

基于改进 YOLOX 的落石检测方法

陈 垦^{1,2}, 欧 鸥³, 杨长志⁴, 龚 帅⁴, 欧阳飞³, 向东升³

(1. 东南大学信息科学与工程学院 毫米波国家重点实验室, 南京 211189;

2. 四川数字交通科技股份有限公司, 成都 610218;

3. 成都理工大学 计算机与网络安全学院 (牛津布鲁克斯学院), 成都 610059;

4. 成理智源科技(成都)有限公司, 成都 610059)

摘要: 山坡地区是落石频发的区域, 凭人力难以及时发现灾害的发生; 为及时检测到落石的发生并做出应对措施, 提出一种基于改进 YOLOX 的落石检测方法, 自动检测并报告落石的发生情况; 通过自制落石数据集训练 YOLOX 网络, 优化空间金字塔池化结构, 获取更多语义信息, 并引入 ECA-Net (Efficient Channel Attention Module, 高效通道注意力模块), 提高特征的提取能力和特征间的信息传播, 同时改进损失函数并使用数据增强, 提高网络训练效果; 实验结果表明, 改进 YOLOX 算法的 mAP@0.5 为 92.50%, 每秒检测帧数为 62.6, 相较于 YOLOX 算法, mAP@0.5 提高 3.45%, 每秒检测帧数上涨 0.3; 与原算法相比, 在不损失性能的情况下, 精度有较大的提升, 同时满足图片与视频数据的实时检测要求。

关键词: YOLOX; 目标检测; 落石检测; 注意力机制; 空间金字塔池化

Rockfall Detection Method Based on Improved YOLOX

CHEN Ken^{1,2} OU Ou³, YANG Changzhi⁴, GONG Shuai⁴, OUYANG Fei³, XIANG Dongsheng³

(1. State Key Laboratory of Millimeter Waves, Southeast University, Nanjing 211189, China;

2. Sichuan Digital Transportation Technology Co., Ltd., Chengdu 610218, China;

3. School of computer and network security (Oxford Brooks College), Chengdu University of Technology, Chengdu 610059, China; 4. Chenglizhiyuan Technology (Chengdu) Co., Ltd., Chengdu 610059, China)

Abstract: Hillside areas are prone to falling rocks, so it is difficult to detect the occurrence of disasters in time by manpower. In order to timely detect the occurrence of falling rocks and take steps, a falling rocks detection method based on improved YOLOX is proposed to automatically detect and report the occurrence of falling rocks. The self-made rockfall data set is used to train the YOLOX network, optimize the spatial pyramid pool structure, and obtain more semantic information. The attention mechanism of efficient channel attention module (ECA-Net) channel is introduced to improve the feature extraction ability and information transmission between features. Meanwhile, the loss function is improved, and the data enhancement is used to improve the network training effect. The experimental results show that the mAP@0.5 of the improved YOLOX algorithm is 92.50%, and the detection frame rate per second is 62.6. Compared with the YOLOX algorithm, the mAP@0.5 is 3.45%, and the detection frame rate per second is increased by 0.3. Compared with the original algorithm, the accuracy is improved greatly without loss of performance, and it meets the real-time detection requirements of image and video data are met.

Keywords: YOLOX; object detection; rockfall detection; attention mechanism; space pyramid pool

0 引言

我国幅员辽阔, 地大物博, 同时也地形复杂, 有风景秀美的山川, 还有大量的丘陵地带。部分城市靠山而建, 山坡区域的地质情况不稳定, 易受烈日与雨水的侵蚀, 发生崩塌、滑坡和落石的情况^[1]。落石的危害程度不一, 轻者阻断交通, 落石灾害发生在国道与隧道洞口^[2], 使行人

和车辆不能正常通行; 重者落石灾害发生在行人经过的区域, 危及人群的生命安全, 或发生在输油管道区域, 对油气运输造成巨大影响^[3]。而落石的发生, 很可能是一些更具危害的灾害如: 崩塌、泥石流灾害的前兆。修建于山坡附近的小区 and 公园等人流量大的场所, 一旦发生落石等情况, 容易造成人员受伤。为了保证人民群众的生命和财产

收稿日期: 2023-01-05; 修回日期: 2023-02-27。

基金项目: 四川省科技厅应用基础研究项目(2021YJ0335)。

作者简介: 陈 垦(1986-), 男, 博士, 高级工程师。

通讯作者: 欧 鸥(1977-), 男, 博士, 教授。

引用格式: 陈 垦, 欧 鸥, 杨长志, 等. 基于改进 YOLOX 的落石检测方法[J]. 计算机测量与控制, 2023, 31(11): 53-59.

安全,防止落石灾害带来的巨大损失,通过实时的落石检测的方法及时提醒民众避难的重要性不言而喻。落石发生时,人们往往无法及时发现,而自动检测识别和告警的方法,可以及时通知人们避难^[4-6]。具体步骤为:通过山坡附近的摄像头对落石现象进行监测,一旦发生落石,可以及时发出警告,提醒人民及时避难,防止人员受到危害。因此,如何利用目标检测方法实时、高效并准确地检测出落石现象,具有一定的研究价值。

传统的落石检测系统,是通过帧差法,检测物体是否产生了移动,再根据移动物体的几何特征判断是否为落石^[7],其检测速度较慢、检测精度不高,比较容易出现错误预报、漏报等现象,因此本文将深度学习的目标检测算法融入落石检测系统中,提高落石检测效率、检测精度。

目标检测算法可分为 Two-Stage 和 One-Stage 模型。Two-Stage 模型包含 R-CNN、Fast R-CNN 和 Faster R-CNN 等^[8-10],整个检测流程分为两大步,即先产生候选框,再进行分类与回归,Two-Stage 模型的识别正确率较高,但实时性能较低,适用对检测速度要求不高的场景。One-Stage 模型包含 YOLO 系列^[11-13]和 SSD 系列^[13-16]等,整个检测流程只需一个步骤,经过单次检测得出目标位置以及置信度得分,其相比于 Two-Stage 模型,准确度稍低,但实时检测性能较高,更适用需要快速检测目标的场景。在 YOLO 系列的版本迭代中,YOLOX 相比于 YOLOv3、YOLOv4 和 YOLOv5,检测帧率和准确率更高,并提供了强大的数据增强,在数据集容量较小的时候更能取得更好的学习效果,意味着更好的检测效果。基于深度学习的落石检测系统中,Two-Stage 模型需要经过产生候选区域—提取特征框—分类与回归的流程,在动态场景下,若目标移动速度较快,可能出现还没来得及检测目标就消失的情况;One-Stage 模型中 SSD 系列和 YOLO 系列检测速度都较快,但是 YOLO 系列版本迭代快,检测效果更好,适合落石检测的性能要求。

本文开展的实验基于 YOLOX 模型,在落石的目标检测问题上,由于落石的大小和形状差异较大,背景环境复杂,为了得到更好的检测和训练效果,首先对主干网络中的空间金字塔结构进行优化,以获得更多的语义信息,其次在骨干网络引入 ECA-Net 注意力机制^[17],使网络更关注落石目标本身的特征,降低周遭环境的干扰,且该注意力机制参数量少,几乎不增加模型的复杂度,在不损失检测性能的情况下,提高对落石目标的特征提取能力,并加强各层特征间的信息传播。最后对损失函数进行改进,降低因样本不均衡带来的训练效果较差的问题通过自制落石数据集,训练出的网络可以准确地检测运动中的落石。实验结果表明,改进后的 YOLOX 落石检测方法相比原算法,对于落石情况发生的检测更加灵敏,并且降低了落石检测的误报率,检测性能与原始网络无异。同一检测场景下,平均精度更高,符合实时目标检测的要求。

1 相关算法

1.1 YOLOX 算法

YOLOX 算法结构如图 1 所示,分为 Backbone、Neck 和 Head 三大部分。

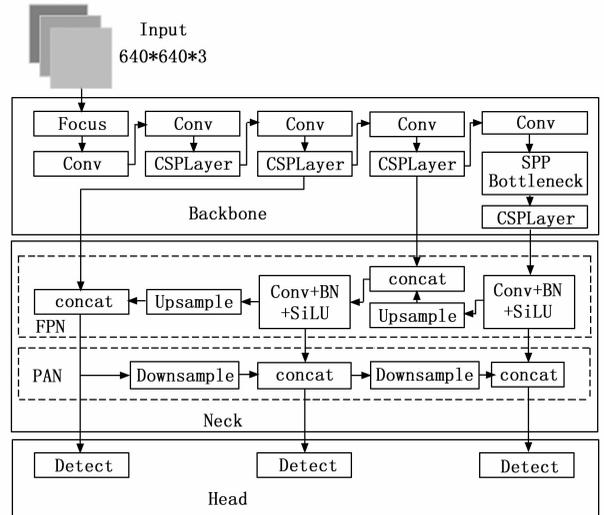


图 1 YOLOX 算法结构

YOLOX 的骨干网络采用 CSPDarknet,该网络基于 ResNet 实现,其中主要模块的功能分别如下。

Focus 模块用于增加特征图的通道维度,输入图片进入后,按照纵横两个方向将特征图切割为像素点,之后再间隔选取像素点,进行拼接与堆叠,将原始的单层特征图变为四层特征图,特征图的通道长宽变为原来的一半,深度扩充为原来的四倍。借此模块可以减少特征图参数量,提升正向传播和反向传播的速度,并且不减少图像的特征信息。Conv 模块是网络中的基本卷积单元,输入特征图进入 Conv 模块后,其中 Conv2D 层进行二维卷积操作,BN(批量归一化)层进行正则化操作,而激活函数层则使用 SiLU 函数进行激活操作。CSPLayer 采用了 ResNet 的残差结构,将 CSPLayer 模块划分为两个不同部分:其中一部分作为残差主干结构,对残差块进行堆叠,增加神经网络的深度;另一部分作为原始结构保留,最后将两部分拼接,构建跳跃连接。这种方法相当于在残差块之间增加了一个捷径,可以重复利用梯度信息,利于增加网络的深度,缓解梯度消失等问题。SPPBottleneck 模块中,除了使用基本卷积单元 Conv 执行卷积操作以外,还有池化层进行池化和拼接操作,该层采用最大池化的方法,通过改变核的大小来提高感受野。骨干网络的各模块图如图 2 所示。

YOLOX 网络的 Neck 部分由 FPN (feature pyramid networks, 特征金字塔)^[18]和 PAN (path aggregation network, 路径聚合网络)^[15]构成,其结构如图 3 所示。FPN 用于自顶向下传播语义特征,而 PAN 则是自底向上传播语义特征,FPN 与 PAN 两者结合,可以构建高层语义特征图以及底层特征图,并利用 PAN 的路径增强方法提高特征信

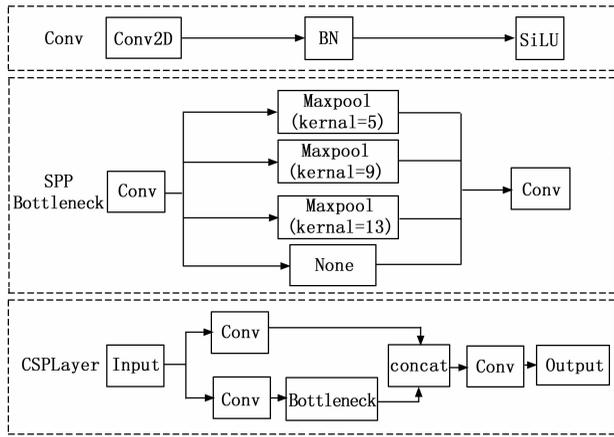


图 2 骨干网络的模块框架

号, 从而加快高低层特征之间的信息传播速度, 实现特征融合。

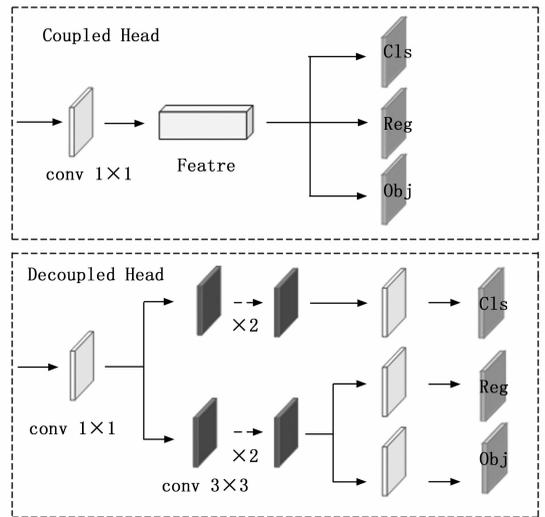


图 4 耦合头与解耦头的结构图

先使用一个卷积层和 3 个池化层对输入的特征图进行处理, 将输入通道缩减为原来的一半, 然后将所得结果和上一步输出的结果进行叠加处理, 此时得到新的特征图的通道数是最开始输入特征图的通道数的两倍。最后, 再进行卷积处理通道数减半, 此时得到的特征图大小和输入特征图大小保持一致, 其结构如图 5 (a) 所示。

SPPCSPC 结构在 SPPBottleneck 基础上加深了网络, 采用残差结构, 其结构如图 5 (b) 所示。

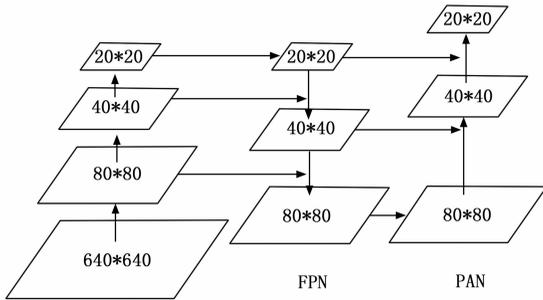


图 3 FPN 与 PAN 结构

YOLOX 网络的 Head 部分为检测结构, 也称为检测头。原始特征图经过上一层结构处理后, 生成三个不同尺寸的特征图, 应对不同尺寸目标的学习或预测。如图 4 所示, YOLOX 使用了 Decoupled Head (解耦头) 结构, 相比以往模型中的 Coupled Head (耦合头) 结构, 耦合头用一个 1×1 卷积涵盖了分类和回归步骤, 而解耦头则将分类和回归步骤解耦, 先使用卷积对输出进行降维, 在分类和回归两个流程中, 各使用两个 3×3 卷积。将检测头替换为解耦头后, 预测的精度提高且网络的收敛速度更快。

1.2 SiLU 激活函数

激活函数对于网络模型学习非线性特征具有非常重要的作用, 可以解决线性模型解决问题的有限性。YOLOX 模型采用了 SiLU 激活函数, 函数如公式 (1) 所示。

$$f(x) = x \cdot \sigma(x); f'(x) = f(x) + \sigma(x)(1 - f(x)) \quad (1)$$

SiLU 是 Sigmoid 和 ReLU 的改进版, SiLU 激活函数具备无上界、有下界、平滑、非单调的特性, 相较于 ReLU 更平滑, 其在深层模型上的效果优于 ReLU 激活函数。

2 算法改进

2.1 空间金字塔池化结构改进

空间金字塔池化结构广泛运用于 YOLO 系列网络中, 在 YOLOX 网络中使用 SPPBottleneck 进行特征提取, 它首

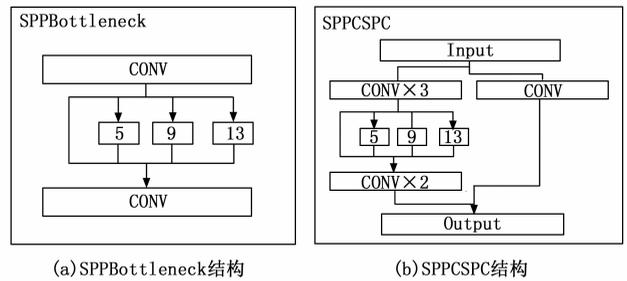


图 5 金字塔池化结构改进

在落石检测中, 由于背景和目标相似, 背景与目标难以完全区分, 通过 SPPBottleneck 提取到的语义信息较少, 不能完全提取出落石目标的特征, 而 SPPCSPC 加深了网络, 通过不同层次网络所提取到的信息将会更多, 获得的语义信息和特征信息更加明显。神经网络的加深容易导致梯度消失和梯度爆炸问题, 学界之前常用数据的初始化和正则化解决梯度问题, 而这种方法并不能解决网络性能退化问题, 而后残差结构的提出能同时解决随着网络加深出现的梯度问题和网络性能退化问题。因此 SPPCSPC 同时采用了残差结构, 提高网络的检测效率。本文通过将 SPPBottleneck 替换为 SPPCSPC 结构, 在落石检测中, 可以将纹理、颜色等外形相似的落石与山坡背景区分开来, 此外也可以减少数据训练产生的误差。

2.2 通道注意力改进

为了提取目标更精确的特征，增加特征信息的传播能力，引入通道注意力模块，可以使网络更关注检测物体本身而不是背景，因此在 YOLOX 中将 ECA-Net 注意力模块引入到骨干网络最后一个残差块，作为特征融合层，命名为 ECALayer。改进后的主干网络结构如图 6 所示。

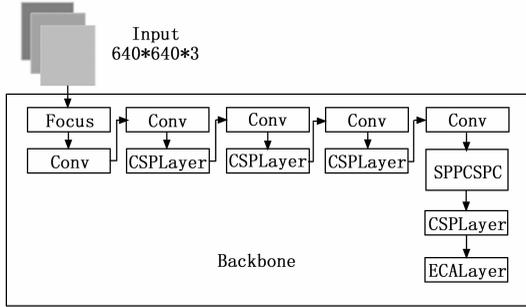


图 6 改进后的骨干网络结构

ECA-Net（有效通道注意力）基于 SENet（挤压激励网络）^[19]改进而来，它避免了 SENet 中因维数降低对通道注意力带来的不利影响，使用适当的局部跨通道交互，同时保持轻量化，ECA-Net 可以在几乎不增加网络模型参数数量的情况下，带来明显的性能提升。

如图 7 所示，SENet（squeeze-and-excitation networks）通过两个非线性的全连接层和一个激活层计算网络通道的权值，而 ECA-Net 将局部信息进行了跨通道交互，提高特征信息的利用率。该结构采用核大小为 k 的快速一维卷积计算通道权值， k 的取值范围受通道维度 C 的影响，其映射关系如公式（2）所示：

$$k = \Psi(C) = \left\lfloor \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right\rfloor_{odd} \quad (2)$$

k 与 C 采用指数函数进行映射， C 通常取值为 2 的整数幂。公式最后的 odd 代表取值最近的奇数。通过映射关系对内核大小 k 调优的方法，可以避免消耗过多的计算资源。

2.3 损失函数优化

深度学习中的损失类型由回归损失和分类损失组成，其函数分别用于衡量各个框之间的位置信息以及减少样本类别不均衡的问题。

YOLOX 中的回归损失函数分别采用 IOU 损失函数和 GIoU 损失函数，IOU 损失函数无法准确反应真实框与目标框的重合度，而 GIoU 损失函数改进了预测框和真实框的相交问题，但当预测框与真实框处于包含关系时，该函数也难以划分两个框的位置关系，降低算法的回归效果。为了改进以上缺点，CIoU 函数直接计算回归预测框和真实框的欧几里得距离来加速收敛，并解决各个框位置不理想的问题，因此在本实验中采用 CIoU 回归损失函数替换 IOU 与 GIoU 损失函数，以降低预测框和真实框位置关系不理想导致的误差。CIoU 损失函数如式（3）所示：

$$Loss_{CIoU} = IoU - \frac{D_1^2}{D_2^2} - \alpha \quad (3)$$

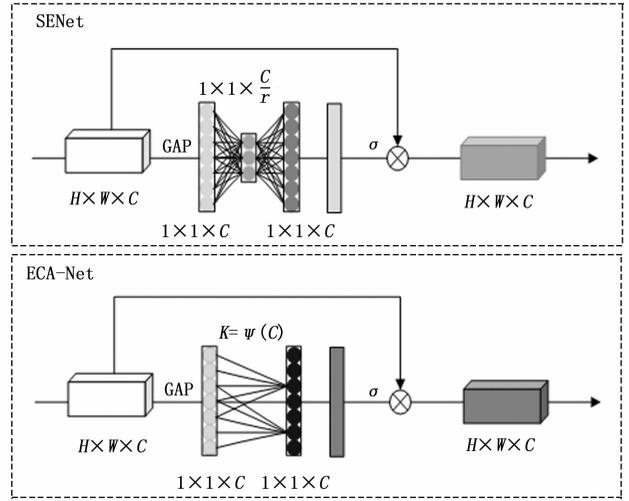


图 7 SENet 和 ECA-Net 模块结构

其中： D_1 为预测框到目标框中心点的欧几里得距离， D_2 为目标最小外接矩形的对角线距离， $\frac{w^{gt}}{h^{gt}}$ 和 $\frac{w}{h}$ 分别为目标框和预测框的宽高比。 α 代表影响因子， ν 代表长宽比一致性的参数，分别如式（4）和式（5）所示：

$$\alpha = \frac{\nu}{(1 - IoU) + \nu} \quad (4)$$

$$\nu = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (5)$$

CIoU 损失函数相比上两种损失函数，考虑了框的相交面积、欧几里得距离和框的长宽比等，预测框的回归更快且更准确。

YOLOX 中的原始分类损失函数采用二元交叉熵，该方法会产生大量的预选框来保证框能够覆盖检测目标，而其中只有小部分能准确涵盖目标，产生大量冗余框，导致没有检测目标的负样本数量过多，无法为算法网络的训练提供作用，导致算法学习无目标的背景，而无法分辨检测目标，减缓函数的迭代过程，并且可能无法优化至最优。因此将二元交叉熵损失函数替换为 Focal Loss 损失函数，如式（6）所示：

$$FL(p_t) = -\alpha_t(1 - p_t)^\gamma \log(p_t) \quad (6)$$

其中： α_t 用于平衡正负样本的重要性，取值范围在 0 到 1 之间，取值在区间范围内越大则更关注正样本，反正则更关注负样本。 γ 用于衡量简单样本和困难样本的重要性，取值范围在 0 到 5 之间，取值在区间范围内越大则越关注困难样本，反正则更关注简单样本， p_t 是 sigmoid 函数输出的预测值。为降低函数的参数数量，提出一种轻量的 Focal Loss 损失函数，其公式如式（7）所示：

$$FL(p_t) = -(1 - p_t)^2 \log(p_t) \quad (7)$$

因此考虑将变量 α_t 设为常量 1，将第变量 γ 设为 2，提高正样本与简单样本的重要性，使该损失函数在本实验中发挥最好作用。

2.4 数据增强

为了保证本次实验的样本数量饱和, 防止训练效果不好的情况发生, 引入 Mosaic^[20] 和 Mixup^[21] 数据增强方式。Mosaic 数据增强方式首先选择四张图片, 再分别进行图片的随机变换, 如使用裁剪或缩放的方式, 最后将这四张图片进行拼接。Mixup 数据增强方式则是随机选择两张图片进行拼接, 主要通过填充和融合的方式变换图片。扩充后的数据集能一定程度上提高算法训练效果, 但是经过以上两种数据增强生成的图片与真实图片有一定区别, 生成的数据量过多, 可能会带来不准确的标注框, 导致检测的结果脱离实际应用。因此本文实验仅在前 90% 的训练流程中采用数据增强。

3 结果与分析

3.1 实验环境

本文的实验环境为一台笔记本电脑, 使用的操作系统为 Linux, CPU 型号为 i7-10750H, GPU 型号为 GTX 2070, 显存大小为 8 GB, 内存大小为 32 GB。使用的所有模型基于 Pytorch1.9.1, cuda 版本为 10.1, cudnn 版本为 7.6.5。

在实际的应用环境中, 在可能发生落石的区域布置摄像头, 作为前端, 使用摄像头二十四小时监测, 远程的监测中心作为后端, 实时地对摄像头捕捉到的画面使用本文落石检测方法检测, 一旦检测到有落石发生的情况, 就通过摄像头监测区域附近的警示器向民众发出警告。实际应用环境下的整个系统的检测流程如图 8 所示。

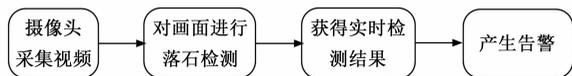


图 8 系统的检测流程图

3.2 数据集概述

本文的实验采用从网上收集的图片以及自己拍摄的真实环境的图片作为数据, 筛选后使用 labeling 软件标注, 一共得到了 518 个样本, 每份样本包含一张图片和一个 xml 文件, 作为训练所用的数据集。经过 Mixup 数据增强后, 共得到数据集 5 439 个, 将所收集数据的 80% 作为训练集, 剩余 20% 均分为验证集和测试集, 最终得到训练集、验证集和测试集的数量分别为 4351、544 和 544。数据集的标注过程如图 9 所示。

3.3 模型训练

本实验使用的输入图片分辨率大小设置为 640×640 , 选择较高输入图片分辨率可以增强网络对细节的学习效果。本文模型基于 YOLOXs 预训练权重, 避免从零训练网络, 可以提高网络训练速度, 缩短网络训练时间, 保证特征提取效果并减少权值随机等问题。

模型的训练分为两个阶段, 经过多次尝试, 决定采用以下模型的训练参数。设定训练的总共次数为 150 个 ep-

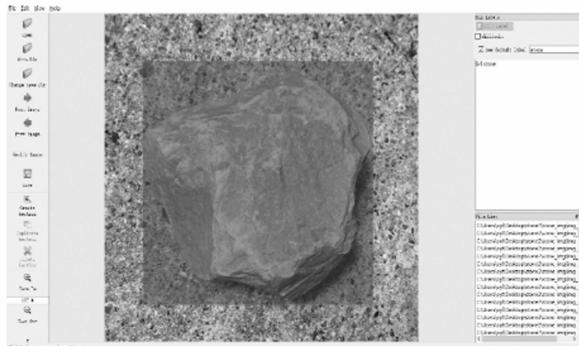


图 9 使用 labeling 标注数据

och, 其中: 0~50 个 epoch 采用冻结训练, 冻结用于图像特征提取的主干网络, 特征提取网络不发生改变, 仅对网络进行微调, 防止权值被破坏, 设置冻结训练的 batch size 为 8, 学习率为默认为 0.001; 51~150 个 epoch 采用解冻训练, 更新图像特征提取的主干网络中的全部参数, 占用显存较大, 为防止显存压力过大导致训练失败, 因此设置解冻训练的 batch size 为 4, 学习率设为默认的 0.000 1。

模型训练的迭代如图 10 所示, 在第 50 个 epoch 时, 由冻结训练到解冻训练的转换, 导致误差值发生波动, 在第 100 个 epoch 之后, 训练误差和验证误差的值保持稳定不变, 趋近于收敛。此时若选取第 150 个 epoch 作为预测权重, 可能出现过拟合的情况, 为防止过拟合情况的发生, 本文实验选用第 100 个 epoch 的权重作为预测权重。

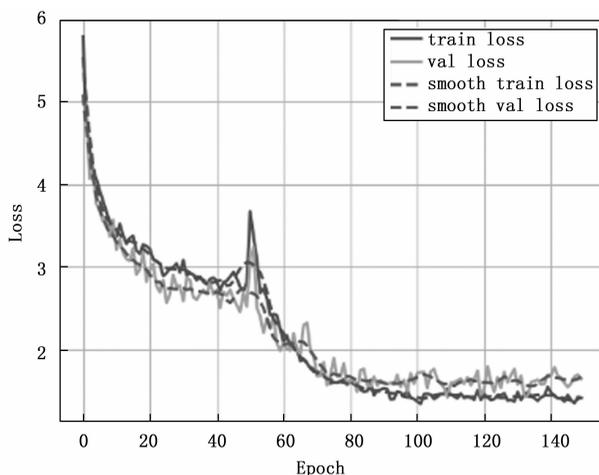


图 10 训练误差与验证误差的迭代

3.4 评估指标

本文采用每秒帧率 FPS (frames per second)、均值平均精度 mAP (mean average precision), 作为模型的评估指标, 其中 $mAP@0.5$ 为主要指标, $mAP@0.75$ 为辅助指标。

mAP 的计算基于 AP 的值, 计算所有类别的 $AP@0.5$ 并求平均值即得 $mAP@0.5$, $mAP@0.75$ 同理。计算公式分别如式 (8) 和式 (9) 所示:

$$AP = \int_0^1 PdR \quad (8)$$

$$mAP = \frac{1}{T} \sum_{t \in T} AP(t) \quad (9)$$

其中： P 代表网络检测精确度， R 代表网络检测召回率， T 代表类别数量，由于本实验中只有石头 1 种类别的样本，故 T 的值为 1。

3.5 结果对比

本文实验首先对比了 YOLOv3、YOLOv4、YOLOv5 和 YOLOX 的检测效果，用 $mAP@0.5$ 作为评估指标。同时本文也采用 YOLOX 算法和改进后的 YOLOX 算法应用于落石检测，作对比实验。

1) 对比实验：

所有实验均采用相同的数据集进行训练、测试和验证。进行模型对比实验，结果如图 11 所示。

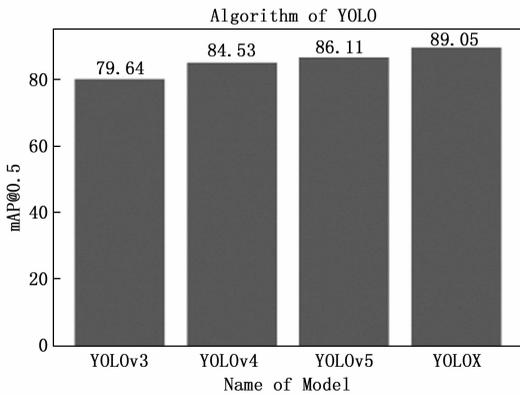


图 11 YOLO 系列模型精度对比

通过算法自带 FPS 性能检测，对同一张图反复检测，测得每秒可以检测的帧数，得到 YOLO 系列模型的对比实验结果如表 1 所示。

表 1 YOLO 系列模型性能对比

算法模型	mAP@0.5	FPS
YOLOv3	79.64%	37.2
YOLOv4	84.53%	45.1
YOLOv5	86.11%	50.4
YOLOX	89.05%	62.3

由实验结果可以看出，随着 YOLO 版本的迭代，检测精度在不断提升，其中 YOLOX 在 $mAP@0.5$ 的指标上相比以往版本有较大的提升，证明其性能较好，检测精度高，在检测结果上具有更高的准确率。

2) 消融实验：

如图 12 (a) 和 12 (b) 所示，分别使用 YOLOX 算法和改进后的 YOLOX 算法对测试集 544 个样本进行预测并评估，图 12 (a) 中，原生 YOLOX 的评估结果 $mAP@0.5$ 为 89.05%，图 12 (b) 中，在原生 YOLOX 上增加了 ECA 注意力机制、SPPCSPC 结构改进和损失函数替换，最终

$mAP@0.5$ 为 92.50%。

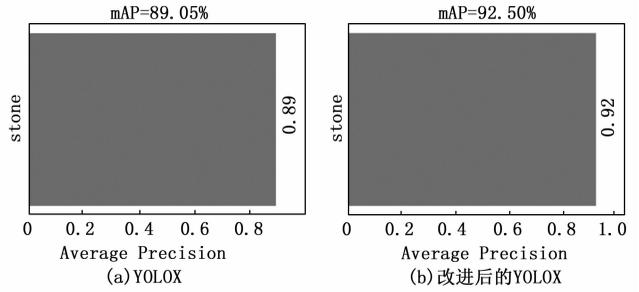


图 12 测试集评估结果对比

采用模型改进的消融实验，YOLOX 改进前后的对比实验结果如表 2 所示。

表 2 YOLOX 改进前后性能对比

算法模型	mAP@0.5	mAP@0.75	FPS
YOLOX	89.05%	84.10%	62.3
YOLOX+ECA	91.48%	85.70%	61.9
YOLOX+ECA+SPPCSPC	92.05%	86.14%	62.5
YOLOX+ECA+SPPCSPC + 损失函数替换	92.50%	86.61%	62.6

由实验数据可知，改进 YOLOX 算法在 $mAP@0.5$ 与 $mAP@0.75$ 在指标上取得了较大提升。相较于 YOLOX 算法，改进 YOLOX 算法的 $mAP@0.5$ 提升 3.45%， $mAP@0.75$ 提升 2.51%，对于形状不规则的石头识别准确率提升较大；FPS 保持稳定，性能损耗较小，改进后的模型依然满足实时检测的要求。

对比结果表明，改进的 YOLOX 落石检测方法的检测精度提升较大。引入空间金字塔池化、通道注意力和损失函数的改进提高网络对目标本身的关注程度，减少复杂背景信息带来的影响。

3.6 检测效果

本实验首先选用图片作为检测对象，如图 13 所示，将落石放置于复杂环境下以模拟落石发生的场景，有树枝、草地、落叶以及遮挡物的干扰，模型仍然能很好地对落石进行检测，且具有较高的置信度。

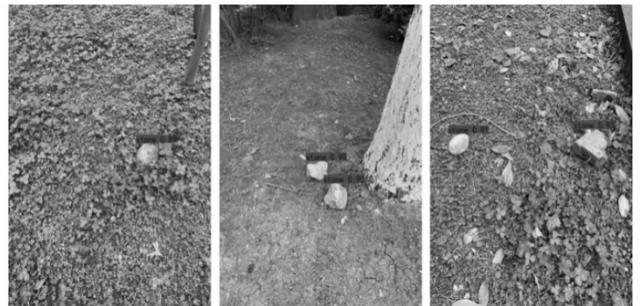


图 13 图片检测结果

本实验最终落实到视频流的检测上, 将输入视频转化为帧, 对每一帧进行检测。如图 14 所示, 截取了落石视频中的部分图片, 以展示改进 YOLOX 模型在视频检测中的效果。

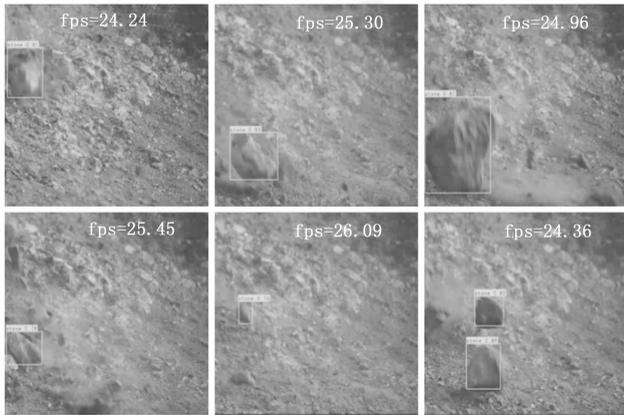


图 14 视频检测结果

用作落石视频检测的实验素材的视频帧率为 30FPS, 视频分辨率为 640×360 , 改进 YOLOX 的视频检测帧率相比于原视频帧率, 性能损失小, 且对于视频中形状不一的落石有较高的检测率。

4 结束语

本文提出了一种基于 YOLOX 的落石检测方法, 为了提高检测准确率和检测效果, 设计了一种改进的 YOLOX 模型, 在模型中对金字塔池化结构进行改进, 引入 ECA-Net 注意力模块, 对 Focal Loss 损失函数进行调整并替换原有函数, 此外采用了数据增强方法, 扩充数据集, 提升网络的学习效果, 训练网络时基于 YOLOXs 预训练模型, 采用冻结和解冻训练的方式, 避免网络的权值被破坏和特征提取不明显等问题。改进后的落石检测算法与原算法相比, mAP@0.5 提高了 3.45%, 达到了 92.50%, 每秒检测帧率提升 0.3, 验证了改进后的算法具有较高的识别精度与较快识别速度, 提升了图像与视频落石检测的检测效果。

因主要采取自拍图片、自收集数据的原因, 目前数据集中石头相关的样本较少, 训练后的网络可能在某些背景下的落石检测精度较低, 且有时会出现漏检或检测错误的情况, 发生误报, 我们后续工作将继续扩充数据集, 推进本方法的实际应用。YOLOX 作为较新提出的模型, 我们后续将探究 YOLOX 的更多的改进方法, 以满足更高标准的实际应用需求。

参考文献:

- [1] 胡厚田. 崩塌落石研究 [J]. 铁道工程学报, 2005 (S1): 387-391.
- [2] 叶四桥. 隧道洞口段落石灾害研究与防治 [D]. 成都: 西南交通大学, 2008.
- [3] 施晓文, 邓清禄, 董国梁. 崩塌落石对管道的危害性 [J]. 油气储运, 2013, 32 (3): 295-299.
- [4] 刘林芽, 吴送英, 左志远, 等. 基于 YOLOv3 算法的山区铁路边坡落石检测方法研究 [J]. 计算机科学, 2021, 48 (S2): 290-294.
- [5] 龙学军, 高 枫. 基于机器视觉的高速道路落石检测 [J]. 中国交通信息化, 2021 (7): 131-133.
- [6] 王 杰, 叶 茂, 马凤山, 等. 基于视频图像识别的崩塌落石监测预警系统设计与实现 [J]. 应用基础与工程科学学报, 2014, 22 (5): 952-963.
- [7] 简云端, 肖 硕. 基于数字图像处理的边坡落石识别算法研究 [J]. 科技创新导报, 2017, 14 (8): 88-89.
- [8] 黄晓红, 李铁锋, 刘祥鑫, 等. 基于改进 Faster R-CNN 算法的岩石裂隙发展方向跟踪预测研究 [J]. 河南理工大学学报 (自然科学版), 2022, 41 (4): 134-141.
- [9] 贾 曲, 田 原. 基于 Faster RCNN 的煤矿井下巷道顶板目标检测 [J]. 煤矿机械, 2022, 43 (5): 174-177.
- [10] 宁 健, 马 森, 柴立臣, 等. 深度学习的目标检测算法综述 [J]. 信息记录材, 2022, 23 (10): 1-4.
- [11] 刘 谱, 张兴会, 张志利, 等. 从 RCNN 到 YOLO 的目标检测综述 [C] // 第十六届全国信号和智能信息处理与应用学术会议论文集 [出版者不详], 2022: 16-23.
- [12] 耿 创, 宋品德, 曹立佳. YOLO 算法在目标检测中的研究进展 [J]. 兵器装备工程学报, 2022, 43 (9): 162-173.
- [13] 欧阳继红, 王梓明, 刘思光. 改进多尺度特征的 YOLO_v4 目标检测方法 [J]. 吉林大学学报 (理学版), 2022, 60 (6): 1349-1355.
- [14] 吴 珊, 周 凤. 基于改进 SSD 的小目标检测算法 [J/OL]. 计算机工程: 1-12 [2023-01-05].
- [15] 蒋 晨, 钱永明, 姚兴田, 等. 基于可变形卷积改进 SSD 算法的目标检测方法 [J]. 电子测量技术, 2022, 45 (16): 116-122.
- [16] 王 岩, 娄树理. 一种改进的 SSD 红外舰船目标检测算法 [J/OL]. 烟台大学学报 (自然科学与工程版): 1-9 [2023-01-05].
- [17] 毛志荣, 都云程, 肖诗斌, 等. 基于 ECA-Net 与多尺度结合的细粒度图像分类方法 [J]. 计算机应用研究, 2021, 38 (11): 3484-3488.
- [18] 杨锦辉, 李 鸿, 杜芸彦, 等. 基于改进 YOLOv5s 的轻量化目标检测算法 [J]. 电光与控制, 2023, 30 (2): 24-30.
- [19] 王春阳, 张英杰, 李长春, 等. 基于“残差-挤压激励”深度混合卷积网络的土地利用分类 [J]. 农业工程学报, 2022, 38 (1): 305-313.
- [20] 周 捷, 徐光辉, 朱东林, 等. 改进 YOLOV4-Tiny 的雨雾道路环境下的实时目标检测 [J]. 信号处理, 2021, 37 (8): 1550-1558.
- [21] 樊星男, 刘晓娟. 一种适用于轴承故障诊断的改进 Mixup 数据增强方法 [J]. 工程机械, 2022, 53 (4): 38-45.