

DOI:10.11918/202112015

轻量化模型的 PeleeNet_yolov3 地表裂缝识别

刘 星,莫思特,张 江,王炜康,杨世基,李 鑫

(四川大学 电气工程学院,成都 610065)

摘 要:为提高地表裂缝检测在低算力运算平台上的稳定性和检测速率,提出了一种 PeleeNet 与 YOLOv3 相结合的目标检测算法。使用 PeleeNet 框架代替 YOLOv3 的 Darknet-53 主体框架,以融合不同的局部特征及提高运算效率;在框架中融合特征注意力模块以提高图像中裂缝区域的显著度,并通过对感受野模块 RFB 卷积核的复用,增大网络的有效视野,提高小目标检测精度;在特征金字塔网络中,通过使用深度可分离卷积代替标准卷积,减少参数计算量;引入 CIoU 损失函数提高模型的分类与回归精度。在测试平台上应用裂缝数据进行算法验证,结果表明:AP₅₀达到了 97.68%,AP₇₅达到了 77.87%,较原始的 YOLOv3 分别提高 8.4%和 12.4%,检测速度达到了 30 帧/s,且模型参数大小仅为原始 YOLOv3 的 30%;可以看出,本研究提出的 PeleeNet_yolov3 轻量化模型对于裂缝目标的检测效果较为明显,并且具有较小的运算量和参数量,适合应用于移动端系统,对于小体积低功耗低算力运算平台具有较大应用价值。

关键词: YOLOv3;裂缝检测;PeleeNet;深度可分离卷积;感受野模块;特征注意力模块;CIoU

中图分类号: TP391.4 **文献标志码:** A **文章编号:** 0367-6234(2023)04-0081-09

PeleeNet_yolov3 surface crack identification with lightweight model

LIU Xing, MO Site, ZHANG Jiang, WANG Weikang, YANG Shiji, LI Xin

(College of Electrical Engineering, Sichuan University, Chengdu 610065, China)

Abstract: To enhance the stability and speed of surface crack detection, a detection algorithm combining PeleeNet and YOLOv3 is proposed. PeleeNet framework is used to replace Darknet-53 framework of YOLOv3 so as to effectively integrate different local features and improve the detection rate. The feature attention module is integrated into the PeleeNet's framework to highlight the saliency of the crack detection in the image, and through the receptive field module RFB broaden the effective field of view of the network, and increase the detection accuracy of small targets. Instead of the standard convolution, the depth separable convolution is employed to reduce the amount of parameter calculation in the feature pyramid network. Then, the CIoU loss function is introduced to strengthen the classification and regression accuracy of the model. The experimental results on the fracture data set show that AP₅₀ and AP₇₅ reach 97.68% and 77.87% respectively, 8.4% and 12.4% higher than the original YOLOv3. Meanwhile, the detection speed reaches 30 frames per second with the size of model parameters being only 30% of YOLOv3. As can be seen, the PeleeNet_yolov3 lightweight model proposed has produced an obvious effect on the detection of crack targets, with a small amount of calculations and parameters involved. Suitable for mobile terminal system, the study presents a great value of application especially for small volume, low power consumption and low computing power computing platforms.

Keywords: YOLOv3; crack detection; PeleeNet; depth separable convolution; receptive field module; feature attention module; CIoU

地表裂缝,也称地裂缝,主要是由于地表岩石层、土体在自然因素(地震、山体滑坡、地面塌陷等)或人为因素(抽水、灌溉、开挖等)作用下产生开裂,并在地面形成一定长度和宽度裂缝的一种宏观地表破坏现象^[1]。如果发生在有人类活动的地区时,便可成为地质灾害。地表裂缝是全世界普遍存在的潜

在危害,快速检测出地表裂缝,可以有效减少地质灾害带来的损失,并起到预测作用。

传统的地裂缝识别方法是人工调查,劳动成本高,很多地方不易到达。随着科技技术的发展,出现了 GPS^[2-3]或 GNSS^[4]、InSAR^[2-3,5]、三维激光扫描^[6]、卫星遥感^[7]等先进检测技术。这些技术可以

进行大面积、精准的扫描,易获密集的、海量的观测数据,工作人员可以全面直观的知道地面裂缝分布情况,但是上述仪器设备价格昂贵,检测成本相对较高,不适合地表裂缝常规排查。随着深度学习的快速发展,特别是卷积神经网络(Convolutional neural networks, CNN)在计算机视觉任务中取得的巨大成功,基于图像的目标辨识得到越来越广泛的应用。卷积神经网络由计算机自动进行视觉特征挖掘,在语义抽象能力上有大幅提升。文献^[8-9,11]分别提出了基于 CNN 的裂缝检测算法。Peng 等^[8]在 CNN 的基础上提出了 DCNN-Inception 网络,对于裂缝分类正确率 98.2%,但是模型训练时间较长。Mao 等^[9]提出 ME-Faster R-CNN,对 Faster R-CNN (Fast region-based CNN)^[10]进行改进,提高了模型的准确率,但是检测速度较慢。Cha 等^[11]利用 CNN 搭建了一个 5 层的网络,后面跟着全连接层对裂缝进行分类,该模型对图像光照要求较高。

目前,目标检测算法按照是否提前生成候选区域,可分为单阶段目标检测算法和双阶段目标检测算法两类。双阶段目标检测算法主要是以 Fast R-CNN 为主,由于需要提前生成候选区域,尽管准确度较高,但检测速度慢。单阶段目标检测算法主要是以 YOLO(you look only once)系列和 SSD(single shot multi box detector)^[12]为主,两者都是基于网格与锚点实现,不需要二次回归^[13],相对于双阶段目标检测算法,检测速度较快,准确率略低。YOLOv3^[14]是 YOLO 系列较优的目标检测算法,框架采用 Darknet-53,输出采用特征金字塔(Feature pyramid networks, FPN)的方式,进行了深层语义特征与浅层语义特征的多尺度融合^[15],能够同时检测大目标、中目标、小目标,大大提高了目标检测的正确率,特别是小目标。地表裂缝目标小,且图形复杂,因此可以选用 YOLOv3 作为地表裂缝识别的检测算法。

由于 YOLOv3 模型参数计算量大,模型复杂,对于硬件要求高,因此不适合用于野外作业的小体积低功耗算力移动端运算。本文提出的轻量化模型,采用移动端网络 PeleeNet 网络框架^[16]代替 Darknet-53,并结合特征注意力模块(Convolution block attention module, CBAM)^[17]、改进的感受野模块(Receptive field block, RFB)^[18]、深度可分离卷积(Depthwise separable convolutions, DSC)^[19]和 CIoU(Complete intersection over union)^[20]损失,提高了目标检测的正确率,减少了运算计算量和参数量,提高了模型检测速率。

1 PeleeNet_yolov3 裂缝检测

1.1 YOLOv3 模型

YOLOv3 是一种端到端的物体检测模型,直接对锚框进行回归和分类,检测速率有大幅提升。YOLOv3 的网络结构包含特征提取网络和检测网络。特征提取网络采用的框架是 Darknet-53,包含了 5 个残差网络(Residual network, ResNet)^[21],全部由 $3 \times 3, 1 \times 1$ 的卷积核组成。每个卷积核都是由卷积、批量归一化(Batch normalization, BN)、激活函数(Leaky ReLU, LR)组成。Darknet-53 将 3 个尺度特征图输出到检测网络的 3 个分支,形成 FPN。最后对 FPN 的输出进行边界框和类别预测。

1.2 PeleeNet_yolov3 网络

PeleeNet 是一种基于 DenseNet^[22]轻量化的变体,主要面向的是移动端应用。现有研究表明对于 COCO 数据集,PeleeNet 不仅比 SSD + MobileNet 更准确,也比 YOLOv2 更准确,算法速度相比 YOLOv3 提升 5.6 倍^[16],因此本文选择 PeleeNet 代替 Darknet-53 框架。PeleeNet 在模型的开始,设计了一种高效率的 Stem 模块,Stem 结构如图 1(a)所示。密集单元 Dense Layer 的设计采用了两种尺度的卷积核同时进行特征挖掘,可以以最小的计算资源有效的提取特征, Dense Layer 模块如图 1(b)所示。

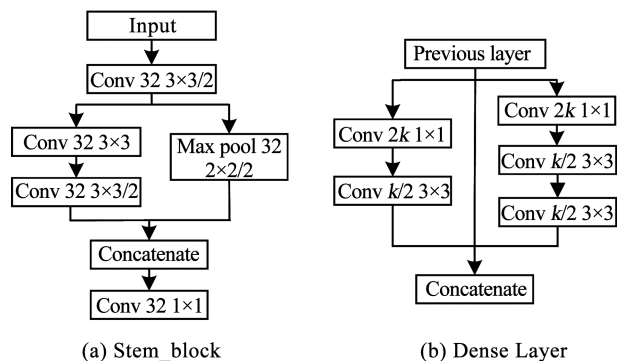


图 1 Stem_block 和 Dense_layer 结构图

Fig. 1 Structure of Stem_block and Dense_layer

针对原始 YOLOv3 的模型参数量、计算量大,模型复杂,对于硬件要求高等问题,本文结合 YOLOv3 和 PeleeNet 两种框架的特点,提出了轻量级的实时目标检测网络,本文将其称为 PeleeNet_yolov3,整体结构如图 2 所示。在 PeleeNet_yolov3 特征提取网络中,使用 PeleeNet 的密集连接单元代替 Darknet-53 的残差单元,在密集单元内,增长率设为 32,即每个密集单元都会被压缩到 32 个特征图,然后将 32 个新的特征图加入到全局特征。随着网络层数的加

深、特征图的尺寸减小,需要更多的特征图来保持图像的特征,因此在整个网络中,采用了 4 级骨干网络,分别是 3、4、8、16,过渡层 (Transition_layer) 通过 1×1 的卷积核进行特征融合,平均池化层进行下采样。PeleeNet 通过密集单元,可以同时进行小目标和大目标的语义特征挖掘,为了进一步的提高模型准确率而不增加较多计算量的前提下,在第 1 个密集单元和检测网络的 3 个输入端添加 CBAM 模块,让模型在训练过程中更加关注特征图信息丰富的通

道和图像区域,加快训练速度,提高模型的准确率。在特征提取网络输出特征图的大小为 13×13 的支路加入针对裂缝特征的 RFB 模块,模拟人类视觉机制,多尺度挖掘特征图的语义特征,让丰富的语义特征进行上采样,弥补浅层网络高级语义特征的不足,以完成裂缝目标识别的任务需求。为满足模型的检测速度而不过多的损失模型的准确率,在模型的检测网络中加入 DSC。下面对网络的各个模块做出简要介绍。

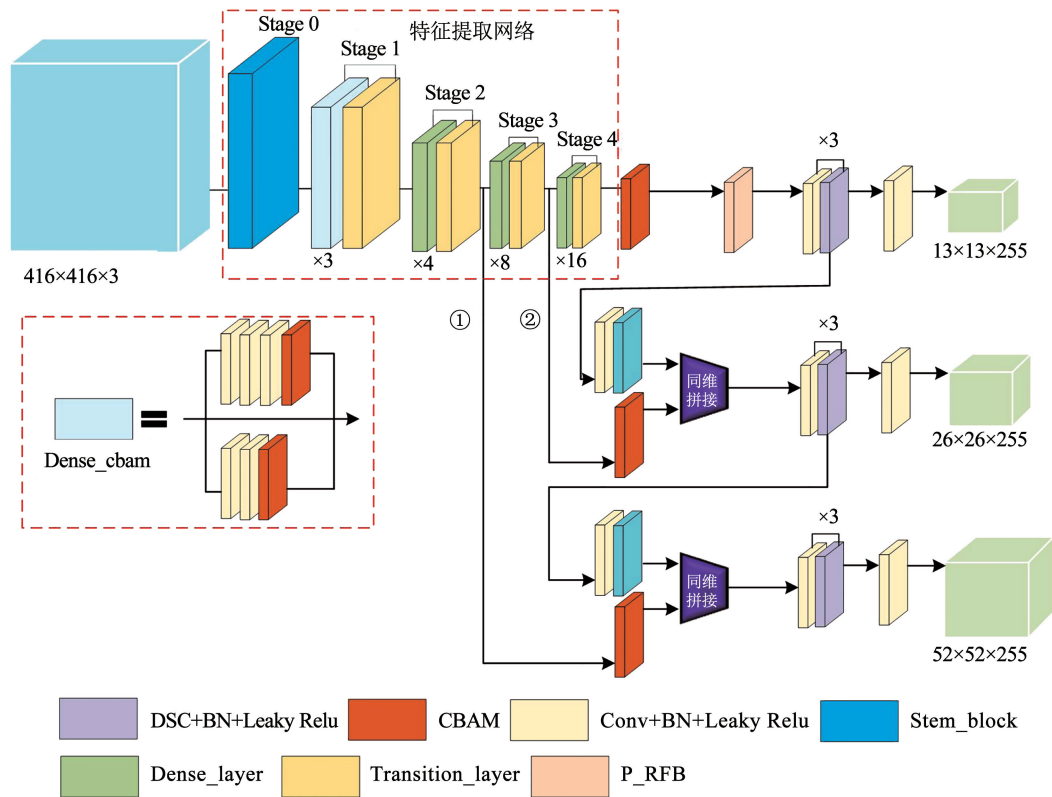


图 2 PeleeNet_yolov3 结构图
Fig. 2 Structure of PeleeNet_yolov3

1.2.1 特征注意力模块-CBAM

CBAM 是一个有效的前馈神经网络特征注意力模块,包含通道注意力模块 CAM (channel attention

module) 和空间注意力模块 SAM (spatial attention module)。CAM 关注通道方向上的信息权重, SAM 关注空间方向上的信息权重。CBAM 如图 3 所示。

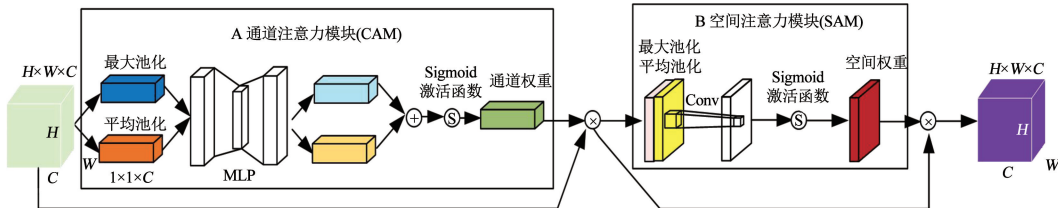


图 3 CBAM 结构图
Fig. 3 Structure of CBAM

CAM 如图 3 中 A 部分所示。输入 $H \times W \times C$ 的特征图,经过最大池化层 (Max pooling, MaxPool) 和平均池化层 (Average pooling, AvgPool) 对特征图进行压缩,然后会形成两个 $1 \times 1 \times C$ 的通道描述,将

两个通道描述经过多层感知机 MLP (Multi-layer perception) 两次全连接后相加,经过 sigmoid 函数转换为权重系数,与原来的特征相乘即可得权重占比不同的特征图。SAM 如图 3 中 B 部分所示。SAM

沿着深度 C 方向进行最大池化和平均池化,得到两个不同的背景描述,对两个描述进行拼接,使用 7×7 卷积生成一个特征注意力卷积图,经过 sigmoid 激活函数,得到空间维度上的权重系数,将该生成的权重系数与输入相乘可得到空间区域占比不同的特征图。

1.2.2 深度可分离模块

为了进一步减少模型的计算量,在模型的检测网络中,使用 MobileNet 的 DSC,如图 4 所示,将模型的标准卷积分为深度卷积 (Depthwise convolution, DC) 和逐点卷积 (Pointwise convolution, PC),以减少模型的数量和大小。DSC 首先进行 DC,DC 对于每个通道分别独立使用一个深度为 1 的卷积核进行特征提取,解决模型参数冗余的问题。经过 DC 的各通道特征之间使用 1×1 的卷积核进行相互融合。现有研究表明,虽然加入 DSC 正确率略有降低,但模型参数大大减少且检测速度得到较大提升。例如,对于一个深度为 C 的特征图,若用 $H_c \times W_c$ 的标准卷积去得到一个深度为 C_1 的特征图,需要的计算量为

$$a = H_c \times W_c \times C \times C_1 \quad (1)$$

如果使 DSC 完成同样的工作,需要的计算量为

$$b = H_c \times W_c \times C + C \times C_1 \quad (2)$$

最后可以推出 DSC 与标准卷积之间的计算量比值为

$$\frac{b}{a} = \frac{H_c \times W_c \times C + C \times C_1}{H_c \times W_c \times C \times C_1} = \frac{1}{C_1} + \frac{1}{H_c \times W_c} \quad (3)$$

由此可见,利用 DSC,计算量和参数相比标准卷积会大幅度下降,这样的改进能够显著地提高检测速度。

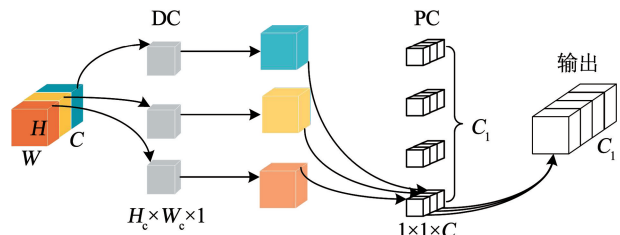


图 4 DSC 结构图

Fig. 4 Structure of DSC

1.2.3 针对裂缝特征 RFB

卷积核的多路复用不会改变 RFB 对不同目标尺度的感知,为了进一步减少模型的参数和运算量,针对裂缝特征,本文对 RFB 进行了改进。将 RFB 分支 2 的 3×3 卷积核,分支 3 的 5×5 卷积核利用 1×3 , 3×1 的卷积核代替,大膨胀率的非对称卷积^[23]适用于多尺度的裂缝形态分布,可以捕获其非局部特征,并让 1×1 的卷积核进行多路复用,由于

P_RFB 输入通道数和输出通道数一致,故文献[24]可以选择直接相连,最终可训练参数相对于 RFB 减少了 38%,同时使模型有更好的非线性表达能力,且从信息传输的角度来看,P_RFB 的多分支结构已经通过不同的视觉大小获取信息,故文献[24]选择直接相连,让信息的传输变得更加流畅,增加了 P_RFB 模块中信息的多样性,提高了模型在浅层中提取高级语义特征的能力,显著提高裂缝检测的性能。改进后的 P_RFB 如图 5 所示。

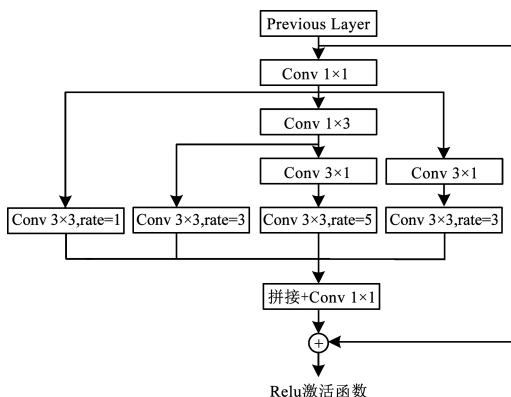


图 5 P_RFB 结构图

Fig. 5 Structure of P_RFB

1.2.4 损失函数的改进

YOLOv3 的损失函数使用的均方差 (MSE) 去判断候选框和真实框之间的匹配程度,但是均方差对于预测框位置的敏感性不够,候选框在不同位置时,其损失值可能相同,导致损失函数不能收敛到最小。本文选用 CIoU 来代替均方差损失。CIoU 是在 DIoU (Distance IoU loss) 上改进而来的,DIoU 损失函数考虑了预测框与真实框之间的重叠比、中心点距离等因素的影响,使候选框时刻保持梯度回传,更好的优化了边界框的位置信息。DIoU 的计算公式为

$$L_{DIoU} = 1 - \frac{|A \cap B|}{|A \cup B|} + \frac{\rho^2(b, b^{gt})}{c^2} \quad (4)$$

式中: c 为能够同时包围候选框 A 和真实框 B 的最小框的对角线距离, d 为两个框中心点的距离, b, b^{gt} 分别为候选框与真实框的中心点, ρ^2 为求其距离,如图 6 所示。

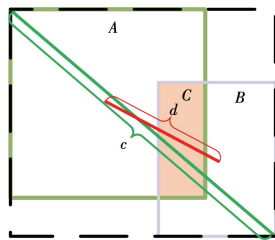


图 6 DIoU 示意图

Fig. 6 Schematic diagram of DIoU

CIoU 相对 DIoU 引入了一个对于候选框与真实框长宽比的惩罚项,CIoU 的计算公式如式(5)所示。裂缝属于小目标,为了深入加强对于小目标的检测,在位置回归损失部分,加上一个惩罚项 $(2-w \times h)$,其中 w 、 h 分别为真实框的宽、高,当真实框较小,其惩罚项系数较大,加强对于小目标的检测,提高模型的准确率。

$$L_{CIoU} = L_{DIoU} + \alpha v \tag{5}$$

式中: v 为衡量长宽比一致性的参数, α 为正权衡参数。 v 、 α 的计算公式如下:

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \tag{6}$$

$$\alpha = \frac{v}{(1 - \left| \frac{A \cap B}{A \cup B} \right|) + v} \tag{7}$$

式中: w 、 h 分别为候选框的宽、高, w^{gt} 、 h^{gt} 分别为真实框的宽、高。最终模型候选边界框与真实边界框的位置部分损失函数为

$$L_{coord} = \sum_{i=0}^{s^2} \sum_{j=0}^k (2 - w_i \times h_i) \times L_{CIoU} \tag{8}$$

式中: w_i 、 h_i 分别为第 i 个真实框的宽、高, s 为划分网格的数目,输出特征图的网格大小为 $s \times s$, k 为候选框的数量。

2 实 验

2.1 配置与训练

实验室平台配置如下: Intel Core i5-9600F @ 2.40 GHz × 4 CPUs, NVIDIA GeForce GTX 960。主要工作内容包括数据集的制作、代码编写等。模型训练采用的是谷歌云平台: Colab, 分配的 GPU: Tesla T4, 深度学习框架: tensorflow 1.5, python 版本为 3.7。

训练时,一个批次包含 16 幅 416 × 416 的图像,动量参量为 0.99,采用随机梯度下降(SGD),设置的最大 epoch 为 140,最大的迭代次数为 19 458,衰减系数为 0.005。为防止梯度爆炸,使用学习率预热的方式,在 276 次迭代后将学习率提升到初始学习率 0.001 大小,预热时学习率较小模型慢慢趋于稳定,加快模型的训练速度。随后的训练学习率采用余弦衰减(cosine_decay)的衰减方式,余弦衰减根据迭代的次数进行学习率的衰减,如图 7(a)所示。在迭代的初期,训练的学习率与学习步长较大,模型的收敛速度较快。在迭代的后期,训练的学习率接近设置的最小学习率,有助于模型的收敛,逐步逼近最优。改进后的模型损失函数下降如图 7(b)所示。由图 7 可知,模型最终在迭代 17 500 次以后,损失值趋于稳定,最终收敛在 0.45 左右。

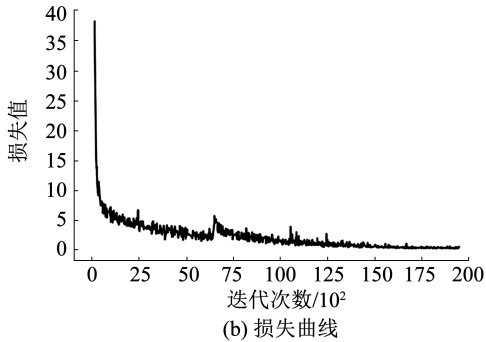
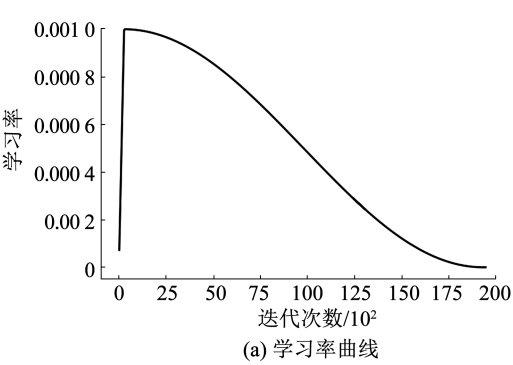


图 7 训练过程可视化
Fig. 7 Visualization of training process

2.2 数据

实验数据来源于无人机在中国易于发生地质灾害的某地进行飞行拍照,共采集到 200 张分辨率为 5 472 × 3 078 的图片,并将图片中含裂缝的区域裁剪成 256 × 256 的大小,并且通过水平、垂直、90°和 180°随机翻转增加数据集的数量,同时加入部分无缺陷的地面图像,提高数据的鲁棒性。数据集的格式采用 xml 对地表裂缝进行标记和定位。图片最终数量为 2 460 张,按照比例为 7:3 划分为训练集和测试集。

地表裂缝数据集匮乏,且具有连续性、延长性的特点,导致标签的制作没有规则,会造成模型预测困难,易导致模型损失函数震荡,且部分地表裂缝的背景颜色复杂,造成模型训练困难。本文,先将图片剪切成 256 × 256 的大小,基于裂缝走向变换和出现裂缝分支,重新绘制标签框并让标签框包含较少的背景色占比,从而可以减少模型训练的困难和模型预测结果的重复性。标签制作如图 8 所示。

2.3 结果及分析

为了验证本文改进算法的有效性,本文在制作的数据集上进行了各类方法的消融实验。各类方法的消融实验结果见表 1。其中 AP (average precision) 为一类的 MAP (mean average precision); AR (average recall) 为平均召回率; AR₅₀、AP₅₀ 分别为当 IOU = 0.50 时的平均检测精度和平均召回率; AR₇₅、AP₇₅ 分别为当 IOU = 0.75 时的平均检测精度

和平均召回率;GFLOPs 为模型的浮点计算量;模型大小给出了模型的参数量;F1(50) 为当 IOU = 0.50 时的调和均值。



图 8 标签制作
Fig. 8 Label making

2.3.1 改进后的模型消融实验

由表 1 可知,编号 2 使用 PeleeNet 作为主体框架,与主体框架为 Darknet-53 的编号 1 对比可知,

模型的平均精度、平均召回率、F1 均有下降,但模型的参数只有编号 1 的 40%,检测速度提高了 64%,达到了 41 帧/s;编号 3 在编号 2 的基础上,将模型的边界框损失函数换为 CIoU, AP₅₀、AP₇₅ 和召回率方面均有提高,说明引入 CIoU 损失对于裂缝检测的各项指标有较为明显的提高,有效解决了训练过程中对于锚框的不敏感性;编号 4 将检测网络中 3 × 3 的标准卷积换为 DSC,模型的参数量进一步减少,各项指标与编号 1 基本能保持,检测速度达到了 46 帧/s;编号 5 是在模型 4 的基础上,添加了改进的 P_RFB,各项评价指标均得到了提升,表明 P_RFB 有效融合了不同尺度的语义信息,提高检测网络的深层语义信息挖掘能力,但也导致模型的参数量小幅增加,使得检测速度降低了 3 帧/s;编号 6 在编号 5 的基础上添加了 CBAM 模块,与之前编号为 1、2、3、4、5 模型对比可知,改进效果提升显著,与编号 1 相比,参数量减少 69.78%,检测速度提高 5 帧/s, AP₅₀ 提升 6.76%, AP₇₅ 提升 9.67%,验证了加入 CBAM 的有效性。综合表 1 中各组实验对比数据,可以验证本文算法的各项改进能有效提高对于地面裂缝的检测。改进网络的部分检测结果如图 9 所示。

表 1 在裂缝数据集上消融实验结果

Tab. 1 Results of ablation experiments on the fracture data set									
方法编号	模型	AP ₅₀ /%	AR ₅₀ /%	AP ₇₅ /%	AR ₇₅ /%	F1(50)	模型大小/MB	浮点运算量/GFLOPs	检测速度/(帧·s ⁻¹)
1	Darknet-53	90.92	85.37	68.20	61.32	0.88	234.0	65.70	25
2	1 + PeleeNet	86.40	84.06	57.62	56.76	0.85	95.0	20.54	41
3	2 + CIoU	94.54	88.72	69.41	65.14	0.92	95.0	20.54	41
4	3 + DSC	94.90	84.53	69.84	62.19	0.90	53.0	12.00	46
5	4 + P_RFB	96.52	90.40	70.75	66.26	0.94	69.0	13.66	43
6	本文	97.68	94.22	77.87	75.12	0.96	70.7	13.68	30

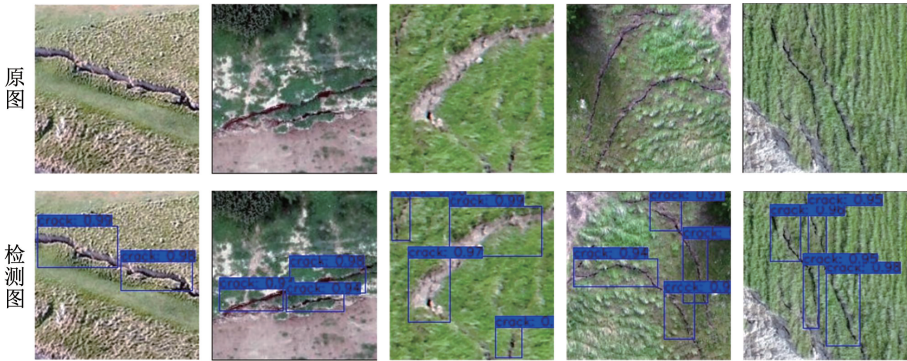


图 9 部分裂缝的检测结果
Fig. 9 Detection results of some cracks

2.3.2 CBAM 位置的验证

为了验证特征注意力模块 CBAM 位置对于网

络框架的影响,在上述编号 5 的基础上,本文对于 CBAM 位置进行了以下研究。CBAM 在特征提取网

络位置见表 2。表中√为加入 CBAM,√* 为检测网络输入端加入 CBAM,×为不加入 CBAM。

表 2 CBAM 不同位置
Tab.2 Different locations of CBAM

实验序号	Stage 1		Stage 2		Stage 3		Stage 4	
	D×3	T	D×4	T	D×8	T	D×16	T
1	×	√	×	√	×	√	×	√
2	×	×	×	√*	×	√*	×	√
3	√	×	√	×	√	×	√	×
4	√	√	√	√	√	√	√	√
5	√	√	×	√	×	√	×	√
6(本文)	√	×	×	√*	×	√*	×	√

注:D 为 Dense_layer, T 为 Transition_layer

编号 1 将 CBAM 加在模型特征提取网络过渡层最大池化层后。编号 2 将 CBAM 位置加在过渡层在模型检测网络的 3 个输入端,如图 2 中①,②两点所示。编号 3 是将 CBAM 位置插入到密集层,如图 2 的 Dense_cbam 所示。编号 4 是在编号 1,3 的基础上进行组合。由表 3 可知,编号 1、2、3、4 在加入特征注意力 CBAM 后,模型的平均召回率、检测

精度都有提升,编号 4 地面裂缝检测的各项指标提升最大,但是模型的检测速度比 YOLOv3 低了 40%,说明在密集层加入 CBAM 对于地面裂缝的检测效果有大幅提升,模型的浮点计算量 GFLOPs 和参数量略有增加,但是网络层数大幅增加,使得检测速度下降了约 20 帧/s;编号 1 和编号 2 的各项指标检测效果近似,为了维持模型的检测速度,提高平均召回率、平均精确度和 F1,在编号 1、2 模型的第 1 个密集块插入 CBAM,通过对比可知,编号 6 比编号 5 的 AP₅₀ 提升了 1.51%,AR₅₀ 提升 2.00%,AR₇₅ 和 AP₇₅ 相对编号 1、2、3、5 提升最大且能保证模型的检测速度。本文最终选择在第 1 个密集块和检测网络的 3 个输入端插入 CBAM。上述实验表明,加入 CBAM 的位置,对于模型浮点计算量 GFLOPs 和参数数量影响较小,对于检测速度具有较大影响。CBAM 不仅能给网络性能带来实质性的提升,而且增加的计算量和参数量很少,全连接层和卷积层带来的计算开销可以忽略不计。

表 3 在裂缝数据集 CBAM 不同位置实验结果
Tab.3 Experimental results at different positions of CBAM in the fracture data set

方法编号	模型	AP ₅₀ /%	AR ₅₀ /%	AP ₇₅ /%	AR ₇₅ /%	F1(50)	模型大小/MB	浮点运算量/GFLOPs	检测速度/(帧·s ⁻¹)
1	1	96.86	91.89	74.17	70.36	0.94	70.7	13.40	35
2	2	96.82	90.77	74.45	69.80	0.94	70.7	13.70	36
3	3	98.92	93.94	72.82	69.15	0.96	69.4	13.42	15
4	4	99.16	99.35	87.07	87.23	0.99	70.8	13.70	14
5	5	96.11	92.17	71.91	68.97	0.94	70.7	13.65	29
6	本文	97.68	94.22	77.87	75.12	0.96	70.7	13.68	30

2.3.3 不同网络的对比实验

为了验证本网络在地表裂缝识别上的优异性,选择经典的模型进行对比,分别选择了 SSD,Faster-RCNN 以及 YOLOv3 及其改进算法,包括 YOLOv3-spp,YOLOv3-IouLoss。实验配置采用相同的配置,在裂缝数据集上进行了对比。结果见表 4。

YOLOv3,SSD 是单目标检测的代表作,这里

SSD 选择的 SSD300,选择 VGG16 作为骨干网络,在 6 个层次特征图上进行多尺度检测,直接对 8 732 个不同尺度的默认框对应的目标位置和类别信息进行预测。Faster-RCNN 是二阶段目标检测的代表性著作,使用区域性生成网络(Region proposal network,RPN)代替了选择性搜索算法,然后在目标区域上进行边框回归和分类,得到目标位置和类别信息。

表 4 在裂缝数据集上不同模型对比实验结果
Tab.4 Comparison of experimental results of different models on the fracture data set

方法编号	模型	AP ₅₀ /%	AR ₅₀ /%	AP ₇₅ /%	AR ₇₅ /%	F1(50)	模型大小/MB	浮点运算量/GFLOPs
1	YOLOv3	90.92	85.37	68.20	61.32	0.90	234.0	65.70
2	YOLOv3-spp	90.80	86.67	68.30	62.44	0.91	288.0	70.56
3	YOLOv3-iou	91.89	87.42	66.90	62.26	0.91	234.0	65.70
4	Faster-RCNN	68.89	87.51	69.71	40.70	0.77	108.0	63.00
5	SSD300(vgg16)	87.38	76.14	52.20	45.50	0.81	90.7	57.00
6	本文	97.68	94.22	77.87	75.12	0.96	70.7	13.68

由表 4 可知,本文提出的平均召回率、平均精度均高于 SSD300、Faster-RCNN、YOLOv3 及其改进算法。本文提供的模型旨在用于无人机地表裂缝检测,模型参数量小、浮点运算量小、准确度高,有益于用于小体积低功耗移动端的地表裂缝辨识。

2.3.4 地表裂缝移动端部署

为了验证本网络在移动端对于地表裂缝的识别效果,本文选择 Google 的 TensorFlow Lite 框架进行模型部署。TensorFlow Lite 是一种在设备端运行 TensorFlow 模型的开源深度学习框架,是一个轻量、快速、跨平台的专门针对移动和 IoT 应用场景的机器学习框架,是开源机器学习平台 TensorFlow 的重

要组成部分^[25]。开发者可以在此基础上进行移动端部署,开发出 APP,将算法在移动端运行。

首先将在 TensorFlow 环境下训练的本文网络模型的 PeleeNet_yolov3.ckpt 文件转换为 tflite 数据格式的 PeleeNet_yolov3.tflite 文件。然后搭建 Android 地表裂缝识别项目工程,将模型载入 VIVO S6 手机。地表裂缝识别算法在 VIVO S6 手机测试结果如图 10 所示。由图 10 可知,本文算法在移动端可以正确识别到地表裂缝,且一张地表裂缝图在 VIVO S6 的 CPU、GPU 上检测时间约为 145、65 ms,其检测速度约为 7、15 帧/s,可以达到移动端实时检测识别的效果。

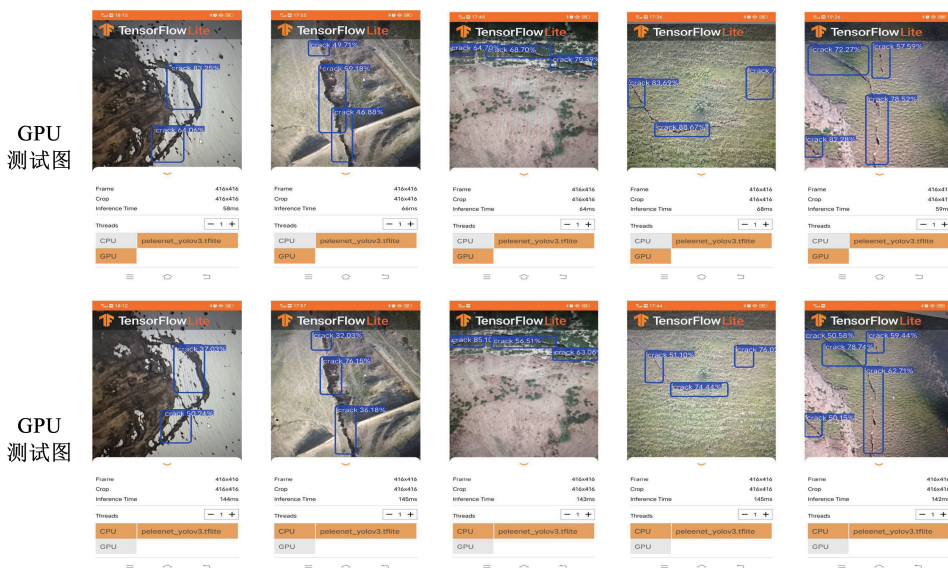


图 10 移动端地表裂缝检测结果

Fig. 10 Mobile terminal surface crack detection results

3 结 论

1)经典的 YOLOv3 网络对地表裂缝缺陷检测的准确度较高,但是模型参数量大,硬件资源要求高,不适合移动端的部署,难以满足实际需求。

2)采用轻量级网络 PeleeNet 作为特征提取网络,结合特征注意力模块 CBAM 和改进的 P_RFB,加入深度可分离卷积进一步对网络结构进行优化调整,并通过 CIoU 对损失函数进行改进,最终模型的准确率 AP_{50} 达到 97.68%, AP_{75} 达到 77.87%,可以解决裂缝目标检测问题。

3)实验表明,PeleeNet_yolov3 在地表裂缝数据集上取得了较优的结果,检测精度和检测速度均优于 YOLOv3 与部分已有的公共目标检测算法,且本文模型在 VIVO S6 移动端可识别图像检测速度约为 15 帧/s,说明本文提出算法的准确率高,模型参数量和计算量小,适用于小体积低功耗低算力运算平台。

参考文献

- [1] 卢全中,李聪,刘聪,等. 地裂缝分类及地面沉降区构造地裂缝防治对策[J]. 地球科学与环境学报, 2021, 43(2): 366
LU Quanzhong, LI Cong, LIU Cong, et al. Classification of ground fissures and Prevention Measures of Tectonic Ground fissures in Land subsidence area[J]. Journal of Earth Sciences and Environment, 2021, 43(2): 366. DOI: 10.19814/j.jese.2020.11050
- [2] 张勤,赵超英,丁晓利,等. 利用 GPS 与 InSAR 研究西安现今地面沉降与地裂缝时空演化特征[J]. 地球物理学报, 2009, 52(5): 1214
ZHANG Qin, ZHAO Chaoying, DING Xiaoli, et al. Research on recent characteristics of spatio-temporal evolution and mechanism of Xi'an land subsidence and ground fissure by using GPS and InSAR techniques[J]. Chinese Journal of Geophysics, 2009, 52(5): 1214. DOI: 10.3969/j.issn.0001-5733.2009.05.010
- [3] FERGASON K C, RUCKER M L, PANDA B B. Methods for monitoring land subsidence and earth fissures in the Western USA[J]. Proceedings of the International Association of Hydrological Sciences, 2015, 372(12): 361. DOI: 10.5194/piahs-372-361-2015
- [4] 许强,董秀军,李为乐. 基于天-空-地一体化的重大地质灾害

害隐患早期识别与监测预警[J]. 武汉大学学报·信息科学版, 2019, 44(7): 957

XU Qiang, DONG Xiujun, LI Weile. Integrated space-air-ground early detection, monitoring and warning system for potential catastrophic geohazards[J]. Geomatics and Information Science of Wuhan University, 2019, 44 (7): 957. DOI: 10. 13203/j. whugis20190088

[5] QIU Zhiwei, JIANG Tingchen, ZHOU Li, et al. Study of subsidence monitoring in Nanjing City with small-baseline InSAR approach[J]. Geomatics Natural Hazards & Risk, 2019, 10(1): 1412. DOI: 10. 1080/19475705. 2019. 1572660

[6] 曾庆鲁, 张荣虎, 卢文忠, 等. 基于三维激光扫描技术的裂缝发育规律和控制因素研究——以塔里木盆地库车前陆区索罕村露头剖面为例[J]. 天然气地球科学, 2017, 28(3): 397

ZENG Qinglu, ZHANG Ronghu, LU Wenzhong, et al. Fracture development characteristics and controlling factors based on 3D laser scanning technology: An outcrop case study of Suohan village, Kuqa foreland area, Tarim Basin [J]. Natural Gas Geoscience, 2017, 28(3): 397. DOI: 10. 11764/ji. ssn. 1672G1926. 2017. 02. 004

[7] 侯恩科, 张杰, 谢晓深, 等. 无人机遥感与卫星遥感在采煤地表裂缝识别中的对比[J]. 地质通报, 2019, 38(2/3): 443

HOU Enke, ZHANG Jie, XIE Xiaoshen, et al. Contrast application of unmanned aerial vehicle remote sensing and satellite remote sensing technology relating to ground surface cracks recognition in coal mining area[J]. Geological Bulletin of China, 2019, 38(2/3): 443. DOI:10. 12097/j. issn. 1671 - 2552. 2019. 2 - 3. 025

[8] PENG Guili, TUO Xianguo, SHEN Tong, et al. Recognition of rock micro-fracture signal based on deep convolution neural network inception algorithm[J]. IEEE Access, 2021, 9(99): 89390. DOI: 10. 1109/ACCESS. 2021. 3086630

[9] MAO Yingchi, PING Ping, CHEN Jing, et al. Crack detection with multi-task enhanced faster R-CNN model [C]//Proceedings of the 6th International Conference on Big Data Computing Service and Applications (BigDataService). Oxford, UK: IEEE, 2020: 193. DOI: 10. 1109/BigDataService49289. 2020. 00044

[10] REN Shaoqing, HE Kaiming, 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. DOI: 10. 1109/TPAMI. 2016. 2577031

[11] CHA Y J, CHOI W, BÜYÜKÖZTÜRK O. Deep learning-based crack damage detection using convolutional neural networks [J]. Computer-Aided Civil and Infrastructure Engineering, 2017, 32(5): 361. DOI: 10. 1111/mice. 12263

[12] LUI Wei, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]//European Conference on Computer Vision. Cham: Springer, 2016: 21. DOI:10. 1007/978 - 3 - 319 - 46448 - 0_2

[13] 姜维, 张重生, 殷绪成. 基于深度学习的场景文字检测综述 [J]. 电子学报, 2019, 47(5): 1152

JIANG Wei, ZHANG Chongsheng, YIN Xucheng. Deeplearning based scene text detection: A survey[J]. Acta Electronica Sinica, 2019, 47(5): 1152. DOI: 10. 3969/j. issn. 0372 - 2112. 2019. 05. 024

[14] REDMON J, FARHADI A. YOLOv3: An incremental improvement [C]. Computer Vision and Pattern Recognition. arXiv Preprint, 2018, arXiv:1804. 02767. DOI: 10. 48550/arXiv. 1804. 02767

[15] 罗会兰, 陈鸿坤. 基于深度学习的目标检测研究综述[J]. 电子学报, 2020, 48(6): 1230

LUO Huilan, CHEN Hongkun. Survey of object detection based on deep learning[J]. Acta Electronica Sinica, 2020, 48(6): 1230. DOI: 10. 3969/j. issn. 0372 - 2112. 2020. 06. 026

[16] WANG R J, LI Xiang, LING C X. Pelee: A real-time object detection system on mobile devices[C]//Proceedings of the 32nd International Conference on Neural Information Processing Systems. Montréal Canada: Curran Associates Inc. , 2018: 1967

[17] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision. Munich, Germany: Springer Verlag, 2018: 3. DOI: 10. 1007/978 - 3 - 030 - 01234 - 2_1

[18] LIU Songtao, HUANG Di, WANG Yunhong. Receptive field block net for accurate and fast object detection [C]//Proceedings of the European Conference on Computer Vision. Cham: Springer International Publishing, 2018: 385. DOI: 10. 48550/arXiv. 1711. 07767

[19] HOWARD A G, ZHU Menglong, CHEN Bo, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Hawaii: IEEE, 2017: 1704. DOI: 10. 48550/arXiv. 1704. 04861

[20] ZHENG Zhaohui, WANG Ping, LIU Wei, et al. Distance-IoU loss: Faster and better learning for bounding box regression [C]//Proceedings of the AAAI Conference on Artificial Intelligence. [S.l.]: AAAI, 2020. DOI: 10. 48550/arXiv. 1911. 08287

[21] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV: IEEE Computer Society, 2016: 770. DOI:10. 1109/CVPR. 2016. 90

[22] HUANG Gao, LIU Zhuang, VAN Der MAATEN L, et al. Densely connected convolutional networks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI: IEEE Computer Society, 2017: 4700. DOI: 10. 1109/CVPR. 2017. 243

[23] YU F, KOLTUN V, FUNKHOUSER T. Dilated residual networks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI: IEEE, 2017: 636. DOI: 10. 1109/CVPR. 2017. 75

[24] SZEGEDY C, LIU Wei, JIA Yangqing, et al. Going deeper with convolutions[C]//Proceedings of the Computer Vision and Pattern Recognition. Boston, MA: IEEE Computer Society, 2015: 1. DOI: 10. 1109/CVPR. 2015. 7298594

[25] 李双峰. TensorFlow Lite: 端侧机器学习框架[J]. 计算机研究与发展, 2020, 57(9): 1839

LI Shuangfeng. TensorFlow lite: On-device machine learning framework[J]. Journal of Computer Research and Development, 2020, 57 (9) : 1839. DOI: 10. 7544/issn1000 - 1239. 2020. 20200291