基于改进 YOLOv3 的小目标检测方法研究 [J]. *传感器与微系统*, 2022, 41(4): 4.
HUANG K Q, LIU X R, HUANG M Y. Research on Small Target Detection Method Based on Improved YOLOv3 [J]. *Sensors and Microsystems*, 2022, 41 (4): 4.
- [9] 周清松, 董绍江, 罗家元, 秦悦, 夏宗佑, 杨建喜. 改进 YOLOv3 的桥梁表面病害检测识别 [J]. *重庆大学学报*, 2022, 45(6): 121-130.
ZHOU Q S, DONG S J, LUO J Y, et al. Improving YOLOv3 for detecting and identifying bridge surface defects [J]. *Journal of Chongqing University*, 2022, 45 (6): 121-130.
- [10] ZHANG J, QIAN S, TAN C. Automated bridge crack detection method based on lightweight vision models [J]. *Complex & Intelligent Systems*, 2022: 1-14.
- [11] HU N, YANG J, JIN X, et al. Few-shot crack detection based on image processing and improved YOLOv5[J]. *Journal of Civil Structural Health Monitoring*, 2023, 13(1): 165-180.
- [12] YU G, ZHOU X. An Improved YOLOv5 Crack Detection Method Combined with a Bottleneck Transformer[J]. *Mathematics*, 2023, 11(10): 2377.
- [13] HU G X, HU B L, YANG Z, et al. Pavement crack detection method based on deep learning models[J]. *Wireless Communications and Mobile Computing*, 2021, 2021: 1-13.
- [14] LIU Y H, YAO J, LU X H, et al. DeepCrack: A deep hierarchical feature learning architecture for crack segmentation[J]. *Neurocomputing*, 2019:139-153.
- [15] LI L F, MA W F, LI L, et al. Research on Detection Algorithm for Bridge Cracks Based on

- Deep Learning[J]. ACTA Automatica Sinica, 2019, 45(9): 1727-1742.
- [16] David E Rumelhart, Geoffrey E Hinton, Ronald J Williams. Learning representations by back-propagating errors[J]. Nature, 1986: 533-536.
- [17] Y Lecun, L Bottou, Y Bengio, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998: 2278-2324.
- [18] LE C, Y, Bengio Y & Hinton, et al. Deep learning[J]. Nature 521, 2015:436-444.
- [19] Krizhevsky A, Sutskever I, Hinton G. ImageNet Classification with Deep Convolutional Neural Networks[J]. Advances in neural information processing systems, 2012, 25(2).
- [20] Kaiming He, X Zhang, Shaoqing Ren, et al. Deep Residual Learning for Image Recognition[J]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015: 770-778.
- [21] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 7132-7141.
- [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] Li C, Li L, Jiang H, et al. YOLOv6: A single-stage object detection framework for industrial applications[J]. arXiv preprint arXiv:2209.02976, 2022.



第一作者：周双喜（1973-），男，教授，博士，硕士生/博士生导师，南昌市 521 学术技术带头人，研究方向为混凝土材料与检测技术研究，1998 年本科毕业于济南大学，2003 年硕士毕业于桂林理工大学，2006 年博士毕业于中国建筑材料科学研究院。E-mail: 18820358448.green.55@163.com。



通信作者：杨丹（1997-），女，学士，无，研究方向为混凝土材料与检测技术研究，2020 年本科毕业于青岛农业大学。E-mail: gryangdan@163.com。