

基于改进 YOLOv5 的城市地下管网缺陷识别算法

完颜健飞¹, 江雅馨², 徐晓龙², 常明¹, 黄英¹

(1. 中建七局第二建筑有限公司, 江苏苏州 215300;

2. 河海大学信息科学与工程学院, 江苏常州 213002)

摘要: 城市地下管网系统作为保障城市排涝安全的重要市政基础设施, 在长期超负荷运行过程中普遍存在着诸多病害问题; 传统检测技术 CCTV 依赖专业人员的专业技能以及先验经验, 因此为实现自动化的城市地下管网缺陷病害, 一种城市管网缺陷病害检测算法被提出并成功运用于实际工程中; 采用自适应 CA 注意力机制, 有效弱化复杂背景负面影响; 缺陷分类与回归的解耦关键方法, 使得检测部分充分利用缺陷纹理和边缘信息, 从而提高小尺寸缺陷的精度; SIOU 损失函数的运用为算法引入角度项权衡, 有效加快收敛速度; 经实验测试得到 71.1% 的平均精确度, 较原始算法提高 5.3%, 并满足了实际应用上的应用。

关键词: YOLOv5; 地下排水管道; 缺陷识别; 注意力机制; 目标检测

Defect Recognition Method of Urban Underground Pipe Network Based on Improved YOLOv5 Algorithm

WANYAN Jianfei¹, JIANG Yaxin², XU Xiaolong², CHANG Ming¹, HUANG Ying¹

(1. The second Construction Co., Ltd. of 7th Division CSCEC, Suzhou 215300, China;

2. School of Information Science and Engineering, Hohai University, Changzhou 213002, China)

Abstract: As an important municipal infrastructure for ensuring urban drainage safety, urban underground pipe network systems commonly suffer from many disease problems during long-term overload operation. The traditional detection technology CCTV relies on professional skills and prior experience of professionals. Therefore, in order to realize automated urban underground pipe network defects and diseases, an urban pipe network defect and disease detection algorithm is proposed, which is successfully used in actual projects. The adaptive CA attention mechanism is adopted to effectively weaken the negative impact of complex backgrounds; the key method of decoupling defect classification and regression enables the detection part to make full use of defect texture and edge information, thereby improving the accuracy of small-sized defects; The SIOU loss function provides the algorithm for the balance of angle, effectively speeding up the convergence. After experimental testing, this algorithm reaches an average accuracy of 71.1%, its average accuracy is 5.3% higher than that of the original algorithm, and it satisfies practical engineering application.

Keywords: YOLOv5 algorithm; underground drainage pipe; defect recognition; attention mechanism; target detect

0 引言

作为城市重要的基础设施之一, 城市地下管网系统承担着日夜交替的城市雨污水输送的重要任务。然而由于长期超负荷运行, 以及管理维护不及时, 管网缺陷成为近年来频发城市次生灾害的重要原因^[1]。闭路电视视觉检测技术 (CCTV, closed-circuit television) 目前是城镇排水管网缺陷病害检测中应用最广的方法之一。这种人工判读的方法不仅工作量大、耗时长, 管道的检测精确度也极大取决于判读人员的主观经验。因此, 推进自动化城市地下管道缺

陷分类系统的建设, 对于智慧城市管理有重大意义^[2]。

在以往研究中, 管网检测任务更依赖于传统机器学习与图像处理方法。文献 [4] 通过灰度值分割的图像处理技术对直方图进行计算实现管道缺陷的自动检测。但由于灰度值分割对噪声抗干扰能力较差, 在弱光、浑水的管网场景下, 往往存在精度较低、无法正常检测的问题。文献 [5] 使用梯度直方图 (HOG, histogram of oriented gradient) 处理缺陷图像获得潜在区域, 提取的 HOG 特征用于训练支持向量机 (SVM, support vector machine), 实现对

收稿日期: 2023-10-19; 修回日期: 2023-11-27。

基金项目: 住房和城乡建设部 2022 年科学技术计划项目 (2022-K-165); 中国建筑第七工程局有限公司局课题 (CSCEC7b-2022-Z-5)。

作者简介: 完颜健飞 (1990-), 男, 硕士研究生, 工程师。

通讯作者: 徐晓龙 (1974-), 男, 硕士研究生, 高级实验师。

引用格式: 完颜健飞, 江雅馨, 徐晓龙, 等. 基于改进 YOLOv5 的城市地下管网缺陷识别算法 [J]. 计算机测量与控制, 2024, 32(11): 258-264.

管网图像区别有无缺陷的检测。然而由于过程较长, 无法用于实时检测。文献 [6] 使用傅里叶变换、小波变换提取的特征训练 SVM, 达到对七种管网缺陷的检测目的。尽管传统方法的检测解决了部分问题, 但检测依旧受特征设计方法以及自身缺陷的影响^[7], 不同的缺陷对应不同的提取方法, 致使检测模型泛化能力弱且鲁棒性较差^[8]。随着计算机视觉领域的不断发展, 基于深度学习的计算机视觉目标检测算法也在各领域取得良好的表现^[9]。基于深度学习的视觉检测技术陆续被提出^[10], 并在许多如交通^[11], 工业工程^[12], 包括特定的地下管网^[13]等场景得到了广泛应用。自学习高维度特征的优势实现真正的端到端检测任务。文献 [14] 利用双阶段算法 Faster R-CNN 作为缺陷检测的跟踪框架, 结合度量学习模型提取特征, 提出一种跟踪缺陷检测的算法, 在连续 CCTV 视频跟踪识别任务中, 达到 57.4% 的 IDF1。文献 [15] 使用 YOLO v2 作为检测网络并部署于嵌入式设备上, 该网络覆盖了管网缺陷分类与检测的功能, 但其检测精度较低。文献 [16] 基于 YOLO v3 算法提出 VIASP 管网缺陷判读算法, 从视频中提取关键信息实现自动缺陷标记, 并最终以表格文本信息形式输出为评估报告。

相较于其余检测单阶段算法, YOLOv5 具有良好的泛化能力, 可以适应管网缺陷复杂的检测场景。本文针对实际地下管网场景提出了一种基于改进 YOLOv5 的城市地下管网的缺陷检测算法, 在原模型基础上对特征融合层, 边框回归损失函数, 以及检测部分做出进一步改进, 最后运用于苏州城市的地下排水管网的实际检测项目中。实验与运用表明本文模型能够对变形、障碍物、浮渣、脱落、树根共 5 种缺陷^[3]进行有效检测, 主要改进如下:

1) 针对地下管网视频采集过程中的弱光性, 以及狭小不稳定环境导致 CCTV 收集数据存在大量噪音的现象, 本算法融入坐标注意力机制模块 (CA, coordinate attention) 弱化噪声对检测结果的负面影响, 使模型更加聚焦于管道缺陷关键部分, 加强对管道内各类缺陷特征提取以及学习的能力。

2) 原始模型中定位和缺陷分类任务之间的相互影响, 导致管网缺陷检测精度下降的问题。针对该问题, 本算法提出了一种解耦的方法将缺陷定位与任务分离, 以加强小尺寸缺陷的检测能力, 同时提高了模型精度。

3) YOLOv5 使用的完全交并比损失函数 (CIoU, complete intersection over union) 未考虑角度因素, 这导致了管网缺陷检测中出现真实值与预测值之间的配准问题。因此本算法引入了 SIoU 损失函数 (SIoU, SCYLLA intersection over union), 从而加快模型整体的收敛速度并保持良好的鲁棒性。

1 城市排污管网缺陷检测算法

1.1 YOLOv5 整体结构

作为 YOLO^[17] 系列中最为广泛使用的算法, YOLOv5 综合性能超过了目前主流端到端检测算法如 YOLO v3^[18]、

SSD^[19], 其具体可以被描述为以下三部分。

1) 输入与特征提取:

首先切片操作将待检测图像分为四块通道数相等的特征图, 四张特征图在通道维度上进行拼接以此减少参数量。输入特征后续被划分为两条路径进行传播, 使得模型整体计算量更小。末端 SPP (Spatial Pyramid Pooling) 模块根据池化等级得到不同的特征图, 特征图随后拼接传递给后续的 neck 网络。

2) 特征融合层:

经由骨干网络所提取的 3 个尺度分别为 (256×20×20, 256×40×40, 256×80×80) 的管网缺陷特征图被送入至路径聚合网络 (PANet, path-aggregation network), 浅层图形特征与深层语义特征将在该网络内进行融合。

3) Head:

检测部分由 3 个卷积模块构成, 对应 3 个输出张量为 80×80×(5+5), 40×40×(5+5), 20×20×(5+5) 的特征层。其中, 5 个通道代 5 种缺陷类别, 4 表示边框位置坐标, 另一个通道表示可能存在缺陷的置信度, 共计 10 个输出通道。

1.2 改进的 YOLOv5 模型

考虑到实际管网内复杂多变环境, 以及现场设备硬件的限制, 模型不仅需要保持结构轻量化, 同时还存在一定的精度要求, YOLOv5 更适用于地下管网的缺陷检测^[21]。因此, 本文以 YOLOv5 作为基础网络, 提出一种改进的城市管网缺陷检测算法, 整体结构如图 1 所示。

1) 融入 CA 注意力机制的骨干网络:

首先, 尺寸固定为的待检测缺陷图像送入特征提取网络, 其中网络由数个卷积模块, C3 模块与 SPP (Spatial Pyramid Pooling) 模块共同构成。

随着网络结构加深, 骨干网络对图像卷积不断抽取特征信息, 特征从初始的图像纹理、边缘等浅层特征转变为深层语义特征, 最终由 FPN 结构将深层语义信息与浅层纹理信息相融合, 从而提升网络检测精度。然而, FPN 结构也存在以下待改进的两方面, 一方面特征融合时仅对特征图降维后简单合并, 然而不同特征层所关注信息不同, 直接融合将可能导致模型对噪音和复杂背景的鲁棒性降低。另一方面原模型采用 SPP 作为顶层模块, 仅包含单尺度的顶层特征经过不同尺寸内核不仅参数量较大, 信息也发生丢失。因此为实现对不同尺寸的同—管网缺陷的识别并保持模型参数轻量化, 本文结合注意力机制模块对多尺度特征捕捉, 合并获得丰富的细化特征图。

注意力机制源于人类视觉系统, 允许网络视野选择性关注重要信息忽略冗余信息。注意力机制使得网络能有聚焦地关注某些特定特征, 同时为主干网络提取基本特征进行筛选。相较于常见的注意力模块, 如 SE 注意力模块 (SE, squeeze-and-extraction) 对输入特征图通道压缩而忽略位置之间的互动; 或后续改进的注意力机制如 BAM 与 CBAM 减少输入张量的通道维度, 并仅通过卷积获得局部

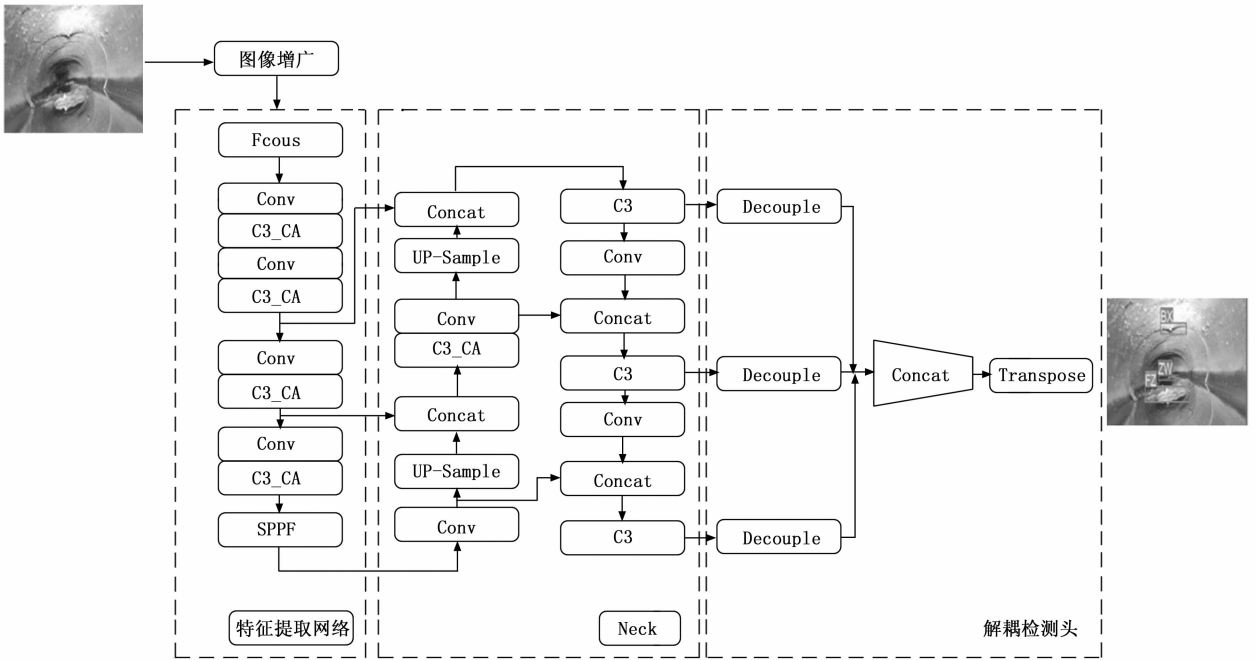


图 1 改进算法整体结构图

位置信息因此无法对远距离关系建模。

在实际任务中，地下管网的检测常常面对低光照度环境和大量背景噪音的挑战。尽管 YOLOv5 算法可以实现对不同尺寸和分辨率目标的检测，然而在复杂的管网背景中，多尺度融合不充分往往会导致缺陷检测不准确问题。因此缺陷检测需要一个能够满足特征融合需求的模块。基于上述考虑，为提高卷积神经网络在整体架构中的性能，本文引入坐标注意力机制 CA (Coordinate Attention)，使模型在特征图内自适应分配权重，以改善主干部分的特征提取能力。

CA 将二维全局池化操作转两个一维操作进行编码，更加关注特征图内像素的重要信息，而非将权重随机分配至整副特征图。注意力模块使模型在较低复杂度的条件下，将计算资源有效分配至管网缺陷特征图的各维度上，从而减弱噪音的负面影响，有利于模型更加专注学习和提取缺陷特征。CA 流程示意图如图 2 所示。

CA 流程主要可以分为两步骤。第一步，坐标信息嵌入模块中将坐标信息嵌入至输入尺寸为的特征图，并沿水平与垂直方向进行池化，得到每个通道的编码，计算如下式：

$$z^h(h) = \frac{1}{W} \sum_{0 \leq i < w} X(h, i) \quad (1)$$

$$z^w(w) = \frac{1}{H} \sum_{0 \leq j < w} X(j, w) \quad (2)$$

z^h , z^w 分别表示为尺寸为 $C \times H \times 1$ 和 $C \times W \times 1$ 的聚合特征张量，两个张量的转换允许模型能够在空间方向捕捉依赖关系，进而聚焦于重要的语义信息。

第二步，坐标注意力生成模块中将坐标信息嵌入部分输出的张量按空间维度进行拼接操作 (concat)，以获得编

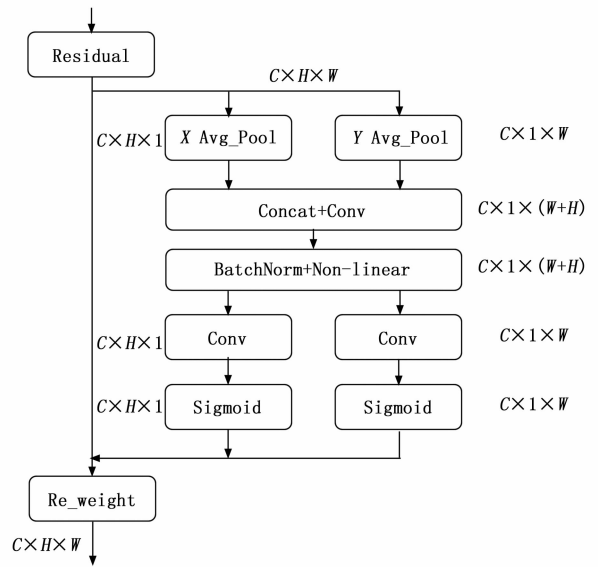


图 2 CA 结构流程

码了水平与垂直方向信息的中间特征图 f 。经过归一化与非线性激活处理后，全局信息 f 再被分割为两个独立向量 f^h 与 f^w ，通过特定的下式的计算过程得到相应权重：

$$g^h = \sigma(F_h(f^h)) \quad (3)$$

$$g^w = \sigma(F_w(f^w)) \quad (4)$$

其中： F_h 和 F_w 表示对 f^h 与 f^w 使用 1×1 卷积核转换，是注意力张量具有与输入相同的通道数。转换后，对张量应用 sigmoid 函数归一化，最终得出 g^h , g^w 用于加权的注意力权重。

针对城市地下管网的复杂多变的实际检测场景，本文

通过在主干网络融入 CA 注意力机制模块, 以一种更优的策略来捕捉位置与通道之间的互动关系, 从而增强网络特征学习能力。

模块的即插即用性质使得注意力能够方便地直接集成于网络架构中, 进一步降低了管网缺陷检测模型的部署难度。对比表 1 中实验数据, 引入 CA 注意力机制的算法检测相比原 YOLOv5 模型, *mAP* 指标有较好的提升。

表 1 引入 CA 注意力机制的性能评估实验

| 算法 | 召回率 | 精确率 | <i>mAP</i> /% |
|-----------|-------|-------|---------------|
| YOLOv5 | 0.661 | 0.657 | 65.8 |
| YOLOv5+CA | 0.685 | 0.668 | 67.1 |

2) 分类回归相解耦的检测头:

原模型在进行管网缺陷检测任务时, 检测部分保持分类与定位之间的耦合, 耦合进而导致的任务不协调, 进而使得检测结果出现偏差。分类任务更注重将提取到的缺陷特征与已知类别进行匹配, 而定位任务则关注于准确定位目标的位置坐标以进行边界框参数修正。YOLOv5 使用相同的神经网络提取, 所得特征同时运用于目标回归和目标分类, 参数共享进而导致两个任务相互影响。

为改善检测管网缺陷的精度问题, 本文引入解耦的检测头结构。通过分离分类与回归, 使得两任务分支可以自适应的在空间中寻找最佳位置, 不同的特征图分别应用于各自的分支。任务的参数独立, 从而避免了彼此之间的干扰。结构如图 3 所示。

解耦检测模块以骨干网络所提取的特征图作为输入, 3 种尺度从大到小进行堆叠为金字塔形式。以最大尺度的特征图输入为例, 首先通过由 1×1 卷积所构成通道缩减模块 CBL, 降低维度以此改善分支的计算量。通道数降低后, 特征信息分别流向分类分支和回归分支。分类分支由 CBL 模块, 1×1 的卷积模块与激活函数所构成。其中, 1×1 卷积负责将通道数转换为对应的类别数, 最终输出张量尺寸为 $80 \times 80 \times 5$, 其中 5 代表缺陷数量。与分类分支相同, 回归

分支也经过一个 1×1 卷积模块将通道数转换为位置信息, 张量尺寸为 $80 \times 80 \times 4$, 其中 4 表示预测框的坐标和尺寸 (x, y, w, h) 。回归分支还包括一个仅由卷积层构成的判断置信度模块, 其输出张量为 $80 \times 80 \times 1$, 单值判断是否存在目标。

解耦的分支获得的不同尺度结果特征按照维度进行拼接, 融合 3 个尺度的张量再次转置 (transpose), 最终得到待解码的预测框信息。针对城市地下管网的复杂环境特点, 通过解耦检测任务, 使模型加强了分类分支对管网缺陷类别的学习能力, 以及回归分支对缺陷纹理和边缘信息的关注, 进一步精炼出高质量的检测结果。对比表 2 数据, 集了解耦检测头的网络架构, 有效的提升了检测性能进一步提高回归精度。

表 2 引入解耦头有的性能评估实验

| 算法 | 召回率 | 精确率 | <i>mAP</i> /% |
|------------|-------|-------|---------------|
| YOLOv5 | 0.661 | 0.657 | 65.8 |
| YOLOv5+解耦头 | 0.668 | 0.670 | 67.5 |

3) 损失函数 Siou Loss:

改进的损失函数将形状损失也考虑在内, 计算如式 (5)。其中, (w, h) , (w^g, h^g) 分别表示预测框与真实框的长宽, θ 是一个用于控制对形状损失的关注程度的超参数。由于不同的训练数据集的先验信息不同, 针对本文所采用的地下管网缺陷数据集, 通过遗传算法得出最优的 θ 取值, 以避免对形状损失过度优化而影响预测框的自由运动:

$$\Omega = (1 - e^{-w\theta})^\theta + (1 - e^{-h\theta})^\theta \quad (5)$$

$$w_w = \frac{|w - w^g|}{\max(w, w^g)}, w_h = \frac{|h - h^g|}{\max(h, h^g)} \quad (6)$$

最终 Siou Loss 损失函数如式:

$$Siou_{Loss} = 1 - IoU + \frac{\Delta + \Omega}{2} \quad (7)$$

CIoU 以及先前的 DIoU, GIoU [20] 等方法未将方向因素融入惩罚指标, 这导致训练过程中网络输出的预测框

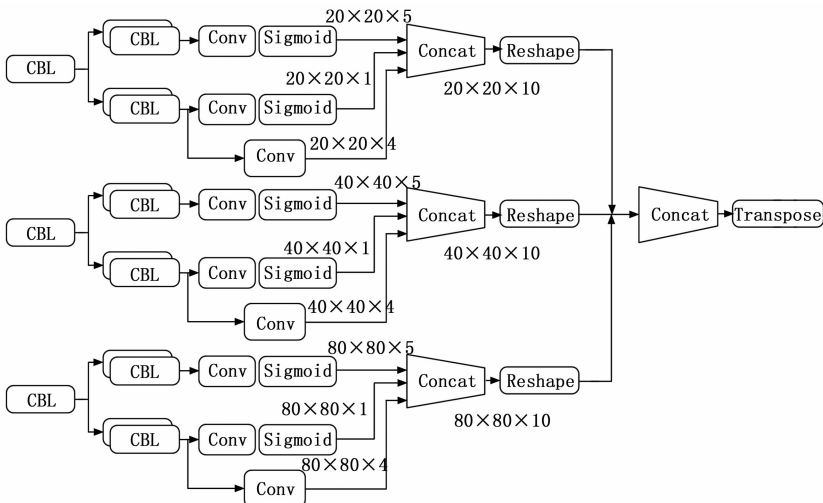


图 3 解耦结构部分

可能会在真实框周围游动，模型进而可能会出现难以收敛的训练结果。采用 SIoU Loss 损失从而引入回归所需的矢量角，对指标进行重新定义并加强对距离损失的惩罚，将真实框与预测框之间的方向匹配性纳入计算，有效缓解回归任务中预测框漂移问题。

2 实验与实验分析

2.1 管网缺陷数据集及增广

1) 缺陷数据集：

本文旨在对城下排水管网的沉积、树根、错位、脱落等 5 种缺陷进行识别和检测。所采用的数据来自于苏州市管网项目中巡检机器人所收集影像。从管网缺陷视频中共截取 2 764 张关键帧，经过数据清洗，最终保留 2 143 张图像，以满足较好的分辨率和信息丰富度要求，部分样本如图 4 所示。

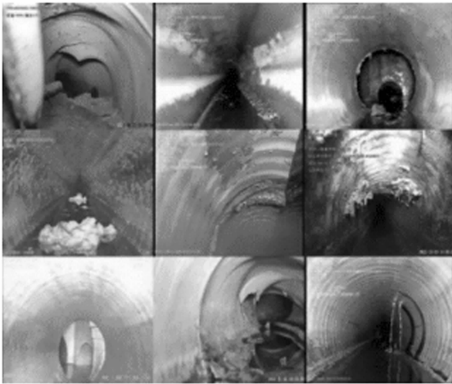


图 4 数据集部分样本

随后使用工具 LabelImg 对图像手工标注，作为地下管网缺陷检测模型的先验知识。为及时对模型进行超参数调整和性能评估，本文采用 k 折交叉验证法 (k-Fold Cross-Validation) 将数据集划分为训练集、验证集与测试集。

2) 图像增广：

为保证模型具备鲁棒性与泛用性，采用光度失真与几何失真两种典型方法数据增强。通过调整对比度、色调、饱和度等参数对图像进行光度失真方面的数据增强；随后添加随即缩放、裁剪、翻转等几何失真的数据增强法。对原始图像进行一系列随机变化映射，生成不同的训练样本，从而增大训练集规模。

为进一步提高数据质量与模型鲁棒性，Mosaic 数据增强方法也同样被运用，如图 5 所示。该方法首先从数据集中随机抽取 4 张图片并裁剪，重新拼接为一张新图像，根据混合区域大小对标签尺寸调整使用混合的新图像进行训练。模型在不消耗多余训练和推理代价的同时，变相增加模型批处理的图像数量，从原来仅处理单张图像提升至融合四张的新图像。此外，随机缩放尺寸较小的缺陷目标，也使得模型的对于小型缺陷检测定位再次得到改善。

2.2 实验环境与参数配置

本文所有实验均在操作系统 Windows10, Intel TM i5-

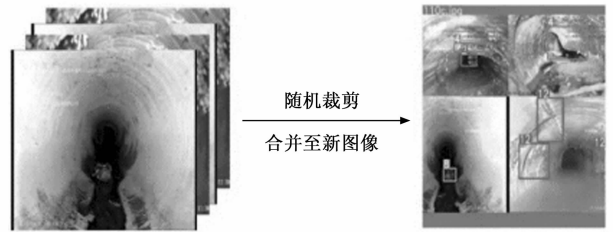


图 5 Mosaic 图像增强

11400F CPU, GeForce RTX1080TI GPU, 11GB 显存的 PC 上完成。采用 PyTorch1.7.0 作为深度学习算法框架，CUDA11.1 进行加速训练，如表 3 所示。

表 3 实验环境配置

| | |
|------|------------------------|
| 操作系统 | Windows10 |
| CPU | Intel TM i5-11400F CPU |
| GPU | GeForce RTX1080TI GPU |
| 训练加速 | CUDA11.1 |
| 训练框架 | Pytorch1.7.0 |

主要训练参数如表 4 所示，训练周期 (epoch) 为 300，单批传递至模型样本数 (batch size) 为 4 张，输入图像尺寸一致调整至 640×640，初始学习率 0.01，采用初始动量 0.937 的随机梯度下降优化器 (SGD, stochastic gradient descent)。

表 4 模型参数

| | |
|--------|---------|
| 配置 | 参数 |
| 训练周期 | 300 |
| 批处理量 | 4 |
| 输入图像尺寸 | 640×640 |
| 初始学习率 | 0.001 |
| 优化器 | SGD |
| 初始动量 | 0.937 |

2.3 评估指标

查准率 (Precision)，召回率 (Recall)，平均精确度值 (mAP, mean Average Precision)^[22] 作为基本指标用于评估模型的检测性能；真实值与预测值之间的交并比 (IoU, intersection over union)^[23] 将用以判断缺陷是否预测成功，具体设定交并比阈值为 0.5，低于阈值的预测框被视为错误预测。

查准率表示正确预测占包含假阳性 FP 与真阳性 TP 的所有阳性检测的比例，用于衡量模型在检测任务中的准确性，如式 (8) 所示：

$$precision = \frac{TP}{TP + FP} \quad (8)$$

召回率是正确预测占所有样本的比值，用于衡量模型对于所有真实缺陷样本的查找能力，即模型可以正确寻找多少真缺陷样本，较高的召回率表示模型具有较好的查找

全面性, 计算如式 (9):

$$recall = \frac{TP}{TP + FN} \quad (9)$$

由查准率与召回率计算求得出模型最终评估指标平均精确度值, 计算如下:

$$AP = \int_0^1 P(r) dr \quad (10)$$

$$mAP = \frac{1}{n} \sum_{i=1}^n AP(i) \quad (11)$$

AP 对应查准率与召回率曲线下的面积, 对 AP 求均值得到 mAP, 其中 N 代表地下管网缺陷类别数, mAP 考虑了模型在不同阈值下的性能, 当数值越高则代表算法检测效果越好。

改进算法旨在运用于城市地下管网巡检任务中, 因此需要达到实时检测缺陷的目标。因此除准确精确度之外, 本文还考虑了模型单位时间内处理的图片帧率 (FPS, frame per second), 用于评估模型动态性能, 并在处理速率上与其余网络模型进行比较。

2.4 检测性能分析

集成各种改进策略的管网缺陷检测算法按照实验参数配置进行多轮训练, 最终得到最优权重。为验证本文所提出的地下管网缺陷检测模型的泛化性能, 选择不同光照度包含多个目标的缺陷图像 (变形、障碍物、浮渣、脱落、树根) 进行测试。图 6 展示了改进算法的可视化输出结果, 每种缺陷类型都由一个不同颜色的边界框突出显示。

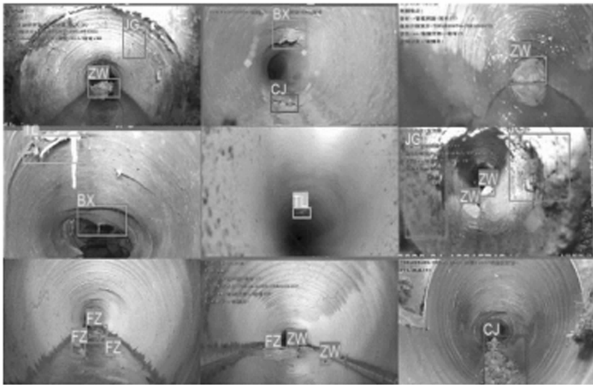


图 6 可视化结果

改进后检测算法 mAP 达到 71.1%, 结果表明即使在复杂的地下管网条件下, 如 CCTV 设备的方向与距离的变化、背景噪音、光照强度差异与不同缺陷纹理特征之间存在重叠等复杂条件下, 该模型也可以准确的检测到各类型的管网缺陷。

在测试集上的算法损失函数曲线如图 7 所示。随着训练轮数的增加, 损失逐渐下降并最终趋近于平稳。根据原模型与改进后的损失曲线对比可以看出, 基于 YOLOv5 的改进算法收敛速度较原模型更快, 并且损失也更小。

从表 5 中可以看出, 改进的管网缺陷算法在 mAP 方面优于主流的单阶段检测算法 SSD、YOLO v3、YOLO v4 以及

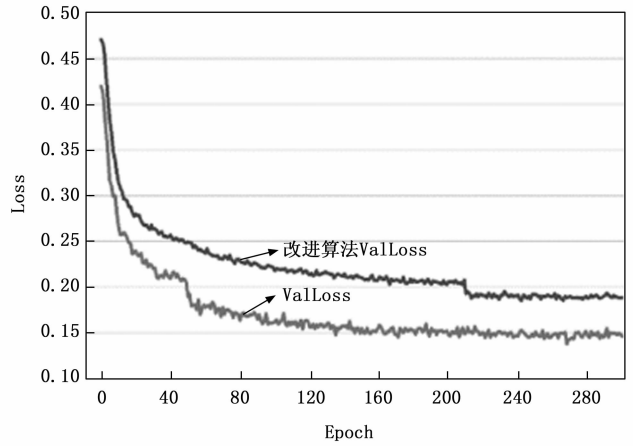


图 7 缺陷检测算法改进前后的 loss 曲线图

双阶段检测算法 Faster R-CNN, 达到了最佳的检测效果。由于双阶段与单阶段的网络结构上有区别, 因此检测结果也有着较大差异。Faster R-CNN 利用 RPN 网络产生的建议框进行判别, 尽管检测精度表现良好, 但效率较低。YOLOv3 与 YOLOv4 做到比 Faster R-CNN 与 SSD 更快的检测速度, 但在本文数据集的整体检测效果仍不太理想。对比其他模型, 本文提出的检测改进算法能有效的准确识别管网内缺陷, 并在不同管道场景下流畅应用于缺陷检测任务。

表 5 不同网络模型性能比较

| 算法 | mAP/% | FPS |
|--------------|-------|-----|
| SSD | 59.7 | 27 |
| YOLO v3 | 62.5 | 34 |
| YOLO v4 | 63.8 | 36 |
| YOLOv5 | 65.8 | 44 |
| Faster R-CNN | 69.4 | 19 |
| 本文算法 | 71.1 | 42 |

2.5 消融实验

本文针对原始模型的回归损失函数, 特征融合, 耦合检测头等方面进行了改进, 为验证所提出的改进算法的有效性, 将基于构建的管网缺陷数据集进行消融实验。“√”表示模型所应用的相应改进策略, “—”代表未改进的原始方法, 不同的改进策略对模型性能影响的实验结果如表所示。

表 6 消融实验数据

| 解耦头 | CA | SiO _U | 召回率 | 精确率 | mAP/% |
|-----|----|------------------|-------|-------|-------|
| — | — | — | 0.661 | 0.657 | 65.8 |
| √ | — | — | 0.668 | 0.670 | 67.5 |
| — | √ | — | 0.685 | 0.668 | 67.1 |
| — | — | √ | 0.674 | 0.683 | 66.1 |
| √ | √ | — | 0.687 | 0.697 | 69.4 |
| √ | — | √ | 0.671 | 0.689 | 68.2 |
| — | √ | √ | 0.696 | 0.709 | 70.6 |
| √ | √ | √ | 0.724 | 0.710 | 71.1 |

单独使用解耦检测头的 YOLOv5 模型相比 YOLOv5-CA 与 YOLOv5-SIoU, 在平均精确度上分别提高了 2.9% 和 2.2%, 说明表明解耦头使得模型的回归和分类分支各自关注到自身的任务, 从而提高了缺陷检测性能。融入 CA 注意力模块的主干网络具有更有效的特征提取能力, 通过抑制强水汽与低光度导致的噪音与无用信息, 使管网内的感兴趣特征更容易被关注到。同时, Siou 损失函数增添了角度惩罚指标, 避免了模型出现收敛结果较差的情况, 进一步减少了回归过程中的损失。将 3 种改进措施同时运用, 模型的各项在评估指标上均产生了正面优化, 性能得到了整体提高, 达到了 71.1% 的平均准确率, 取得了所有算法中最好的检测结果。

以上实验结果表明, 改进的管网缺陷算法在检测任务中取得了显著的效果, 为实现更准确和高效的缺陷检测提供了有力的支持。

3 结束语

地下排水管网的检测不及时可能会引发连锁的安全隐患问题, 在传统检测流程中, 人工判读缺陷效率较低, 且检测结果极大地取决于主观因素。因此本文提出一种基于 YOLOv5 的城市地下排水管网缺陷检测的改进模型, 将 CA 坐标注意力机制、解耦检测部分、Siou 损失函数 3 种改进措施联合用以训练模型, 从而实现对城市地下管网缺陷的自动化检测。实验表明本文所提出的算法能够有效提高缺陷检测性能, 多种改进策略相融和至算法内, 效果更为显著, 最终在本文数据集上达到平均精确度 71.1%。

目前, 城市管网检测这类任务所面对的复杂场景仍然存在许多问题, 譬如管道内部状况极为复杂、检测技术受各方面限制, 使得收集到的影像不够清晰, 上述问题将会是后续工作的研究重点。

参考文献:

- [1] 王淑梅, 王宝贞, 曹向东, 等. 对我国城市排水体制的探讨 [J]. 中国给水排水, 2007, 200 (12): 16-21.
- [2] 王和平, 安关峰, 谢广永. 城镇排水管道检测与评估技术规程解读 [J]. 给水排水, 2014, 50 (2): 124-127.
- [3] 中华人民共和国住房和城乡建设部. 城镇排水管道检测与评估技术规程 [M]. 北京: 中国建筑工业出版社, 2012.
- [4] DURAN O, ALTHOEFER K, SENERIRATNE L D. Automated sewer pipe inspection through image processing [C] // IEEE, 2002; 2551-2556.
- [5] HALFWAY M, HENGME J. Automated defect detection in sewer closed circuit television images using histograms of oriented gradients and support vector machine [J]. Automation in Construction, 2013; 1-13.
- [6] YE X, ZUO J, LI R. Diagnosis of sewer pipe defects on image recognition of multi-features and support vector machine in a southern chinese city [J]. Frontiers of Environmental Science, 2019, 13 (2).
- [7] 华夏, 王新晴, 王东, 等. 基于改进 SSD 的交通大场景多目标检测 [J]. 光学学报, 2018, 38 (12): 221-231.
- [8] 刘威, 靳宝, 周璇, 等. 基于特征融合及自适应模型更新的相关滤波目标跟踪算法 [J]. 智能系统学报, 2020, 15 (4): 714-721.
- [9] SALMAN M, YUKSEL S E. Fusion of hyperspectral image and LiDAR data and classification using deep convolutional neural networks [C] //2018 26th Signal Processing and Communications Applications Conference (SIU). IEEE, 2018; 1-4.
- [10] LOWE D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60 (2): 91-110.
- [11] 吕禾丰, 陆华才. 基于 YOLOv5 算法的交通标志识别技术研究 [J]. 电子测量与仪学报, 2021, 35 (10): 137-144.
- [12] 杨永波, 李栋. 改进 YOLOv5 的轻量级安全帽佩戴检测算法 [J]. 计算机工程与应用, 2022, 58 (9): 201-207.
- [13] SHUKLA H, PIRATLA K. Leakage detection in water pipelines using supervised classification of acceleration signals [J]. Automation in Construction, 2020, 117: 103256.
- [14] WANG M, KUMAR S, CHENG J. Automated sewer pipe defect tracking in CCTV videos based on defect detection and metric learning [J]. Automation in Construction, 2021, 121: 103438.
- [15] KLUSEK M, SZYDLO T. Supporting the process of sewer pipes inspection using machine learning on embedded devices [C] //ICCS, 2021; 77979.
- [16] YIN X, MA T, BOUFERGUENE A. Automation for sewer pipe assessment? CCTV video interpretation algorithm and sewer pipe video assessment (SPVA) system development [J]. Automation in Construction, 2021, 125: 103622.
- [17] REDMON J, DIVALA S, GIRSHICK R. You only look once: Unified, real-time object detection [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016; 779-788.
- [18] REDMON J, FARHADI A. YOLOv3: An incremental improvement [J]. ArXiv Preprint ArXiv: 1804.02767.
- [19] 陈幻杰, 王琦琦, 杨国威, 等. 多尺度卷积特征融合的 SSD 目标检测算法 [J]. 计算机科学与探索, 2019, 13 (6): 13.
- [20] REZATOFIGHI H, TSOIN, GWAK J Y, et al. Generalized intersection over union: A metric and a loss for bounding box regression [J]. CVPR, 2019, 00075.
- [21] 吕禾丰, 陆华才. 基于 YOLOv5 算法的交通标志识别技术研究 [J]. 电子测量技术, 2021, 35 (10): 137-144.
- [22] 刘颖, 刘红燕, 范九伦, 等. 基于深度学习的小目标检测研究与应用综述 [J]. 电子学报, 2020, 48 (3): 590-601.
- [23] 祝文韬, 谢宝蓉, 王琰, 等. 光学遥感图像中的飞机目标检测技术研究综述 [J]. 计算机科学, 2020, 47 (11A): 165-171.