

基于深度特征金字塔的路面病害检测

张伟¹, 刘宁钟¹, 寇金桥²

(1. 南京航空航天大学 计算机科学与技术学院, 江苏 南京 211106;

2. 北京计算机技术及应用研究所 方舟重点实验室, 北京 100854)

摘要:道路为人们的生活和工作提供了方便,路面作为道路最重要的组成部分,直接影响了道路的使用,但由于车辆行驶和风吹日晒,路面病害层出不穷。利用目标检测技术对路面病害进行快速检测,可以及时筛选出病害,降低日常人工检查的负担,提高养护效率。然而,路面病害特征比较细微,随着神经网络深度的不断增加和下采样,细节信息损失比较多。通过将通道注意力集成到特征金字塔网络,可以从通道和空间两个维度上提高网络对路面病害的表征能力,同时提出了一种新的路面病害特征提取器,使得网络更关注低层次特征。实验部分,将改进后的特征金字塔分别应用在 Road Damage Dataset 2018 数据集和自制的沥青路面病害数据集上,并与其他经典的目标检测模型进行了比较,实验结果证明了基于改进后的特征金字塔的模型在路面病害检测上的有效性。

关键词:路面病害;目标检测;特征金字塔;通道注意力;特征提取器

中图分类号: TP31

文献标识码: A

文章编号: 1673-629X(2022)12-0173-06

doi: 10.3969/j.issn.1673-629X.2022.12.026

Pavement Disease Detection Based on Depth Feature Pyramids

ZHANG Wei¹, LIU Ning-zhong¹, KOU Jin-qiao²

(1. School of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China;

2. Ark Key Laboratory, Beijing Institute of Computer Technology and Application, Beijing 100854, China)

Abstract: The road provides convenience for people's life and work, and the pavement, as the most important component of the road, directly affects the use of the road. Due to vehicle movement and wind and sun, pavement diseases are endless. Rapid detection of pavement diseases using object detection technology allows timely screening of diseases, reduces the burden of daily manual inspection, and improves maintenance efficiency. However, the pavement disease features are relatively subtle, and more detailed information is lost as the depth of the neural network continues to increase and downsample. By integrating channel attention into the feature pyramid network, the network's ability to characterize pavement distress can be improved in both channel and spatial dimensions, and a new pavement disease feature extractor is proposed to make the network more focused on low-level features. In the experimental part, the improved feature pyramid is applied to the Road Damage Dataset 2018 dataset and the homemade asphalt pavement disease dataset, respectively, and compared with other classical object detection models. The experimental results show the effectiveness of the model based on the improved feature pyramid for pavement disease detection.

Key words: pavement disease; object detection; feature pyramid; channel attention; feature extractor

0 引言

道路基础设施是中国一项重要的公共资产,也是国家整个交通系统的基本组成部分。截至2021年初,中国高速公路运行长度约十六万公里,稳居全球首位^[1],而到了2035年,中国的农村公路更要全面做到乡镇通的三级公路,总计里程数能够超过五百万公里,

这就要求道路养护当局提供必要的支持和强有力的工具,以准确实时地应对道路状况问题^[2]。传统的道路检查以日常检查和年度检查相结合,在日常检测中,以人工巡查为主,然而现场目视检查是劳动密集型的。对于年度检查,由权威机构牵头,采用专用设备和人员进行操作检测,但是这种检测过程成本极高^[3-4]。因

收稿日期: 2021-12-27

修回日期: 2022-04-28

基金项目: 中央高校基本科研业务费专项资金(3082020NZ2020017)

作者简介: 张伟(1997-),男,硕士,CCF会员(J7554G),通讯作者,研究方向为计算机视觉;刘宁钟,博士,教授,研究方向为计算机视觉;寇金桥,硕士,研究员,研究方向为嵌入式系统架构设计、异构加速及软件开发。

此基于图像的道路表面病害识别深受研究人员的青睐^[5],早期研究侧重于传统的图像处理技术,这一类方法主要是利用路面病害像素和路面背景像素的差异^[6-8]。一些研究人员也将机器学习方法与图像处理技术相结合,可以进一步提高表面病害检测的鲁棒性和通用性^[9]。

近年来,卷积神经网络已经证明它们可以在复杂的图像上表现稳定,因为它们可以从输入图像中挖掘出更具有代表性的特征^[10],所以越来越多的研究人员尝试使用深度神经网络来检测道路表面的病害。分类网络可以在只包含一种预定义病害类型的图像中获得相当高的准确性^[11]。分割神经网络在病害的量化上具有明显优势,因为其检测粒度是像素级别的^[12-13]。使用基于目标检测的方法对道路表面病害进行检测并分类是更好的选择,因为目标检测数据集的标注与制作更加简单高效,且这些方法可以提供实例级别的病害检测^[14]。

在目标检测网络的前馈推理过程中,考虑到模型的速度和大小,必须对原始输入图像和后续特征图进行降采样。然而,采集到路面图像中的路面病害本身目标信号就较弱,一些小裂缝可能只有少数像素,因此,随着网络深度的增加,不可避免地会造成信息的持续损失。在以往的目标检测研究中,研究者倾向于使用更深层次的特征图,因为这些特征在经过大量神经网络层的提取后,具有丰富的语义信息。然而这样的特征选择却忽略了一定的空间信息,尤其对于道路表面病害来说,这种信息忽略可能是致命的,因为原始图像中一条明显的裂缝,随着网络深度的加深,可能不再为深层网络所见。特征金字塔网络 (Feature Pyramid Network, FPN) 让目标检测模型也关注到了浅层特征,并将特征图的各个层次进行了融合,空间信息和语义信息进一步得到了提升^[15]。因此将 FPN 引入到道路表面病害的目标检测网络中,可以为整个网络提供更多的空间信息。此外,在使用深度神经网络处理图像时,原始图像以及后续特征图是三个维度的,空间信息仅仅利用了二维的信息,因此通道信息也是不能被忽略的。与此同时,如何更好地利用 FPN 所生成的多层次特征图并提取出更有用的信息也是该文的研究重点。

1 相关工作

1.1 FPN

2017 年, Lin 等人提出了特征金字塔网络 FPN, 金字塔的构建包括自下而上的路径、自上而下的路径和横向连接。图 1 显示了构建 FPN 的过程。对于最高层次的特征图,其包含的细节信息较为粗糙,将该特征

图的空间分辨率向上采样 2 倍。然后,通过对应元素直接相加的方式,将上采样后的特征图与相应的自下向上的特征图通过横向连接融合。重复此过程,直到最大分辨率的最底层特征图。最后,在每个融合后的特征图上加上一个 3×3 的卷积,生成最终的输出特征图,这可以减小融合后的混叠效果。最后一层输出特征图 P6 是通过 P5 上采样生成的。FPN 是一种可以用于多尺度目标检测的实用且准确的解决方案,在两阶段目标检测模型 Faster R-CNN^[16] 上集成 FPN 并取得显著的效果之后,FPN 成为许多新目标检测器和模型的基础。

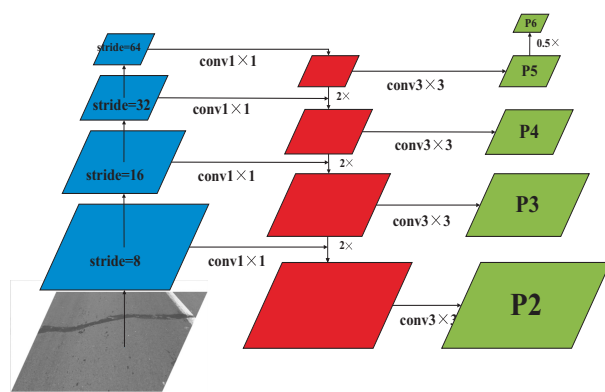


图 1 FPN 的网络结构

对于两阶段目标检测模型来说, FPN 不仅是生成候选目标的特征来源,也是最终输出预测目标的特征来源。目前的两阶段检测模型大多使用 ROI 特征提取器,该特征提取器根据候选区域生成网络 (Region Proposal Network, RPN) 生成的候选边界框来选择主干网生成的特征区域。引入 FPN 以后,特征提取不再直接操作于主干网特征,而是操作于由 FPN 生成的特征金字塔上。对于 ROI 特征提取器来说,会将每个 ROI 映射到 FPN 的某一个单层次特征图,然后进行池化等操作。公式 1 就是最基本的映射规则,将宽度为 w 、高度为 h 的 ROI 分配到特征金字塔的第 k 层。以 Faster R-CNN 为例,在未使用 FPN 之前,仅在 C4 阶段的单尺度特征图上运行目标检测与分类,因此 k_0 一般设置为 4。而公式 1 同时也说明如果某个 ROI 的尺度较小 (例如, 224 的一半), 那么该 ROI 会被映射到一个更低层次的特征图 (例如, $k = 3$), 而该特征图刚好也包含更多的细节信息,这也表明了公式 1 设计得比较合理,因此该映射规则一直被使用。

$$k = \lfloor k_0 + \log_2 \left(\frac{\sqrt{wh}}{224} \right) \rfloor \quad (1)$$

1.2 通道注意力

最经典的通道注意力架构单元是挤压和激发块 (Squeeze and Excitation, SE)^[17]。图 2 显示了通道注意力机制的挤压和激励过程。第一步挤压是通过全局

平均池化将形状为 $W \times H \times C$ 全局特征图压缩成 $1 \times 1 \times C$ 形状的通道描述符,以此来收集全局空间信息。第二步是激励,将第一步中聚集的信息,通过 FC-ReLU-FC 变换来重新学习通道之间的权重分布,并引

入一个超参数 r 来限制复杂性,然后用一个 sigmoid 激活函数作为一个简单的门控机制输出一个注意力向量。注意力向量最终通过相乘的方式来重新调整原始特征图的输入通道。

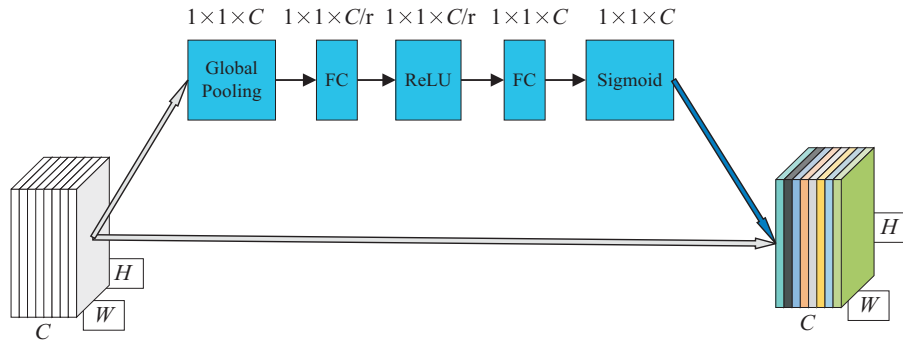


图2 SE通道注意力机制结构单元

1.3 路面病害目标检测

在 Maeda 等人的研究中,开发了分别以 MobileNet 和 Inceptionv2 为主干网的 SSD 检测模型,用来检测 8 类道路表面病害,同时公布了公共道路损伤数据集 2018 (Road Damage Dataset 2018, RDD2018)^[18]。Wang 等人在该 RDD2018 上使用 Faster R-CNN 来检测和分类受损的道路。Kluger 为了解决道路损坏检测和分类问题,应用 Faster R-CNN、RetinaNet、随机森林与神经网络以及集成学习的方法。

2 模型与方法

FPN 是一种等价于从不同图像分辨率提取特征的体系结构,它已经被大量实验证明具有维持空间信息的有效能力。该文进一步考虑通道特征对于路面病害表征的有效性。长久以来,通道特征注意力的发展一直集中在语义分割和图像分类任务上,在目标检测领域应用了更多的空间注意力机制,这是因为对于目标检测的直观感受上,空间信息更重要。而对于路面病害检测来说,经过主干网的一系列卷积处理,然后通过 FPN 生成特征金字塔,每两个层次的特征图之间的分辨率相差 2 倍,即使分辨率最高的特征图,仍然只有输入图像的四分之一大小。然而,图像中的道路病害,如裂缝,本身可能只有几个像素,因此,随着网络深度的增加,一条明显的裂缝可能不再为深层网络所见。笔者认为此时增加空间注意力并不是最佳选择,因为卷积本身就是一种广义的空间注意力模块,因此将重点放到特征图的通道维度,如图 3 所示,对 FPN 的每一层输出图都使用一个 SE 进行通道矫正。在以往的模型中集成 SE 时,都是将 SE 加到主干网的某一些堆叠块后面,比如残差块,该文开创性地将通道注意力机制应用到颈部网络,集成方式只是简单的串联,使模型具有松耦合的特性,简单高效,并且几乎不增加模型的

复杂度。

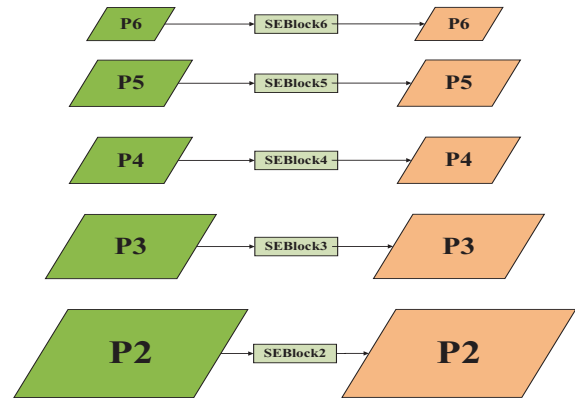


图3 FPN与通道注意力的集成

集成了通道注意力的特征金字塔,分别从通道维度和空间维度增强了模型对于路面病害的表征能力,并生成了多层次特征图,然后由特征提取器根据 ROI 选择特征层次进行特征提取。在传统的 FPN 中只考虑了单层次特征图提取的方法,这存在一些信息相关的问题。公式 2 可以表示为 RPN 中的回归子网对锚框的修正, x 是在特征金字塔上输入的某一层特征图, w_* 是要学习的回归权重参数($*$ 表示 x, y, w 和 h , 对应锚框的平移变换和尺度缩放), $d_*(x)$ 是网络的输出,用来修正候选锚框。可以看出,这里可能存在特征没有对齐的问题,如果特征图 x 来源于第 i 层,候选锚框 P_i 位于来自特征金字塔的第 i 层,然后候选锚框 P_i 经过 RPN 的第一次修正后变为 P_2 ,在头部网络进行第二次修正时,公式 1 将 P_2 映射到 FPN 的第 j 层,如果 $i \neq j$,那么该 ROI 经过两次修正后很明显存在依赖的特征基础不一致。

$$d_*(x) = w_* \cdot x \quad (2)$$

在过往的研究中都指出,在设计路面病害的检测模型时,如何利用好低层次的特征是非常重要的,高层次的特征融合可能会造成冗余的干扰信息^[19]。受此

启发,提出了一种自下而上的 ROI 路面病害特征提取器,将其命名为 RDROIE (Road Disease ROI Extractor)。图 4 显示了这一提取框架。 k 是通过公式 1 计算而来,认为由公式 1 计算出来的第 k 层次是某个尺度病害的理想特征层次,具有较好的语义特征,然而由于道路病害的细微性,在卷积下采样的过程中,特征信息损失严重,因此选择 k 层及 k 层以下的 ROI 特征进行融合。首先是 ROI 池化模块,它是一个对非规则 ROI 执行最大池化以获得固定大小表示的模块,ROIAlign 是现阶段最流行的池化操作,它通过将原始

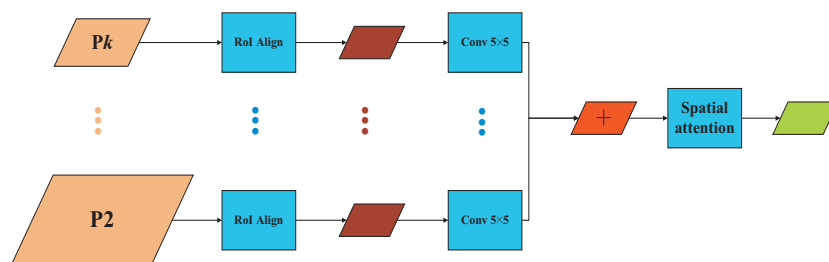


图 4 路面病害 ROI 提取框架

3 实验

3.1 实验细节

对于路面病害检测来说,准确检测到病害的重要

特征图上的 ROI 划分为相等的网格,并在每个网格内使用双线性插值来缩小矩形特征图区域。经过 ROIAlign 池化模块生成 ROI 特征都具有相同的形状,随后是一个预处理模块,其目的是对池化后的区域进行初步细化。该文使用了一个 5×5 卷积层,然后使用一个求和操作,聚合来自每个层次分支的单个 ROI 特征。在最终返回聚合特征之前应用了一个额外的细化步骤,这一步应该允许网络学习全局特征,加强最终 ROI 特征的信息能力。该文选择应用空间注意力层,这一层的主要目标是消除无用信息。

性是高于检测速度的,而两阶段模型在准确率上大多数时候是优于一阶段模型的,因此采用两阶段的检测模型作为基线模型,选取了带有 FPN 的 Faster R-CNN,图 5 展示了所使用的路面检测框架。

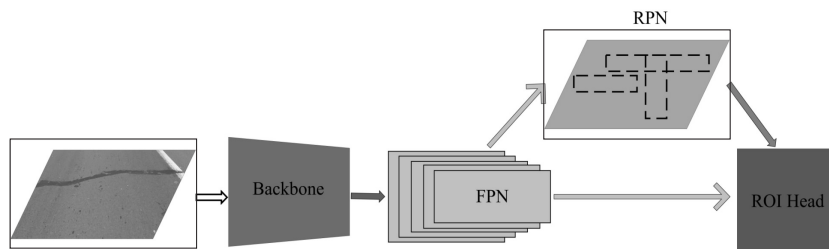


图 5 路面病害检测模型

该文将改进后的特征金字塔网络在两个数据集进行了验证,一个是 RDD2018,图像来自日本的 7 个地方 (Ichihara, Chiba, Nagakute, Sumida, Muroran, Adachi 和 Numazu),共 9 053 张图像样本,有 7 240 张和 1 813 张图像被分别用作训练样本和测试样本。该数据集包含 8 个不同的道路病害类别,分别是纵向裂缝 (D00),纵向接缝 (D01),横向裂缝 (D10),横向接缝 (D11),龟裂 (D20),车辙、露骨、坑槽和分裂归为一类的 D40,人行道模糊 (D43) 和白线模糊 (D44)。一个是笔者自己制作的沥青路面病害数据集,包含 2 297 张图像,随机选择了其中 2 072 张用于训练,225 张用于验证,共六种病害,分别是纵向裂缝 (Type1),横向裂缝 (Type2)、灌封修补 (Type3)、块状修补 (Type4)、坑槽 (Type5) 和块状裂缝 (Type6)。图 6 分别给出了两个数据集中的图像示例,(a) 是文中的沥青路面病害数据集的示例,(b) 是 RDD2018 中的示例。在训练期间,对于自己制作的数据集,用于训练和测试的图像原

始大小为 $4\,000 \times 3\,000$ 的高分辨率图像,如果直接使用原始分辨率的图像作为训练输入,需要大量的 GPU 内存以及训练时间,造成模型训练和优化困难。针对这种情况,将图像大小调整为 $1\,333 \times 800$,这是最常用的目标检测图像输入分辨率,既可以保持病害在图像中的完整性,也可以减少在训练阶段对硬件的需求。尽管如此,由于使用的是一块 GTX 1080 Ti 11 GB GPU,批量大小不能设置得太大,但是对于批量归一化来说,小批量会使批量统计的估计不准确,所以在训练过程中固定 BN 层的参数,不更新它们。在将每个图像馈送到网络之前,对其进行了随机水平翻转,随机概率为 0.5,然后将输入图像的 RGB 通道上的均值和标准差值进行归一化。在后续的比较中,与之比较的所有模型都不是来自原始论文的实现,但它们是在相同的环境上进行训练并测试,除了提出的改进的 FPN 模块外,其余都使用相同的配置,因此该文使用 MMDetection 作为基础工具箱来开发文中方法。



图 6 路面病害图像示例

3.2 实验结果

采用基于 IOU 的方法来匹配预测边界框与真实标注 (Ground Truth, GT) 边界框,从而判断定位是否准确,如果 $\text{IOU} \geq 0.5$,则认为该检测边界框是 TP。如果 $\text{IOU} < 0.5$,则认为该检测边界框是 FP。由于目标检测中不涉及真阴性 (TN) 这一概念,因此评估指标主要是基于 TP、FP 和 FN 的精确率 P 和召回率 R ,分别定义为:

$$P = \frac{\text{TP}}{\text{TP} + \text{FP}} = \frac{\text{TP}}{\text{所有预测边界框}} \quad (3)$$

$$R = \frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{\text{TP}}{\text{所有 GT 边界框}} \quad (4)$$

在此基础上使用 F1 分数来评价检测结果, F1 分数定义如下:

$$\text{F1} = 2 \cdot \frac{P \times R}{P + R} \quad (5)$$

将 Faster R-CNN ResNet-50 with FPN 作为基线结果,然后在此基础上引入提出的集成通道注意力的特征金字塔和改进的 ROI 特征提取器。表 1 显示了在逐渐引入提出的方法后,模型在沥青路面病害数据集和 RDD2018 测试集上的 F1 分数变化。

表 1 每个启发式方法在测试集上的 F1 分数

组件	RDD2018	沥青路面数据集
基线	0.614	0.397
+SE	0.619	0.401
+RDROIE	0.630	0.415

如表 1 所示,在基线模型上的 FPN 上集成了 SE, RDD2018 上的 F1 分数从 0.614 上涨到了 0.619,在沥青路面数据集上由 0.397 上涨到 0.401,这证明了引入通道注意力机制的有效性,也说明特征通道维度非常值得关注。然后使用所提道路病害特征提取器 RDROIE,尽管模型复杂度略有增加,但从表 1 中可以看出, RDD 上 F1 分数上涨了 1.1%,精度上升较为明显,达到 0.630,相似地,在沥青路面病害数据集上上涨了 1.4%,达到了 0.415,这说明对于路面病害的目标检测任务来说,融合更多低层次的特征更有效果。表 2 展示了不同模型的 F1 分数对比。

表 2 不同模型的 F1 分数对比

模型	RDD2018	沥青路面数据集
SSD	0.526	0.148
YOLOv4	0.483	0.279
RetinaNet	0.456	0.324
FCOS	0.533	0.385
Faster R-CNN	0.610	0.363
文中模型	0.630	0.415

在验证了提出的模型改进方法的有效性后,与其他的流行的目标检测模型在 ARSDD 测试集进行了比较。从表 2 中可以看出,两阶段模型 Faster R-CNN 和所提模型在结果上明显好于一阶段模型 SSD、YOLOv4、RetinaNet 和 FCOS,这说明了两阶段目标检测在现阶段的路面病害检测中的应用应该是更有效的。FCOS 是近期比较热门的一个目标检测模型,该模型实现了无候选边界框的创新思路,在诸如 COCO 等公开数据集上的表现好于许多基于锚框的目标检测算法。从表 2 可以看出,该模型在两个数据集的 F1 分数都明显优于其余的模型。

3.3 结果分析

图 7 显示了文中模型的在两个数据集上正确的检测结果,图 7(a)检测到了一些沥青路面数据集中的横向裂缝。同理,图 7(b)检测到了一些龟裂 (D20)。裂缝作为路面上最常见的病害,对其进行精准检测是十分重要的,无论是图 7(a)近处的裂缝,还是图 7(b)中远端的裂缝,基于改进后的特征金字塔网络模型都能够有效地检测到。

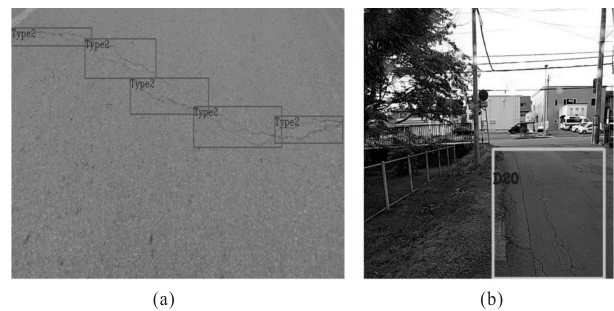


图 7 模型的检测结果可视化

4 结束语

道路路面的损坏会影响道路正常运行,如果放任不管,会使得整条道路持续恶化,该文使用基于目标检测的方法来检测道路表面病害。由于路面病害目标具有特征信号微弱的特点,首先介绍了 FPN 以及特征提取器,然后阐述了目标检测中通道注意力机制,在此基础上,介绍了针对路面病害检测所提出的两个方法,一个是将通道注意力机制集成到多层次特征图上,增强网络的语义表征能力。第二个方法是对特征提取器进

行了重新设计,让路面病害的 ROI 不再仅仅只依赖于某一层次的特征图,而是融合了更多低层次的特征图信息。接着采用基于 Faster R-CNN 在 RDD2018 和沥青路面病害上分别进行了实验,结果表明带来了显著的提升,同时分别在提出的数据集以及公共数据集与其他模型进行了对比和验证,结果表明该模型都优于其他的模型。

参考文献:

- [1] 徐 鹏,祝 轩,姚 丁,等. 沥青路面养护智能检测与决策综述[J]. 中南大学学报:自然科学版,2021,52(7):2099-2117.
- [2] 马 建,赵祥模,贺栓海,等. 路面检测技术综述[J]. 交通运输工程学报,2017,17(5):121-137.
- [3] 薛红军,周 敏,朱志超,等. 智能养护巡查用 AI[J]. 中国公路,2020(3):66-68.
- [4] 叶家玮,吴 鹏,郑国梁. 基于路面雷达的路面结构缺陷检测方法[J]. 华南理工大学学报:自然科学版,2004(9):82-85.
- [5] 啜二勇. 国外路面自动检测系统发展综述[J]. 交通标准化,2009(17):96-99.
- [6] KIRSCHKE K, VELINSKY S. Histogram-based approach for automated pavement-crack sensing[J]. Journal of Transportation Engineering, 1992, 118(5):700-710.
- [7] OLIVEIRA H, CORREIA P L. Automatic road crack segmentation using entropy and image dynamic thresholding[C]//2009 17th European signal processing conference. Glasgow: European Association for Signal, Speech, and Image Processing, 2009:622-626.
- [8] KAMALIARDAKANI M, SUN L, ARDAKANI M K. Sealed-crack detection algorithm using heuristic thresholding approach[J]. Journal of Computing in Civil Engineering, 2016, 30(1):04014110.
- [9] CHEN P H, SHEN H K, LEI C Y, et al. Support-vector-machine-based method for automated steel bridge rust assessment[J]. Automation in Construction, 2012, 23:9-19.
- [10] GOODFELLOW I, BENGIO Y, COURVILLE A, et al. Deep learning[M]. Cambridge: MIT Press, 2016.
- [11] KIM H, AHN E, SHIN M, et al. Crack and noncrack classification from concrete surface images using machine learning[J]. Structural Health Monitoring, 2019, 18(3):725-738.
- [12] LI S, ZHAO X, ZHOU G. Automatic pixel-level multiple damage detection of concrete structure using fully convolutional network[J]. Computer-Aided Civil and Infrastructure Engineering, 2019, 34(7):616-634.
- [13] 李 画. 基于深度学习的路面缺陷检测方法研究[D]. 长春: 长春大学, 2021.
- [14] 陈旭丹,姜 男,董丽娜,等. 基于深度学习 YOLO 模型的高原山区沥青路面病害自动识别[J]. 公路交通科技:应用技术版, 2019, 15(11):75-78.
- [15] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Honolulu: IEEE, 2017:2117-2125.
- [16] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6):1137-1149.
- [17] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Washington: IEEE, 2018:7132-7141.
- [18] MAEDA H, SEKIMOTO Y, SETO T, et al. Road damage detection and classification using deep neural networks with smartphone images[J]. Computer-Aided Civil and Infrastructure Engineering, 2018, 33(12):1127-1141.
- [19] 刘宪明,辛公锋. 国内基于深度学习的道路路面病害检测研究综述[J]. 电子世界, 2021(8):96-98.