文章编号:1674-2974(2024)12-0107-12 DOI:10.16339/j.cnki.hdxbzkb.2024288

多病害并发复杂场景下的道路病害检测RGT-YOLOv7模型

罗向龙,王彦博†,蒲亚亚,刘若辰
(长安大学 信息工程学院,陕西 西安 701164)

摘 要:随着我国公路网的不断扩展,道路病害检测已经成为道路养护与行车安全保障必不可少的组成部分,基于深度学习的道路病害检测已经成为该领域的研究热点.针对多种病害并发的复杂场景下道路病害识别精度不高、泛化能力不足的问题,提出了一种复杂场景下的道路病害检测模型RGT-YOLOv7(Receptive Ghost Triplet-YOLOv7).在主干网络部分引入三重注意力机制,提高病害特征在不同通道与空间的相关性,解决了特征提取效率不高的问题;将原有的全连接空间金字塔卷积模块替换为快速全连接空间金字塔卷积模块,并加入幻影卷积模块,提高冗余特征的使用率,将原有的冗余特征与新提取到的特征融合,得到包含不同尺度的特征信息;为了扩大模型感受野,在特征增强部分加入改进的感受野模块,利用不同尺寸的空洞卷积从不同方向对特征图进行提取,加强对横向和纵向特征的提取.实验结果表明,与YOLOv7相比,识别的平均正确率(mean average precision,mAP)和平衡F分数分别提升了6.9、3.9个百分点,尤其是对纵向裂缝危害识别的平均正确率提高了22.3个百分点,与Faster R-CNN、YOLOv5等模型相比也有良好的性能提升,表明RGT-YOLOv7是一种有效的复杂场景下的道路病害检测模型.

关键词:目标检测;道路病害检测;深度学习;YOLOv7;RFB
中图分类号:TP391 **文献标志码**:A

Road Disease Detection RGT-YOLOv7 Model under Multiple Diseases Complicated Scenarios

LUO Xianglong,WANG Yanbo†,PU Yaya,LIU Ruochen
(School of Information Engineering, Chang'an University, Xi'an 710064, China)

Abstract: With the continuous expansion of China's road network, road disease detection has become an indispensable part of road maintenance and traffic safety, and road disease detection based on deep learning has become a research hotspot in this field. Aiming at the problems of low accuracy and generalization ability of road disease identification in complex scenes with multiple diseases, a road disease detection model called Receptive Ghost Triplet-YOLOv7 (RGT-YOLOv7) in complex scenes is proposed in this paper. A triplet attention mechanism

* **收稿日期**:2023-12-08
基金项目:国家自然科学基金资助项目(62001059), National Natural Science Foundation of China(62001059);陕西省自然科学基金资助项目(2022JM-056), Natural Science Foundation of Shaanxi Province(2022JM-056)
作者简介:罗向龙(1978—),男,陕西乾县人,长安大学教授,博士
† 通信联系人,E-mail:546249828@qq.com

is introduced in the backbone network to improve the correlation of disease features in different channels and spaces, and to solve the problem of low feature extraction efficiency. The original SPP module is replaced by the SPPF module, the Ghost module is added to improve the utilization rate of redundant features, and the original redundant features and the newly extracted features are fused to get more diverse and rich feature information with different scales. In order to improve the model perception field, RFBs module is added in the feature enhancement part, and the feature map is extracted from different directions by using cavity convolution with different sizes to enhance the extraction of horizontal and vertical features. Experimental results show that the average accuracy and balanced F score are improved by 6.9 percentage points and 3.9 percentage points, respectively, compared with YOLOv7, especially the longitudinal fracture identification is improved by 22.3 percentage points, and it also has good performance improvement compared with Faster R-CNN, YOLOv5, and recently proposed algorithm models. It is an effective road disease detection algorithm for proposed RGT-YOLOv7 under complex scenes.

Key words: object detection; road distress detection; deep learning; YOLOv7; receptive field block (RFB)

公路是促进国家发展与提高人民生活质量的重要基础设施，随着我国公路网规模的不断增大以及车流量的飞速增长，公路因自然因素与长时间的使用不可避免会出现不同类型的路面病害，这已成为威胁行车安全的严重隐患之一. 近些年伴随着信息技术的蓬勃发展，如何利用信息化的手段，特别是通过图片对路面病害进行自动检测与识别成为公路养护管理领域研究重点.

目前国内外针对道路病害检测研究方法主要分为基于传统特征提取的图像分割算法和基于深度学习的方法两大类. Ding等[1]为解决强光条件下车辙检测困难，采用全局灰度校正、特征融合分割和自适应中值滤波算法减少了路面标记的干扰，很好保留了线结构光条，通过最小二乘法对光心曲线进行校正，得到了反映路面实际变形的相对光心曲线，测试结果表明该方法的平均相对检测误差明显低于人工检测. Zhang等[2]针对沥青路面裂缝检测困难的问题，利用自适应阈值分割图片中的裂缝信息，提出了一种基于灰度相近空间聚类像素的裂缝特征提取方法，引入了信念区域(Region of Belief, ROB)的概念，通过定义可信度因子提高检测速度，使用区域增长算法，以ROB种子作为起点对不同区域进行搜索，最终实现道路裂纹检测，并在自建的数据集上表现出了较好的识别率. 基于传统特征提取图像分割的路面病害检测算法更关注图像分割任务，并未对识别结果进行分类，且过于依赖病害特征的提取效果，而在真实场景中存在诸多因素影响特征提取结果，从而导致在复杂环境中识别准确率低，限制了其在实际场景中的应用.

随着信息技术的不断发展，机器学习特别是深度学习研究浪潮的兴起，基于深度学习的路面病害检测成为该领域的热点. Song等[3]为提高路面损伤检测精度和速度对原有的区域卷积神经网络(region-based convolutional neural networks, R-CNN)[4]进行改进，提出了一种快速区域卷积神经网络(faster region-based convolutional neural networks, Faster R-CNN)[5]的新型检测方法，可自动识别和定位路面裂缝、坑洞、溢油等路面病害，测试结果表明Faster R-CNN相较于卷积神经网络(convolutional neural network, CNN)平均检测精度提高了2.1%. 肖力炀等[6]针对现有裂缝识别算法中识别精度不高的问题，提出了一种基于改进型掩膜区域卷积神经网络(mask region-based convolutional neural network, Mask R-CNN)模型的路面裂缝识别方法，采用CNN及特征金字塔网络(feature pyramid network, FPN)的结构进行图像特征提取，通过从上到下的下采样和从下到上的上采样，获得不同尺度的特征信息，最后使用连接全卷积网络来对特征图中的每个像素点进行分类，其检测精度相比于Faster R-CNN提高了6%. 基于深度学习的目标检测，通常分为单阶段模型和双阶段模型两大类，而基于Faster R-CNN的路面病害检测属于双阶段模型，该类模型在精度上虽有很大优势，但检测速度较慢，难以满足实际工程中实时性的要求.

为了提高路面病害检测的速度，同时保证检测精度，单阶段目标检测模型成为目前的热门研究方法. Yi 等[7]提出了一种基于改进 YOLOv7(you only look once version 7, YOLOv7)[8]的高效路面破损检测方法，首先采用无参数注意力机制(simple attention machanism, SimAM)[9]关注模块对特征图像进行加权处理，以幻影卷积(ghost convolution, Ghost)模块代替部分深度卷积模块，提高模型预测精度和速度,在公开数据集上的测试结果与其他传统模型对比,平均正确率(mean average precision, mAP)和平衡 F 分数(balanced F score, F_1)都有所提升. Liu 等[10]以 YOLOv5(you only look once version 5, YOLOv5)作为基准模型，提出了一种改进的YOLOv5模型来解决道路病害检测问题，通过改进的快速空间金字塔池化(spatial pyramid pooling-FAST, SPPF)模块来提高特征融合能力，增加卷积块注意力模块(convolutional block attention module, CBAM)关注关键信息，使用双三次插值代替上采样层的最近邻插值，缓解由于上采样减小图像大小导致的图像特征丢失过多问题,最终在GRDDC2020数据集上测试，平均精度达到58.87%.何铁军等[11]为进一步提升道路病害检测精度,在YOLOv5基础上提出了针对路面病害特征改进的检测模型 PD-YOLO (pavement damage-YOLO, PD-YOLO).将原有的空间金字塔池化与全连接空间金字塔卷积(spaitial pyramid pooling and fully connected spatial pyramid convolution, SPPCSPC)模块替换为快速空间金字塔池化与全连接空间金字塔卷积(spaitial pyramid pooling-FAST and fully connected spatial pyramid convolution, SPPFCSPC)模块,增强了特征提取效果，在特征融合层引入自适应空间特征融合(adaptively spatial feature fusion, ASFF)模块增强不同特征元素直接的联系,提高对道路病害检测的敏感度. 在RDD2020数据集上的平均检测精度相较YOLOv5模型提高了4.8%,达到了55.9%,但在其他数据集上精度有所降低.虽然深度学习的单阶段检测算法在道路病害识别方面有了很大的提升,但在多种不同病害目标并存的场景下,模型的检测精度及鲁棒性仍需提高.

针对多类病害并发的复杂场景下道路病害识别精度不高、模型泛化能力不足的问题,本文基于YOLOv7模型提出一种新的道路病害检测模型(receptive ghost triplet-YOLOv7,RGT-YOLOv7).为了加强特征提取效率,突出路面病害特征，在主干网络特征输出时引入三重注意力(Triplet Attention)机制,增强病害特征在不同通道与空间的相关性;为提高特征融合效果，使用SPPFCSPC模块替换SPPCSPC模块，并加入幻影卷积模块增强对冗余特征的利用率，在保留原有的小尺度特征的基础上提高对宏观特征的提取，通过Ghost模块的残差链接将原有特征与新提取特征融合，增强上下文特征信息之间的依赖关系;为了提高模型泛化能力，增强对横向、纵向裂缝病害的识别能力，在特征增强部分加入改进的感受野模块(receptive field block, RFB)，通过不同尺寸的空洞卷积进行横向与纵向的特征提取，扩大模型感受野，提高特征信息的全局相关性. 在公开道路病害数据集RDD2022上对提出的RGT-YOLOv7模型进行测试，结果表明在多类道路病害并发的复杂场景下具有良好检测效率，检测精度提升显著，是一种有效的复杂场景道路病害检测方法.

1　相关理论

1.1　YOLOv7模型

YOLO(you only look once, YOLO)算法作为深度学习目标检测单阶段算法的经典代表，以其运行速度快、分类精度高被广泛应用于目标检测的相关工作中. YOLOv7作为YOLO系列新推出的算法，集成了YOLOv5[12]的跨网格搜索及YOLOX[13]的匹配策略精华，采用新的标签分配策略，有效地将每个目标与其所属的网格相对应，降低标签分配误检率;同时利用重参数VGG网络(re-parameterization visual geometry group, REPVGG)[14]中重参数化的思想，将重参数化引入网络结构中，并提出了辅助头训练法，在保证精度提升的基础上不影响效率;借鉴高效资源计算网络VoVNet(vortex of vectorized neural network)[15]和跨级局部网络(cross stage partial network, CSPNet)[16]的思想，通过控制最长梯度路径，提取到更多有效特征，让网络具有更强的鲁棒性. YOLOv7网络主要由输入(Input)、主干(Backbone)网络、颈部(Neck)网络和头部(Head)网络组成，网络结构如图1所示.

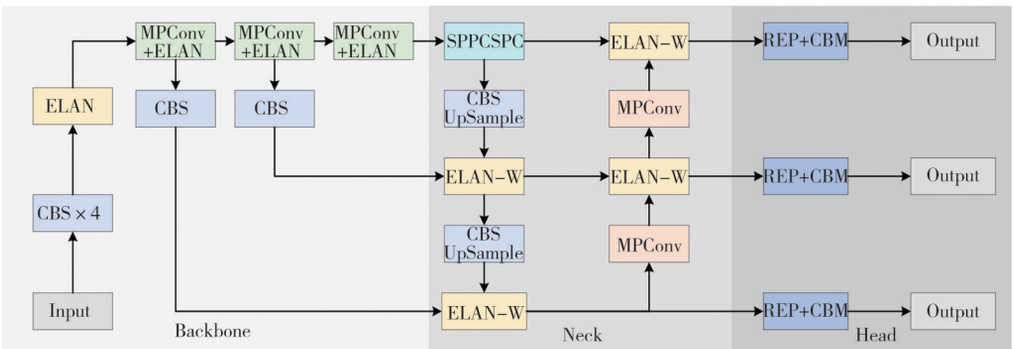


图 1 YOLOv7 结构图

Fig.1 YOLOv7 structure diagram

1.2 Ghost 卷积

传统卷积为了提取更加全面的特征,通常会使用卷积核对特征图进行反复卷积,这会产生大量冗余特征,对模型精度影响较大且计算量庞大.Ghost 卷积是由 Han 等[17]提出的一种轻量化卷积模块,原理图如图 2 所示.首先使用普通卷积产生少量特征信息,然后对特征信息进行线性运算,产生新的特征信息,最后将新的特征信息与之前普通卷积产生的特征信息拼接,得到与传统卷积输出数量相同的特征图.Ghost 模块减少了模型计算量的同时利用相似度高、数量多的冗余特征提升特征提取能力.

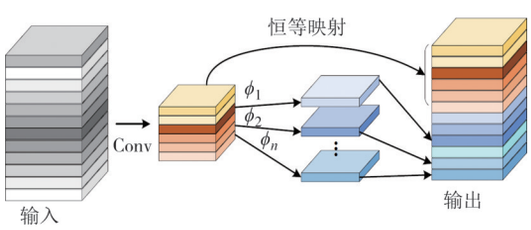


图 2 Ghost 模型原理图

Fig.2 Schematic diagram of Ghost model

假设输入特征图为 $\boldsymbol{X}^{h\times w\times c}$,$h$、$w$ 和 c 分别为特征图高、宽和通道数,经过第一步卷积计算后输出特征图为 $\boldsymbol{Z}^{h'\times w'\times n}$,$h'$、$w'$ 和 n 分别为特征图高、宽和通道数,常规卷积核大小为 k,线性变化卷积核大小为 d,则经过 s 次变换后普通卷积与 Ghost 卷积的计算量比值如式(1)所示:

$$r_s=\frac{n\cdot h'\cdot w'\cdot c'\cdot k\cdot k}{\frac{n}{s}\cdot h'\cdot w'\cdot c\cdot k\cdot k+(s-1)\cdot\frac{n}{s}\cdot h'\cdot w'\cdot d\cdot d}\approx s \tag{1}$$

由式(1)可以看出 Ghost 卷积计算量约为普通卷积 $\frac{1}{s}$,Ghost 模块相比普通卷积减少了模型计算量,降低了模型复杂度,同时保留了原始特征图中的部分特征信息,输出与普通卷积相同的特征图维数;同时利用了普通卷积产生的冗余特征提升特征提取能力,产生了更多与原特征图相关的特征信息.

1.3 Triplet 注意力机制

YOLOv7 模型结构复杂、提取特征种类与数量多,要准确地关注到路面病害特征比较困难,经常因为模型层数的增加忽视许多有用的细小特征,对病害检测精度有很大的影响,针对以上问题,本文引入 Triplet 注意力[18]机制,通过三分支结构捕获不同维度与通道之间的特征相关性,利用旋转操作和残差结构建立维度间的依存关系,保留有效特征信息的同时细化需要关注的特征信息,提高识别分类性能.Triplet 注意力原理图如图 3 所示,顶部分支负责计算通道维度 c 与空间维度 w 间的跨通道交互关系,中部分支负责计算通道维度 c 与空间维度 h 间的交互权重,底部分支思路与空间注意力类似,用于计算空间维度 h 和 w 的依赖关系,最后采用平均计算对三条分支输出的权重进行聚合.

2 RGT-YOLOv7 路面病害检测模型

2.1 Ghost-SPPFCSPC 模块

YOLOv7 中的 SPPCSPC 模块由空间金字塔池化(spatial pyramid pooling, SPP)模块和全连接空间金字塔卷积(fully connected spatial pyramid convolution, CSPC)结构组成,SPP 通过对特征图的不同尺度池化,获得不同尺度的感受野,有效解决了图像由于裁剪、缩放引起的特征细节丢失问题,CSPC 结构给原有的 SPP 模块加上了残差结构,提升特征提取效果,但由于传统卷积在特征提取方面的局限性,许多病害特征无法得到有效的利用.本文借鉴 YOLOv5 中的 SPPF 层结构,引入 Ghost 模块,提出 Ghost-SPPFCSPC 模块替换原有的 SPPCSPC 模块,模块结构如图 4 所示.

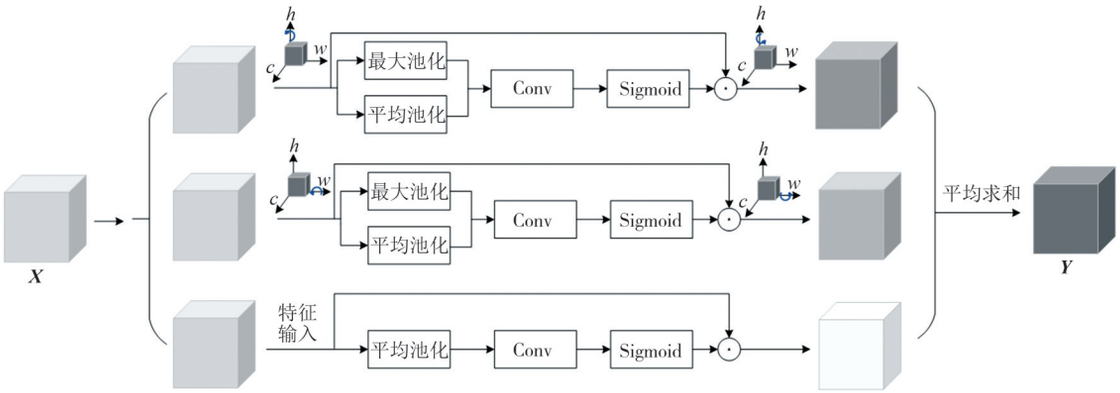


图 3　Triplet 注意力原理图

Fig.3　Triplet attention schematic diagram

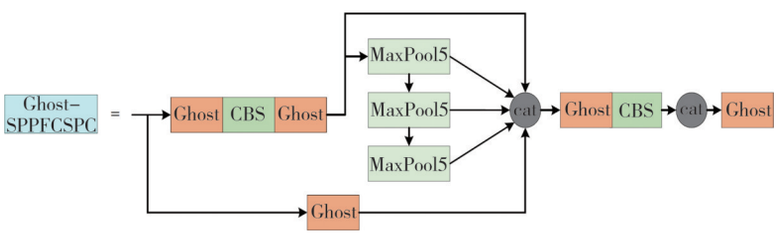


图 4　Ghost-SPPFCSPC 模块结构

Fig.4　Ghost-SPPFCSPC module structure

首先将部分卷积替换为 Ghost 模块，强化特征提取时大量冗余特征利用率，提高特征提取效率，其次用三个池化核为 5 的最大池化层替换原池化核为 5、9、13 的三个最大池化层，在保留细节特征信息的前提下，增强宏观特征的提取，提升模块运算速度，将 Ghost 模块中提取出的特征信息进行多次池化操作，获得不同感受野下的特征信息，最后将这些特征信息同残差链接输出拼接起来，从不同尺度获得更多类型的特征信息，极大地扩大模型感受野，提升模型识别精度.

2.2　RFBs 模块

YOLOv7 网络结构复杂，输入道路病害图像在经过大量卷积后会产生许多与路面病害无关的冗余特征向量，对所需特征的提取与模型鲁棒性有很大影响，因此本文在特征增强阶段引入 RFBs 模块，增强病害特征的分辨能力，以广阔的感受野提升病害特征的全局相关性.RFBs 模块结构如图 5 所示，RFBs 模块是借鉴 Inception[19] 模块思想提出的，利用三种不同尺寸的空洞卷积构成多分支结构，加入 1×1 卷积核减少运算量，使用卷积核尺寸为 1×3 与 3×1 的空洞卷积分别从横向和纵向提取特征，提高横向裂缝和纵向裂缝的识别准确率，扩大特征提取感受野，更好地融合不同尺度特征信息，特征融合部分增加残差链接，将输入特征与空洞卷积组提取的特征充分融合，减少病害特征信息的丢失.

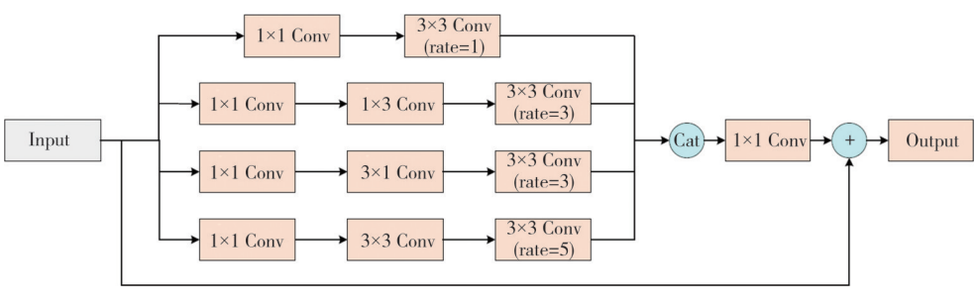


图 5　RFBs 模块结构图

Fig.5　RFBs module structure diagram

2.3　RGT-YOLOv7 模型

道路病害包括裂缝、坑槽、松散、塌陷等多种类型，实际工程中要求准确识别各类病害，及时采取不同措施修复维护道路，但在拍摄的路面图片中，

经常出现多种病害并存的复杂场景，此外，每种病害尺寸还经常发生变化，例如纵向裂缝呈现尺寸变化的细长条状. YOLOv7 网络是目标检测常用算法之一，但检测目标大多有固定形状特征，在特征检测时按照[12,16,19,36,40,28]、[36,75,76,55,72,146]、[142,110,192,243,459,401]三种锚框尺寸对目标进行特征提取，而对于路面病害尤其是裂缝病害形状特征比较特殊的细长条状，传统的特征提取受限于锚框尺寸固定而出现特征提取位置不精准，导致现有的目标检测模型难以有效提取病害特征，经常出现漏检，导致模型检测精度不高，泛化能力差. 针对现有模型特征提取效率不高，多种病害并存难以全部识别，如何同时提高检测精度与模型泛化性的问题，本文提出一种基于 YOLOv7 的改进模型 RGT-YOLOv7，模型结构图如图 6 所示.

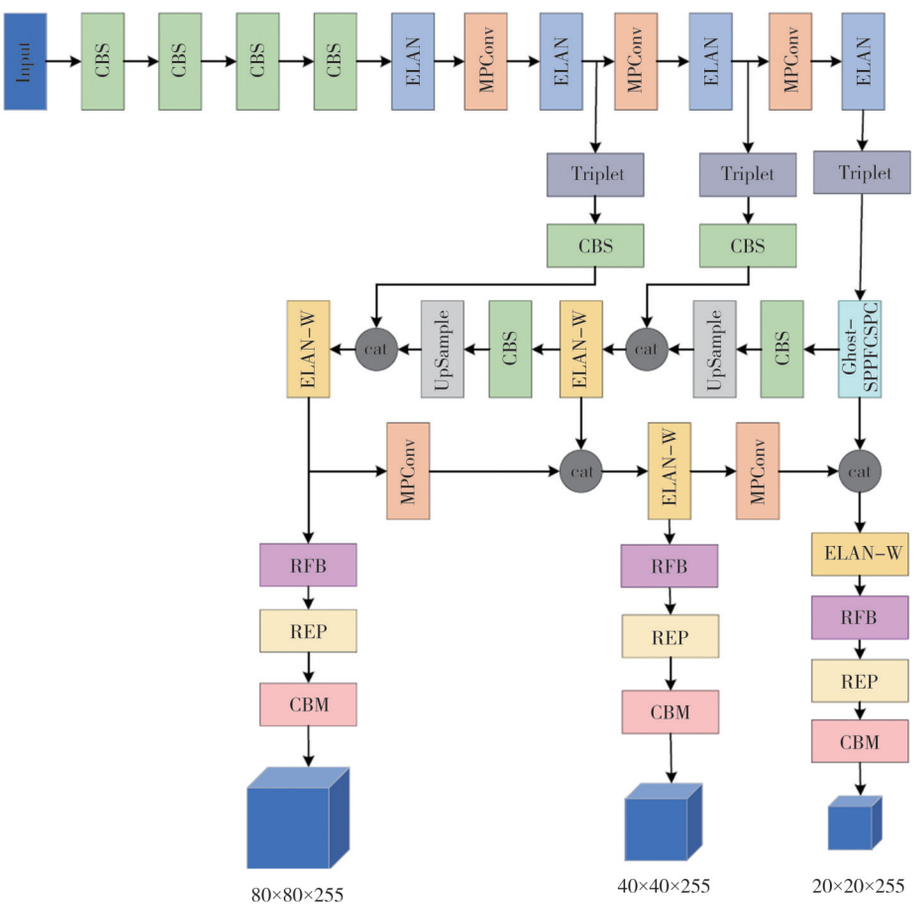


图 6 RGT-YOLOv7 模型结构图

Fig.6 RGT-YOLOv7 model structure diagram

针对 YOLOv7 模型对于所需特征信息敏感度不高的问题，在主干特征提取层三个输出位置引入 Triplet 注意力模块，在不降低维度的前提下利用多维交互提高路面病害特征重要性，提高网络特征提取能力，解决因网络不断加深，特征信息消失的问题；针对传统卷积计算量大，无法利用冗余特征产生更有效特征信息的问题，本文设计了一种新的 Ghost-SPPFCSPC 模块，通过利用 Ghost 模块替换部分卷积的方式，将普通卷积产生的大量相似的冗余特征充分利用，将原有的多种病害特征保留，强化特征之间的相关性；针对 YOLOv7 模型对纵向裂缝检测精度不高的问题，在三个不同尺度的特征输出位置加入 RFBs，模块利用三个不同尺寸的空洞卷积分别从横向和纵向提取病害信息，加强横向裂缝与纵向裂缝的特征能力，同时扩大模型感受野，提高病害特征的全局相关性，提高模型泛化能力.

假设 $\boldsymbol{X}$ 为输入图片，$\boldsymbol{X} \in \mathbb{R}^{h \times w \times c}$，具体如式(2)所示，$h$、$w$ 和 c 分别为输入图片高、宽和通道数；n 表示训练集的大小；$\varphi_i(\cdot)$ 是激活函数；$\boldsymbol{Z}_{il}$ 与 $\tilde{\boldsymbol{Z}}_{il}$ 为第 i 组特征输入经过不同卷积步长后的输出；$\boldsymbol{Z}_n$ 表示每张图片的特征输出，具体计算如式(3)所示：

$$\boldsymbol{X} = \{\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_i, \cdots, \boldsymbol{X}_n, 1 \leqslant i \leqslant n\} \tag{2}$$

$$\begin{cases} \boldsymbol{Z}_{il} = \varphi_i\left(\boldsymbol{W}_{e,i}\mathrm{conv}_{3\times3}\left(\boldsymbol{X}_i\right)\right), l = 1, i = 1, 2, \cdots, n \\ \tilde{\boldsymbol{Z}}_{il} = \varphi_i\left(\boldsymbol{W}_{e,i}\mathrm{conv}_{3\times3}\left(\boldsymbol{Z}_{il}\right)\right), l = 2, i = 1, 2, \cdots, n \\ \boldsymbol{Z}_n = \left[\tilde{\boldsymbol{Z}}_{12}, \tilde{\boldsymbol{Z}}_{22}, \cdots, \tilde{\boldsymbol{Z}}_{n2}\right] \end{cases} \tag{3}$$

式中:$\boldsymbol{W}_{e,i}$为经过BatchNorm层产生的输入权重;e代表对特征图进行池化的维度;l为卷积步长.

输入图片在经过四个卷积核大小一致、卷积步长不同的CBS模块后,将提取出的特征信息送入ELAN模块中进行特征增强,ELAN模块通过对不同特征组合处理,得到更多与原特征相关的特征信息,同时可以保留特征的梯度信息,具体计算如式(4)所示:

$$\begin{cases} \tilde{\tilde{\boldsymbol{Z}}}_{il} = \varphi_i\left(\boldsymbol{W}_{e,i}\mathrm{conv}_{1\times1}\left(\boldsymbol{Z}_{i2}\right)\right), l = 1,\ i = 1, 2, \cdots, n \\ \hat{\boldsymbol{Z}}_{il} = \tilde{\boldsymbol{Z}}_{il}\left(\tilde{\boldsymbol{Z}}_{il}\left(\tilde{\tilde{\boldsymbol{Z}}}_{i1}\right)\right), l = 2,\ i = 1, 2, \cdots, n \\ \bar{\boldsymbol{Z}}_{il} = \tilde{\boldsymbol{Z}}_{il}\left(\tilde{\boldsymbol{Z}}_{il}\left(\tilde{\boldsymbol{Z}}_{il}\left(\tilde{\tilde{\boldsymbol{Z}}}_{i1}\right)\right)\right), l = 2,\ i = 1, 2, \cdots, n \\ \boldsymbol{Y}^i = \tilde{\tilde{\boldsymbol{Z}}}_{i1} + \hat{\boldsymbol{Z}}_{i2} + \bar{\boldsymbol{Z}}_{i2}, i = 1, 2, \cdots, n \end{cases} \tag{4}$$

式中:$\tilde{\tilde{\boldsymbol{Z}}}_{il}$为经过卷积步长为$l$的CBS模块后的特征信息输出;$\hat{\boldsymbol{Z}}_{il}$与$\bar{\boldsymbol{Z}}_{il}$为经过不同次CBS模块后的特征信息输出;$\boldsymbol{Y}^i$为特征信息经过ELAN层增强后的特征输出;$\boldsymbol{Y}^i$亦作为Triplet模块特征输入,通过三分支结构捕获不同维度与通道之间的特征相关性,利用旋转操作和残差结构建立维度间的依存关系,具体计算如式(5)和式(6)所示:

$$\boldsymbol{Z} - \mathrm{Pool}_e\left(\boldsymbol{Y}_i\right) = \left[\mathrm{MaxPool}\left(\boldsymbol{Y}_i\right), \mathrm{AvgPool}\left(\boldsymbol{Y}_i\right)\right] \tag{5}$$

$$\begin{cases} \boldsymbol{T}_w = \varphi_i\left(\boldsymbol{W}_{e,i}\mathrm{conv}_{7\times7}\left(\boldsymbol{Z} - \mathrm{Pool}_w\left(\boldsymbol{Y}_i\right)\right)\right) \\ \boldsymbol{T}_h = \varphi_i\left(\boldsymbol{W}_{e,i}\mathrm{conv}_{7\times7}\left(\boldsymbol{Z} - \mathrm{Pool}_h\left(\boldsymbol{Y}_i\right)\right)\right) \\ \boldsymbol{T}_c = \varphi_i\left(\boldsymbol{W}_{e,i}\mathrm{conv}_{7\times7}\left(\boldsymbol{Z} - \mathrm{Pool}_c\left(\boldsymbol{Y}_i\right)\right)\right) \\ \boldsymbol{T} = \boldsymbol{T}_w + \boldsymbol{T}_h + \boldsymbol{T}_c \end{cases} \tag{6}$$

式中:e代表对特征图进行池化的维度;$\boldsymbol{T}_w$、$\boldsymbol{T}_h$、$\boldsymbol{T}_c$代表对输入特征分别进行空间维度h与通道维度c、空间维度w与通道维度c、空间维度h与w之间特征相关性捕获;$\boldsymbol{T}$代表三条分支特征融合输出.

将与路面病害具有很高相关性的特征信息$\boldsymbol{T}$作为Ghost-SPPFCSPC模块特征输入,通过Ghost模块的冗余特征利用,保留特征信息$\boldsymbol{T}$中很多高效特征信息,并通过残差链接与多尺度的其他特征信息融合产生出更多具有丰富特征的病害特征信息,将Ghost-SPPFCSPC模块输出与其他尺度特征拼接与特征增强后的输出特征$\boldsymbol{G}$作为RFBs模块输入,通过不同卷积核大小的空洞卷积对输入特征进行多尺度的特征提取,扩大了模型感受野,增强特征的全局相关性与模型泛化能力,具体计算如式(7)所示:

$$\begin{cases} \boldsymbol{R}_1 = \xi_s\left(\mathrm{conv}_{1\times1}\left(\boldsymbol{G}\right)\right), s = 1 \\ \boldsymbol{R}_2 = \xi_s\left(\mathrm{conv}_{1\times3}\left(\mathrm{conv}_{1\times1}\left(\boldsymbol{G}\right)\right)\right), s = 3 \\ \boldsymbol{R}_3 = \xi_s\left(\mathrm{conv}_{3\times1}\left(\mathrm{conv}_{1\times1}\left(\boldsymbol{G}\right)\right)\right), s = 3 \\ \boldsymbol{R}_4 = \xi_s\left(\mathrm{conv}_{3\times3}\left(\mathrm{conv}_{1\times1}\left(\boldsymbol{G}\right)\right)\right), s = 5 \\ \boldsymbol{R}_5 = \boldsymbol{R}_1 + \boldsymbol{R}_2 + \boldsymbol{R}_3 + \boldsymbol{R}_4 \\ \boldsymbol{R} = \mathrm{conv}_{1\times1}\left(\boldsymbol{R}_5\right) + \boldsymbol{G} \end{cases} \tag{7}$$

式中:$\boldsymbol{R}_i = \left\{\boldsymbol{R}_1, \boldsymbol{R}_2, \cdots, \boldsymbol{R}_5, 1 \leqslant i \leqslant 5\right\}$为RFBs模块在不同分支上的特征输出;$\boldsymbol{R}$为RFBs模块对不同分支特征提取结果进行融合后与残差链接的特征输出;$\xi_s(\cdot)$为空洞卷积;s为空洞卷积扩张率.

对其他两条特征提取分支进行相同操作,通过上采样产生不同尺度大小的特征映射,最终输出三种不同尺度的病害特征映射信息,利用其实现路面病害的分类与识别.

3　实验及结果分析

3.1　数据集介绍

本文使用数据集为开源数据集RDD2022[20],主要来自中国、日本、美国、捷克、挪威和印度等国家收集的道路破损图像,路面病害类型有D00(纵向裂缝)、D10(横向裂缝)、D20(网状裂缝)、D40(坑洞)四种,以来自中国的4 378张和日本的10 506张图片,共计14 884张图片作为测试数据集,针对上述四种常见路面病害进行训练识别,使用随机划分的方法按照8∶2的比例划分训练集和验证集,最终划分结果训练集11907张图片,验证集2 977张图片.

3.2　实验环境及参数

本文实验是在装载Ubuntu18.04的Linux操作系统下进行的,实验所使用的硬件配置为:GPU为3块11 GB显存的NVIDIA GeForce RTX 2080Ti,CPU为Intel Xeon E5-2620 v4@ 2.10 GHz,32 GB的运行内存.编程环境:Python 3.7以及Pytorch 1.7版本的深度学习框架,训练时使用YOLOv7预训练权重,设置Batch size为4,Epoch为100.

3.3 评价指标

为评估算法模型的效果，通常采用召回率(R)、准确率(P)、平衡F分数(F_1)、平均正确率(mAP@0.5)、帧率(FPS)、每秒浮点运算次数(FLOPs)和模型参数量作为评价指标，其中召回率、准确率、平衡F分数、平均正确率、帧率越高代表模型检测效果越好，浮点运算量和模型参数量越低代表模型复杂度越低. 召回率代表正确检测样本在总样本中的比例，准确率代表正确检测样本在所有检测到的样本中的比例，F_1分数为二者加权平均，mAP是衡量模型准确性优劣的重要标准，计算公式分别如式(8)~式(11)所示：

$$R = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{8}$$

$$P = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{9}$$

$$F_1 = \frac{2 \times P \times R}{P + R} \tag{10}$$

$$\mathrm{mAP} = \frac{\sum_{i=1}^{K} \mathrm{AP}_i}{K} \tag{11}$$

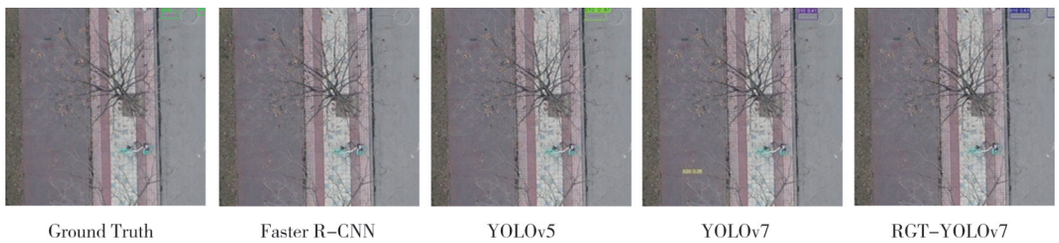
式中：TP代表被网络检测到并正确分类的病害数量；FN代表属于病害但是没有被正确检测出来的病害数量；FP代表被错误识别成病害的背景区域数量；AP表示P与R所围曲线的面积；K为检测目标类别数.

3.4 实验结果分析

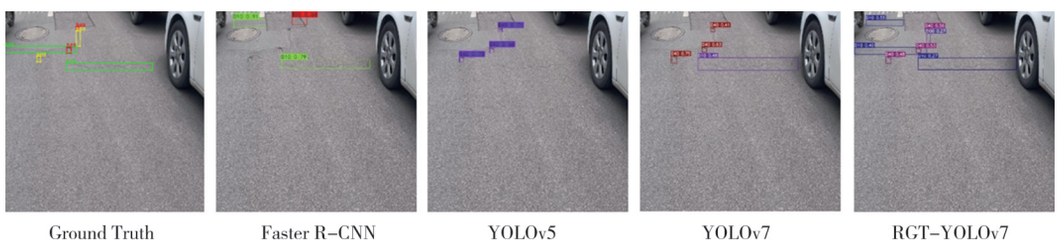
3.4.1 对比实验

为了验证所提出模型在多种道路病害并发的复杂场景下检测识别的有效性和可行性，在相同设备环境条件下，分别选取了病害目标较小、多类病害并发、病害形状特殊、病害图像光照不足四种多病害并发复杂场景下的道路病害图片的检测结果，并与目前一些主流深度学习目标检测模型进行对比，结果分别如图7所示. 图7(a)中Faster R-CNN漏检了位于图像右上角的横向裂缝，YOLOv5与YOLOv7均检测出了裂缝病害，但YOLOv7错将树枝识别成了网状裂缝，RGT-YOLOv7模型拥有更广阔的感受野，不仅检测出了其他模型检测出的横向裂缝，还检测出了位于井盖右边的部分横向裂缝，并且识别准确度更高；图7(b)中Faster R-CNN识别出了部分横向裂缝，但未识别出三个坑槽，YOLOv5仅识别出了坑槽，但裂缝病害均未识别成功，YOLOv7相较前两种方法有所提升，但仍忽略了很多横向裂缝，RGT-YOLOv7识别出了每种病害并且有比较高的准确率；图7(c)中Faster R-CNN识别效果较好，但仍对同一目标进行了多次识别，YOLOv5未检测出左侧纵向裂缝病害，YOLOv7未检测出右侧纵向裂缝病害，该纵向裂缝占整张图片很大的部分，可能因为感受野受限或特征提取不充分，YOLO系列算法都对纵向裂缝有漏检现象，RGT-YOLOv7利用两种大小不同的卷积核分别从纵向和横向提取病害特征信息，因此准确地识别出了图7(c)中各类病害并准确地画出了每一个锚框；图7(d)中Faster R-CNN识别出了位于图像中间部分的两条横向裂缝，未能识别图像上方较小的横向裂缝，YOLOv5未识别出路面病害，YOLOv7漏检了其中一条颜色较浅的横向裂缝，RGT-YOLOv7提取病害特征时联系全局上下文信息，有较好的特征信息利用率，因此很好地将三条裂缝病害识别成功. 在真实场景中，经常会在某一路段交织出现多种病害，从四组病害图像中可以看出RGT-YOLOv7模型在多种病害并发的复杂场景中的病害检测相较于其他现有模型有更高的识别精度.

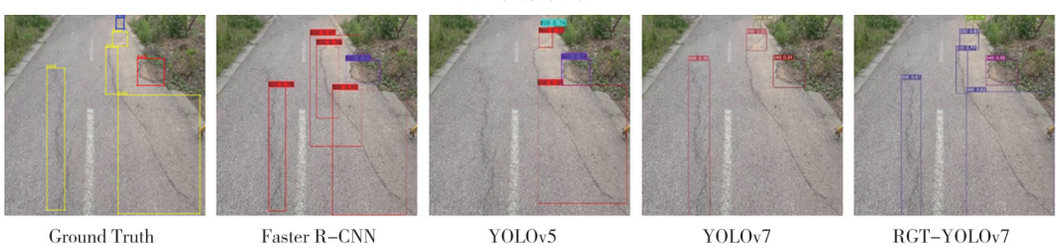
为了进一步说明本文所提出模型在道路病害检测中的有效性，对本文模型与目前现有的主流和最新提出的方法在相同数据集中进行对比，对比结果如表1所示. 从表中可以看出其他网络模型mAP均值为52.8%，F_1值平均为54.3%，本文提出算法相较其他网络模型mAP平均提高8.96个百分点，F_1值平均提高7.07个百分点，在mAP与F_1值上都高于其他模型，尤其是对比最新提出的YOLO-LRDD[21]、LE-YOLOv5[22]、PD-YOLO[10]、MN-YOLOv5[23]、GSC-YOLOv5[24]五种方法，本文提出的RGT-YOLOv7模型在准确率上相较YOLO-LRDD、PD-YOLO、MN-YOLOv5、GSC-YOLOv5模型分别提高3.2、3.0、5.2、5.2个百分点，相较LE-YOLOv5模型降低1.5个百分点；相较YOLO-LRDD、LE-YOLOv5、PD-YOLO、MN-YOLOv5、GSC-YOLOv5模型召回率分别提高2.2、7.2、5.3、4.4、2.6个百分点；mAP分别提高4.2、4.9、5.9、8.2、7.2个百分点；F_1分数分别提高2.7、3.3、4.2、4.8、3.9个百分点，在帧率上RGT-YOLOv7略低于部分其他模型，但检测精度相对于其他模型显著提升，可应用于多种路面病害并存的复杂环境下的



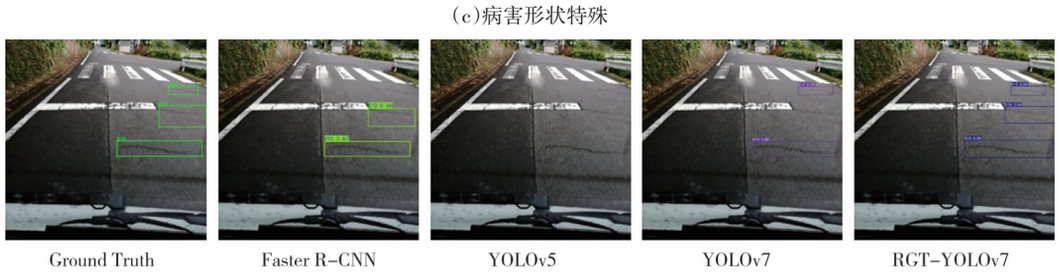
(a)病害目标较小



(b)多类病害并发



(c)病害形状特殊



(d)病害图像光照不足

图7　不同模型测试对比图

Fig.7　Comparison diagram of different models test

路面裂缝检测.

3.4.2　消融实验

为了验证提出模型RGT-YOLOv7用于道路病害检测时各部分的有效性，本文将YOLOv7模型作为基准模型，保证其他参数与实验环境不变的条件下，将各模块分别组合进行实验，在检测精度等方面进行对比，消融实验结果如表2所示.

表1　模型性能对比表

Tab.1　Model performance comparison table

模型	*P*/%	*R*/%	mAP@0.5/%	F_1/%	帧率/(帧·s^{-1})
Faster R-CNN	26.1	45.1	37.4	38.4	19
YOLOv5	56.5	50.4	51.8	50.6	72
YOLOv7	59.7	55.5	54.9	57.5	56
YOLO-LRDD[11]	59.2	58.2	57.6	58.7	86
LE-YOLOv5[12]	63.9	53.2	56.9	58.1	71
PD-YOLO[7]	59.4	55.1	55.9	57.2	56
MN-YOLOv5[23]	57.2	56.0	53.6	56.6	42
GSC-YOLOv5[24]	57.2	57.8	54.6	57.5	53
RGT-YOLOv7	62.4	60.4	61.8	61.4	54

从表2可以看出，在仅添加Ghost模块后，模型参数量降低，mAP没有变化，可能是由于Ghost模块中产生的微小病害特征信息在后续特征提取中未被保留,故特征信息缺失，在后续特征提取中没有起到作用.仅添加SPPFCSPC模块后，模型参数量不变，mAP提高0.8个百分点，可能是因为SPPFCSPC模块利用三个连续的最大池化,扩大模型感受野,加强宏观特征的提取.加入Ghost和SPPFCSPC模块后，mAP和F_1分数分别提高3.0、0.8个百分点，将Ghost模块加入SPPFCSPC后，Ghost模块中利用冗余特征产生的许多细小特征信息在后续的特征提取

表2 消融实验性能对比表

Tab.2 Ablation experiment performance comparison table

组别	网络设置				mAP@0.5/%	F_1/%	Parameters/M	FLOPs/G
	Ghost	SPPFCSPC	RFBs	Triplet				
1	基线网络				54.9	57.5	37.2	105.2
2	√				54.9	56.3	35.8	104.0
3		√			55.7	56.8	37.2	105.2
4			√		59.2	59.6	59.0	145.2
5				√	56.8	57.4	37.2	105.3
6	√	√			57.9	58.3	35.8	104.0
7	√	√	√		59.3	59.8	57.6	144.0
8			√	√	59.4	59.9	59.0	145.2
9	√	√		√	58.3	58.7	35.8	104.1
10	√	√	√	√	61.8	61.4	57.6	144.1

中得以保留，并与不同尺度的特征信息进行融合，得到更多的具有全局意义的特征信息，扩大了模型特征提取感受野，增强特征提取效果. RFBs模块利用三种不同尺寸的卷积核提取三种不同尺度的特征信息，使用残差链接增强提取特征与原始特征之间的相关性，将卷积替换为空洞卷积，更大程度地扩大模型感受野，使模型提取到的特征拥有更多不同的特征信息，因此在添加RFBs模块后，模型mAP与F_1分数相较YOLOv7分别提升4.3、2.1个百分点，检测精度提升明显. 与Ghost-SPPFCSPC模块结合后缩减了部分参数量，并且二者都可以扩大模型感受野，可以更准确地识别病害图像中的网状裂缝、纵向裂缝等形状特征比较大的病害. 加入Triplet注意力模块后提升了特征提取时对于路面病害特征的处理优先级，更高效地提取出各类路面病害的不同特征，有助于区分不同病害类型，减少误检率. 添加Triplet注意力模块后，模型mAP提高1.9个百分点，F_1分数降低0.1个百分点，提升不是特别明显，但加入RFBs与GhostSPPFCSPC模块后，Triplet注意力模块对模型检测精度的提升效果明显，仅添加Triplet注意力模块可能因为后续特征提取的局限性，许多含有丰富特征的信息无法与各个尺度的其他特征信息很好地融合，特征信息与上下文联系不紧密，甚至丢失部分特征，因此提升幅度较小，而加入Ghost-SPPFCSPC与RFBs模块可以很好地解决这个问题，Ghost-SPPFCSPC模块在扩大模型感受野的同时，保留了部分原始特征信息，让Triplet注意力增强提取出的特征信息得以保留，并与不同尺度的其他特征信息充分融合，获得更多不同类别的特征信息，RFBs模块利用卷积核大小不同的空洞卷积分别对图像进行横向和纵向的特征提取，提升模型对横向裂缝、纵向裂缝的特征提取效果，利用多个尺度的不同特征信息相互拼接融合，使特征联系上下文信息，让纵向裂缝的特征能够被很好地提取出来，最终模型mAP与F_1分数分别提高6.9、3.9个百分点，可以看出本文提出算法对于各种尺寸病害的识别精度都有比较大提升.

3.4.3 *泛化性实验*

为验证模型对不同病害类型的识别有效性，本文利用RGT-YOLOv7模型与YOLOv7原始模型对四种常见路面病害进行检测，各类病害平均检测精度结果如表3所示.

表3 多类病害检测mAP@0.5对比

Tab.3 Multi-disease detection mAP@0.5 comparison %

模型	D00	D10	D20	D40
YOLOv7	39.3	54.2	67.8	58.2
RGT-YOLOv7	61.6	56.4	69.5	59.3

从表3可以看出，YOLOv7模型对于D00(纵向裂缝)检测精度仅有39.3%，对数据集进行分析后发现，纵向裂缝经常占据整张图片很大部分，或与网状裂缝共同出现，YOLOv7模型因锚框尺寸与传统特征提取方式的局限性，对这类病害的识别难度比较高，本文提出的RGT-YOLOv7模型相较YOLOv7模型有更大的感受野，分别从横向和纵向提取病害特征信息，有更强的特征提取能力，因此在纵向裂缝识别方面提升比较明显，提高了22.3个百分点，在横向裂缝、网状裂缝、坑槽病害检测方面分别提高了2.2、1.7、1.1个百分点，可以看出本文提出模型在

不同类型路面病害检测方面有着良好的性能提升.

为进一步评估提出的RGT-YOLOv7模型的泛化能力,本文在相同参数设置与环境下,选用公开道路病害数据集RDD2022-美国和RDD2022-印度进行验证,验证结果如表4所示.

表4　其他数据集泛化性实验

Tab.4　Other data set generalization experiments　%

数据集	模型	mAP@0.5	F_1
RDD2022-美国	YOLOv7	50.7	54.3
	RGT-YOLOv7	57.1	55.2
RDD2022-印度	YOLOv7	35.8	40.0
	RGT-YOLOv7	41.7	43.1

从表4可以看出本文提出模型表现仍优于YOLOv7,但因数据集不同,图片质量与数量与原数据集存在差异,因此性能提升也有不同.本文提出算法在美国数据集上mAP与F_1分数相较YOLOv7分别提高了6.4、0.9个百分点,在印度数据集上mAP与F_1分数提高了5.9、3.1个百分点,综合以上结果可以得出,本文提出模型具有较强的泛化能力.

4　结 论

针对道路病害检测中因多种病害并发、影响因素多、病害尺寸不统一、病害数目多等原因导致的模型检测精度低、泛化能力差的问题,本文在YOLOv7模型基础上提出一种增大感受野,提高模型特征提取利用率的检测算法RGT-YOLOv7.在主干网络部分加入Triplet注意力模块提高对路面病害特征提取效率,将SPPCSPC模块改为SPPFCSPC模块,同时引入Ghost模块,保证感受野不变的前提下,保留细小特征,提高对宏观特征的提取效果,并将输入特征通过Ghost模块与残差链接和最终输出特征拼接,提高特征相关性,加入RFBs模块,从多个方向和尺度提取病害特征,增强模型感受野,提高不同方向特征信息提取能力.实验结果显示,相较YOLOv7算法,RGT-YOLOv7模型mAP和F_1分数方面分别提高6.9、3.9个百分点,与YOLOv5、Faster R-CNN和最近提出的新模型相比也有较好的性能提升,并在不同数据集上的测试结果相较YOLOv7都有不错的准确度提升,表明RGT-YOLOv7网络模型具有良好的泛化能力,为多种病害并发的复杂场景下进行道路病害检测提供了一种可能,未来将进一步研究在进一步提升网络检测精度的同时,如何对病害的尺寸进行识别,以更好满足实际工程的需求.

参考文献

[1] DING S, XING Y Y, LANG H, et al. Line-structured light rut detection of asphalt pavement with pavement markings interference under strong light[J]. Journal of Transportation Engineering, Part B: Pavements, 2022, 148(2): 04022007.

[2] ZHANG D J, LI Q Q, CHEN Y, et al. An efficient and reliable coarse-to-fine approach for asphalt pavement crack detection[J]. Image and Vision Computing, 2017, 57: 130-146.

[3] SONG L, WANG X C. Faster region convolutional neural network for automated pavement distress detection[J]. Road Materials and Pavement Design, 2021, 22(1): 23-41.

[4] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition. June 23-28, 2014, Columbus, OH, USA: IEEE, 2014: 580-587.

[5] 薛亚东,高健,李宜城,等. 基于深度学习的地铁隧道衬砌病害检测模型优化[J]. 湖南大学学报(自然科学版),2020,47(7):137-146.
XUE Y D, GAO J, LI Y C, et al. Optimization of shield tunnel lining defect detection model based on deep learning[J]. Journal of Hunan University (Natural Sciences), 2020, 47(7): 137-146. (in Chinese)

[6] 肖力炀,李伟,袁博,等. 一种基于改进实例分割模型的路面裂缝检测方法[J]. 武汉大学学报(信息科学版),2023,48(5):765-776.
XIAO L Y, LI W, YUAN B, et al. A pavement crack identification method based on improved instance segmentation model [J]. Geomatics and Information Science of Wuhan University, 2023, 48(5): 765-776. (in Chinese)

[7] YI C C, LIU J, HUANG T, et al. An efficient method of pavement distress detection based on improved YOLOv7[J]. Measurement Science and Technology, 2023, 34(11): 115402.

[8] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 17-24, 2023, Vancouver, BC, Canada: IEEE, 2023: 7464-7475.

[9] YANG L, ZHANG R, LI L, et al. SimAM: A Simple, Parameter-Free Attention Module for Convolutional Neural Networks[C]// Proceedings of the 38th International Conference on Machine Learning (ICML). June 18-24, Cambridge: PLMR, 2021: 11863-11874.

[10] LIU Y H, DUAN M L, DING G G, et al. HE-YOLOv5s: efficient road defect detection network[J]. Entropy, 2023, 25(9): 1280.

[11] 何铁军, 李华恩. 基于改进YOLOv5的路面病害检测模型[J].

土木工程学报,2024,57(2):96-106.
HE T J, LI H E. Pavement damage detection model based on improved YOLOv5[J]. China Civil Engineering Journal, 2024, 57(2):96-106.(in Chinese)

[12] 余加勇,刘宝麟,尹东,等. 基于YOLOv5和U-Net3+的桥梁裂缝智能识别与测量[J]. 湖南大学学报(自然科学版),2023,50(5):65-73.
YU J Y, LIU B L, YIN D, et al. Intelligent identification and measurement of bridge cracks based on YOLOv5 and U-Net3+[J]. Journal of Hunan University (Natural Sciences), 2023, 50(5): 65-73. (in Chinese)

[13] GE Z, LIU S T, WANG F, et al. YOLOX: exceeding YOLO series in 2021[EB/OL]. https://arxiv.org/abs/2107. 08430v2

[14] DING X H, ZHANG X Y, MA N N, et al. RepVGG: making VGG-style ConvNets great again [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 20-25, 2021, Nashville, TN, USA: IEEE, 2021: 13728-13737.

[15] LEE Y, HWANG J W, LEE S, et al. An energy and GPU-computation efficient backbone network for real-time object detection[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). June 16-17, 2019, Long Beach, CA, USA: IEEE, 2019: 752-760.

[16] WANG C Y, MARK LIAO H Y, WU Y H, et al. CSPNet: a new backbone that can enhance learning capability of CNN[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). June 14-19, 2020, Seattle, WA, USA: IEEE, 2020: 1571-1580.

[17] HAN K, WANG Y H, TIAN Q, et al. GhostNet: more features from cheap operations [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 13-19, 2020, Seattle, WA, USA: IEEE, 2020: 1577-1586.

[18] MISRA D, NALAMADA T, ARASANIPALAI A U, et al. Rotate to attend: convolutional triplet attention module[C]//2021 IEEE Winter Conference on Applications of Computer Vision (WACV). January 3-8, 2021, Waikoloa, HI, USA: IEEE, 2021: 3138-3147.

[19] SZEGEDY C, LIU W, JIA Y Q, et al. Going deeper with convolutions[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 7-12, 2015, Boston, MA, USA: IEEE, 2015: 1-9.

[20] ARYA D, MAEDA H, GHOSH S K, et al. RDD2022: a multi-national image dataset for automatic Road Damage Detection[EB/OL]. https://arxiv.org/abs/2209.08538v1

[21] WAN F, SUN C, HE H Y, et al. YOLO-LRDD: a lightweight method for road damage detection based on improved YOLOv5s[J]. EURASIP Journal on Advances in Signal Processing, 2022(1):98.

[22] DIAO Z, HUANG X F, LIU H, et al. LE-YOLOv5: a lightweight and efficient road damage detection algorithm based on improved YOLOv5[J]. International Journal of Intelligent Systems, 2023: 8879622.

[23] GUO G G, ZHANG Z Y. Road damage detection algorithm for improved YOLOv5[J]. Scientific Reports, 2022, 12(1): 15523.

[24] WANG J H, GAO X H, LIU Z, et al. GSC-YOLOv5: an algorithm based on improved attention mechanism for road creak detection[C]//2023 IEEE 12th Data Driven Control and Learning Systems Conference (DDCLS). May 12-14, 2023, Xiangtan, China: IEEE, 2023: 1664-1671.