

DOI:10.19651/j.cnki.emt.2210174

基于改进 YOLOX 的地下排水管道缺陷识别算法^{*}陆绮荣^{1,2} 丁 昕^{1,2} 梁雅雯^{1,2}

(1. 桂林理工大学信息科学与工程学院 广西 541000; 2. 广西“嵌入式技术与智能系统”重点实验室 广西 541000)

摘要: CCTV 检测技术在地下排水管道缺陷检测被广泛应用,但 CCTV 收集的缺陷图像需要依赖专业的检测人员进行检测识别,结果具有一定主观性且耗费大量时间。为了实现地下排水管道缺陷检测识别自动化,提出了一种基于改进 YOLOX 的地下排水管道缺陷识别方法。首先针对数据集过少的问题,通过 StyleGAN2 对原始图像进行预处理,生成多缺陷图像。其次,为了提高检测精度,对 YOLOX 的特征融合层进行改进,借鉴空洞卷积池化金字塔思想并引入 SE 注意力机制解决顶层特征仅包含单尺度特征且不与其它特征图进行融合的问题,同时设计了一种基于权重的特征融合模块,解决不同特征层融合带来的特征混叠问题。最后,将 YOLOX 边界损失函数改为 CIOU,提高目标检测框回归的效率。实验结果表明,所提的算法能够很好对沉积、渗漏、树根入侵、裂缝和错口 5 种缺陷进行识别,mAP 达到 68.76%,相比原始 YOLOX 算法提升了 1.62%。

关键词: 地下排水管道;YOLOX;缺陷识别;StyleGAN2;特征融合;目标检测

中图分类号: TP391.41 **文献标识码:** A **国家标准学科分类代码:** 520.6

A defect recognition method of the underground drainage pipe based on improved YOLOX algorithm

Lu Qirong^{1,2} Ding Xin^{1,2} Liang Yawen^{1,2}

(1. Guilin University of Technology, College of Information Science and Engineering, Guangxi 541000, China;

2. Guangxi Key Laboratory of Embedded Technology and Intelligent System, Guangxi 541000, China)

Abstract: CCTV inspection technology is widely used in underground drainage pipe defect detection, but the defect images collected by CCTV need to rely on professional inspectors for inspection and identification, and the results are subjective and time-consuming. In order to automate underground drainage pipe defect detection and identification, an improved YOLOX-based underground drainage pipe defect identification method is proposed. Firstly, for the problem of too small data set, the original image is preprocessed by StyleGAN2 to generate multi-defect images. Second, to improve the detection accuracy, the feature fusion layer of YOLOX is improved by borrowing the idea of convolutional pooling pyramid in the null space and introducing the SE attention mechanism to solve the problem that the top layer features contain only single-scale features and are not fused with other feature maps, and a weight-based feature fusion module is designed to solve the feature blending problem brought by the fusion of different feature layers. Finally, the YOLOX boundary loss function is changed to CIOU to improve the efficiency of target detection frame regression. The experimental results show that the proposed algorithm can well identify five defects, namely, deposition, leakage, tree root invasion, cracks and misalignment, with an mAP of 68.76%, which is 1.62% better than the original YOLOX algorithm.

Keywords: underground drainage pipe; YOLOX; defect recognition; StyleGAN2; feature fusion; object detection

0 引 言

目前地下排水管道检测应用最广的方法是闭路电视(closed-circuit television, CCTV)视觉检测技术,专业人员

通过 CCTV 收集的视频和图像对管道缺陷进行判读和评级。但人工判读容易受到人眼主观性影响,并且需要大量人力和成本。建立智能化的管道缺陷检测和分类系统,已成为城市排水设施建设和管理的迫切需求^[1]。

收稿日期:2022-05-31

^{*} 基金项目:国家自然科学基金(62166012)

在以往的研究中,一些研究者采用传统的机器学习方法,比如基于形态、几何和纹理特征来诊断缺陷^[2-4],随着算法的不断发展,基于深度学习(deep learning, DL)的方法也被应用在地下排水管道缺陷检测研究中,以提高污水检测的整体性能。基于深度学习的目标检测方法利用卷积神经网络来自动提取有效特征,克服了手工特征泛化能力差以及鲁棒性差等问题^[5]。Kumar 等^[6]将卷积神经网络(convolutional neural network, CNN)应用在对地下排水管道的树根入侵、沉积物和裂缝的 3 种缺陷进行检测识别,平均准确率达到 87.7%;Wang 等^[7]采用了 Faster R-CNN 网络对地下排水管道的六种缺陷进行检测识别,平均准确率达到 88.99%;Li 等^[8]设计了一种新的基于两阶段的目标检测算法,利用多层全局特征融合技术,对地下排水管道的缺陷进行检测识别,模型的 mAP 达到 50.8%。但随着检测的缺陷类型增多时,检测精度并不高。另外地下排水管道缺陷图像难以获取,存在数据样本较少的问题。为此,本文采用 YOLOX 的算法对地下排水管道的 5 种缺陷进行检测识别,5 种缺陷分别是沉积、裂缝、渗漏、树根入侵以及错口。为了进一步提高模型对地下排水管道缺陷检测的精度,在原 YOLOX 算法的基础上,本文对 YOLOX 算法的特征融合层进行了改进,并采用了 CIUO 作为边界框损失函数,同时针对数据样本不足,本文使用了 StyleGAN2 网络对部分缺陷类型进行数据扩充并使用 5 种数据增强方法

进一步扩充数据集。最后在自制的地下排水管道缺陷数据集上进行实验,实验结果表明本文模型能够很好地检测识别到地下排水管道的 5 种缺陷。

1 YOLOX 算法及改进

1.1 YOLOX 算法

YOLOX 算法^[9]是旷视科技研究院 2021 年提出来的新一代 YOLO 系列,它将解耦头、数据增强、无锚点以及标签分类等目标检测领域的优秀进展与 YOLO 进行了巧妙的组合^[10],通过一系列的改进,YOLOX 超越了目前主流的目标检测算法 YOLOV3^[11],YOLOV4^[12]以及 YOLOV5。本文采用 YOLOX-S 版本作为基础网络,其整体网络结构如图 1 所示,它由 4 部分组成:输入层、主干特征提取层、特征融合层、预测层。其中主干特征提取层是由 Focus 模块以及叠加一定数量的 CSP 模块,并且使用 Silu 激活函数,最后经过 SPP^[13]模块得到主干网络特征提取结果,该结构使得局部特征与全局特征相融合,丰富了特征图的表达能力。特征融合层采用特征金字塔网络^[14](feature pyramid networks, FPN)结构,FPN 结构自顶向下传达强语义特征,并将不同尺度的特征图进行融合。预测层仍然使用大、中、小的 3 种不同比例尺特征图对目标进行检测,同时采用解耦的检测头分别计算分类和定位的结果。

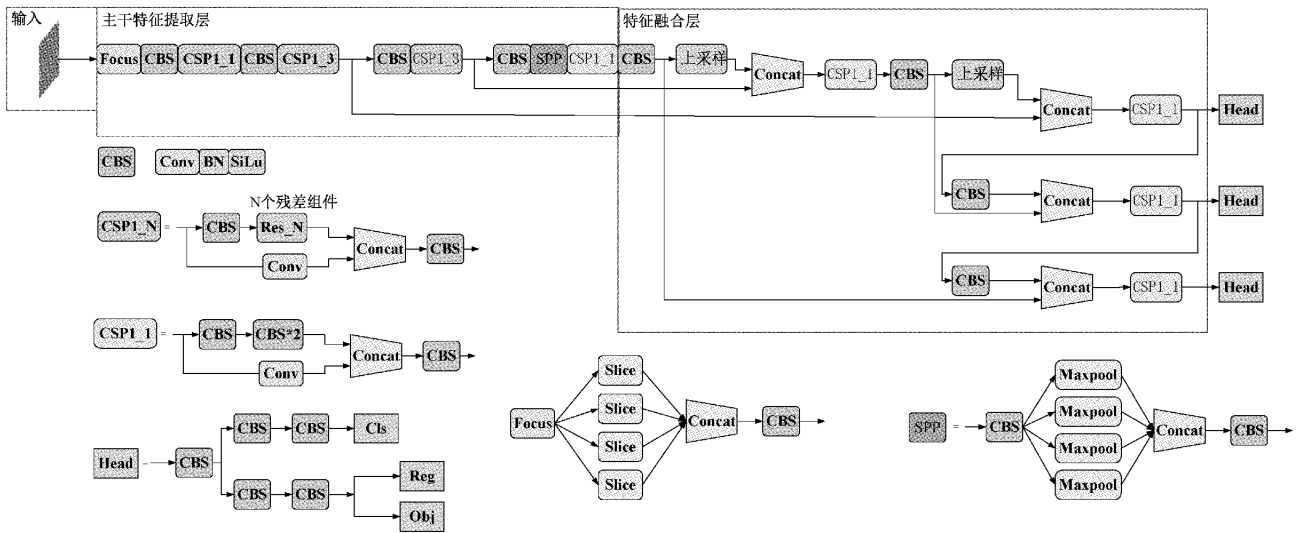


图 1 YOLOX 网络结构图

1.2 网络改进

FPN 网络将深层网络的语义信息与浅层网络的位置信息结合起来,提升了网络的检测精度^[15]。FPN 结构见图 2 所示。然而传统的 FPN 结构有以下的严重问题,一方面在进行对特征图进行融合时只是简单的进行直接合并,由于特征之间的语义差距和相似性,这种方法是不可行的。为了防止不同层的特征图直接融合导致特征混叠问题,本文设计了一种基于权重的特征融合模块,利用权重

系数来融合不同层的特征。另一方面,信息丢失主要发生在顶层特征,顶层特征仅包含单尺度特征且不与其它特征图进行融合,因此本文利用空洞空间卷积池化金字塔^[16](atrous spatial pyramid pooling, ASPP)的思想同时引入压缩激励模块(squeeze and excitation block, SE block)^[17],对顶层特征进行多尺度特征捕获,然后经过通道注意力机制获取更丰富的细化特征图最后合并。改进后的 FPN 结构如图 3 所示。

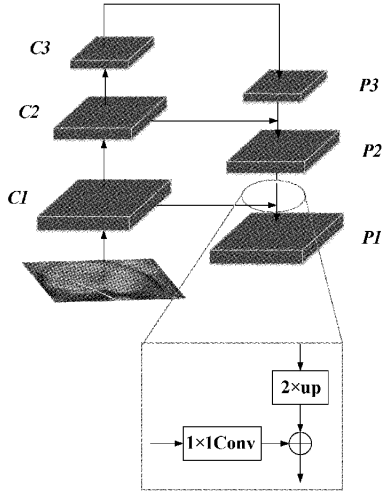


图2 FPN网络结构

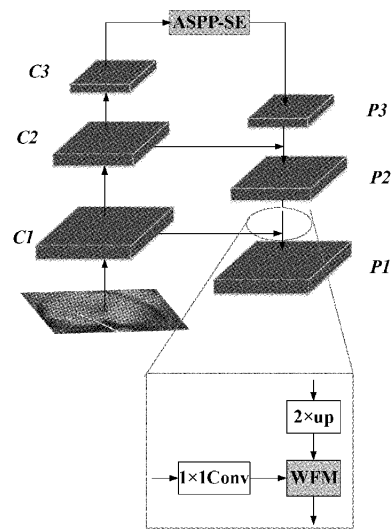


图3 改进的FPN网络结构

1.3 基于改进的空洞空间金字塔模块

在特征融合层中,最高层的特征图由于通道数的减少导致信息丢失,并且仅包含单尺度信息且不与其它特征图融合,因此需要进行改进。ASPP通过不同空洞率的空洞卷积并行采样,有效地提取了多尺度物体的特征,但是该模块没有考虑到特征图通道之间的相互关系。对此本文利用ASPP并进行了一定的改进,改进的空洞空间金字塔

模块结构如图4所示。首先本文将输入的特征图分别输入到5个分支,通过引入4个3x3大小不同空洞率的空洞卷积,以获取不同尺度的感受野,提取多尺度的丰富上下文信息,再通过将各分支输出的特征图经过SE模块,得到通道间的注意力向量,然后将原始的特征图进行乘积运算,得到一个更丰富的细化特征图作为输出,最后将各分支的输出进行合并。

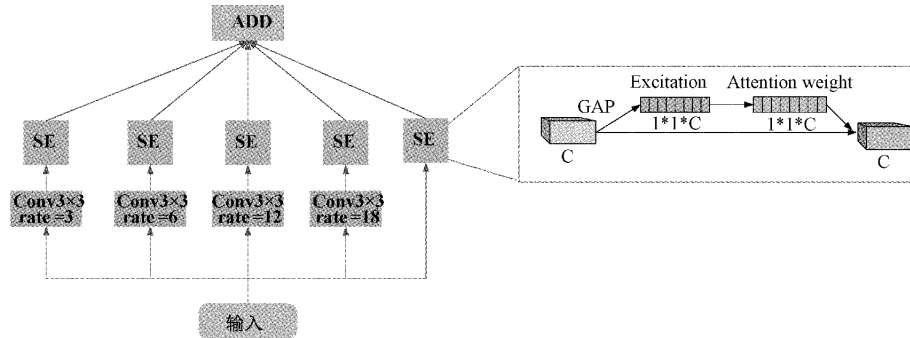


图4 改进的空洞空间金字塔模块结构图

1.4 基于权重的特征融合模块

在FPN结构中,特征融合操作可以用式(1)表示为:

$$P2 = F1 \oplus F2 \tag{1}$$

F1表示图2中C2经过1x1卷积降维的特征,F2表示图2中P3经过下采样的输出。由于直接将两层特征进行简单的融合,并没有考虑到特征之间的语义差别,会导致特征混叠问题,因此本文设计了一种基于权重的特征融合模块,结构如图5所示。该特征融合方式用式(2)可表示为:

$$P2 = F2 \otimes (W(F2, F1)) \odot F1 \tag{2}$$

其中W(F2, F1)是一种权重函数,对两个特征进行融合并获得权重系数,结构如图6所示。其主要步骤是先将输入的两个特征图F2和F1进行融合,融合后大小为HxWxC,将融合后的特征经过置换注意力机制模块(shuffle attention, SA)^[18],以获取更多所需要关注目标的细节信

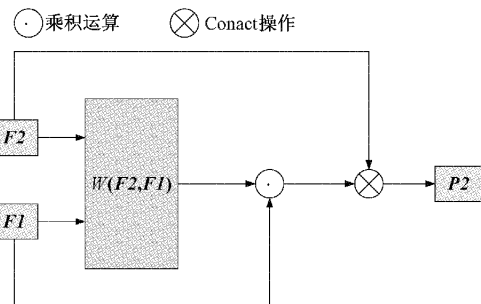


图5 基于权重的特征融合模块结构图

息,抑制其他无用信息。然后将其分别输入到maxpool和avgpool并将通道减少到1,最后本文将这两个特征进行相加并使用sigmoid函数得到权重函数W(F2, F1),大小为HxWx1。

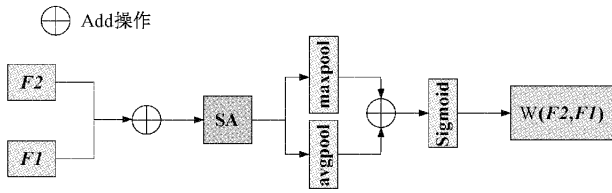


图 6 权重函数结构图

1.5 SA 注意力

SA 是一种轻量且高效的注意力机制模块,SA 结构图如图 7 所示,SA 首先将输入的特征进行分组,假设输入的特征为 $P \in R^{C \times H \times W}$, 首先将特征 P 的通道拆分为 G 组: $P = [P_1, P_2, \dots, P_G]$, 然后将每一组特征 P_k 分成两个分支: $P_{k1}, P_{k2} \in R^{C/2G \times H \times W}$, 最后将 P_{k1} 和 P_{k2} 特征分别通过通道注意力机制和空间注意力机制以生成不同的重要性系数。

对于通道注意力机制模块,首先通过全局平均池生成基于通道的统计信息 $s \in R^{C \times H \times W}$, 然后经过 Fuse 线性函

数增强特征表示,最后通过 Sigmoid 函数激活后与原特征 P_{k1} 进行相乘。具体计算如式(3):

$$s = F_{sp}(P_{k1}) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W P_{k1}(i, j) \quad (3)$$

$$P'_{k1} = \sigma(F_c(s)) \times P_{k1} = \sigma(W_1 s + b_1) \times P_{k1}$$

其中, F_c 为线性函数, σ 是 Sigmoid 激活函数, W_1 和 b_1 为线性函数的参数。对于空间注意力机制模块,首先采用组归一化(group norm, GN)对 P_{k2} 进行获取空间统计数据,再利用 Fuse 线性函数增强特征表示,最后通过 Sigmoid 函数激活后与原特征 P_{k1} 进行相乘。具体计算如式(4)所示。

$$P'_{k2} = \sigma(F_c(F_{GN}(P_{k2}))) \times P_{k2} = \sigma(W_2 S_2 + b_2) \times P_{k2} \quad (4)$$

其中, F_{GN} 是 GN 函数, S_2 是经过归一化的特征, W_2 和 b_2 为线性函数的参数。

将经过通道注意力模块得到的 P'_{k1} 和空间注意力模块得到的 P'_{k2} 进行合并,最后将每一组特征进行合并。

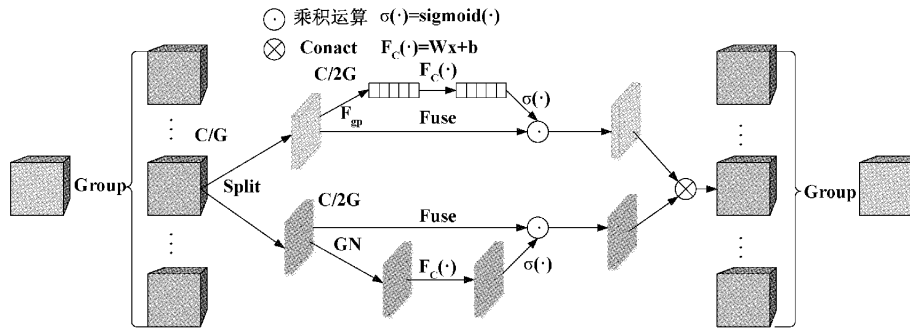


图 7 SA 注意力机制结构图

1.6 边界损失函数改进

YOLOX 使用交并比(intersection over union, IOU)作为边界损失函数,它是衡量真实目标框与预测框之间的重合度,重合度越大,IOU 的值也就越大。IOU 的定义如下:

$$IOU_{(A,B)} = \frac{A \cap B}{A \cup B} \quad (5)$$

其中, A 为预测框的面积, B 为真实目标框的面积。由公式可以看出,当预测框和真实框不相交时,两者重叠的面积为 0,不能衡量出两框之间的距离,这会导致梯度不能回传,学习训练不能进行网络无法收敛。并且 IOU 的值只跟两框之间的重合面积有联系,无法反映出两框重合度的大小。为了解决上述问题,本文采用 CIOU^[19] 作为边界损失函数,CIOU 直接对两框中心点距离进行优化,大大加快两框靠近的速度,使得网络收敛加快。CIOU 的定义如下:

$$CIOU = IOU - \left(\frac{p^2(b, b^{gt})}{c^2} + \alpha v \right) \quad (6)$$

其中, α 表示权重函数, v 是衡量长宽比一致性的函数, $p^2(b, b^{gt})$ 是预测框和真实目标框两框中心点的欧式

距离, c 表示能够同时包含预测框和真实目标框的最小包围区域的对角线距离。 a 和 v 的公式如下:

$$\begin{cases} a = \frac{v}{(1 - IOU) + v} \\ v = \frac{4}{\pi^2} \left(\arctan \frac{\omega^{gt}}{h} - \arctan \frac{\omega}{h} \right)^2 \end{cases} \quad (7)$$

CIOU 边界损失函数定义为:

$$L_{CIOU} = 1 - IOU + \frac{p^2(b, b^{gt})}{c^2} + \alpha v \quad (8)$$

2 实验过程

2.1 数据扩充

本文针对地下排水管道的沉积、裂缝、渗漏、树根入侵、错口这 5 种缺陷类型进行检测,数据集来源于某市地下排水管道检测项目中 CCTV 拍摄的视频和图片,从中筛选了一共 1 361 张原始图片,其中沉积类型 332 张,裂缝 163 张,渗漏 351 张,树根 326 张,错口 189 张。然而裂缝和错口的数据较少,会造成样本分布不均匀,将导致模型对少样本类型检测性能较低。因此为了解决上述问题本

文使用GAN^[20]网络来生成裂缝和错口这两种缺陷类型的图片。GAN网络包含生成器G和判别器D两部分,如图8所示,生成器G通过从输入的图片进行学习真实图像的分布特点,从而进行生成相似度逼近真实的假图像,而判别器D通过将生成的图像与真实图像进行判别,生成器G和判别器D通过不断的对抗博弈来进行迭代优化,最后生成高质量的合成图像。原始GAN生成的图像有很多缺陷,而且仅能从噪声生成图像,在应用上有一定的局限性^[21]。因此本文采用StyleGAN2^[22]对地下排水管道数据集进行生成以解决数据集不足。

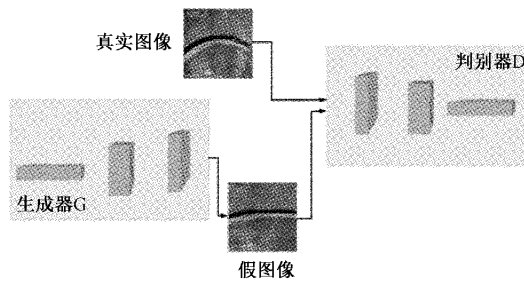


图8 GAN网络结构图

2.2 StyleGAN2 网络训练

StyleGAN2网络训练采用Adam优化器,生成器G和判别器D的初始学习率都为0.001。将裂缝和错口两种缺陷类型作为网络输入,由于样本数据过少,本文采用迁移学习的思想,StyleGAN2利用文献[23]的数据集进行迁移学习,经过长时间的训练,最后生成了700张图片。为了科学的评估StyleGAN2网络生成的缺陷图像的多样性和质量,本文采用文献[24]提出的弗雷谢特起始距离(Frechet inception distance, FID)评价指标,FID评价指标公式如下:

$$FID(x, g) = \|\mu_x - \mu_g\|_2^2 + Tr(\sum_x + \sum_g - 2(\sum_x \sum_g)^{\frac{1}{2}}) \quad (9)$$

其中, x 和 g 表示原始图像和生成图像, μ_x 表示原始图像特征分布均值, μ_g 表示生成图像特征分布均值, \sum_x 表示原始图像特征分布方差, \sum_g 表示生成图像特征分布方差, Tr 表示矩阵的迹。

StyleGAN2网络生成的地下排水管道缺陷图像的FID值如表1所示,从表中可以看出,生成的缺陷图像清晰度高且多样性。最终本文从生成的图片中挑选了166张裂缝图片,169张错口图片。图9展示了经过StyleGAN2生成的部分地下排水管道缺陷图。

表1 StyleGAN2的FID值

GAN网络模型	FID
StyleGAN2	48.27

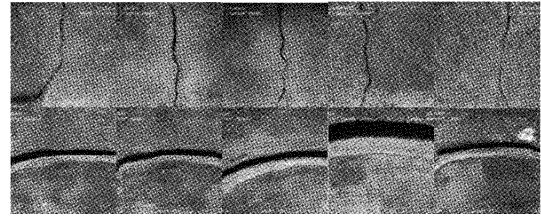


图9 StyleGAN2网络生成的部分缺陷图

2.3 数据集增强及数据集划分

为了进一步扩大数据集,采用平移、改变亮度、加噪声、旋转以及镜像这5种数据增强方式对缺陷图片进行增强,最后一共得到了6784张地下排水管道缺陷图像。将数据集按照8:1:1划分为训练集,验证集和测试集,图10展示了数据集部分图像。

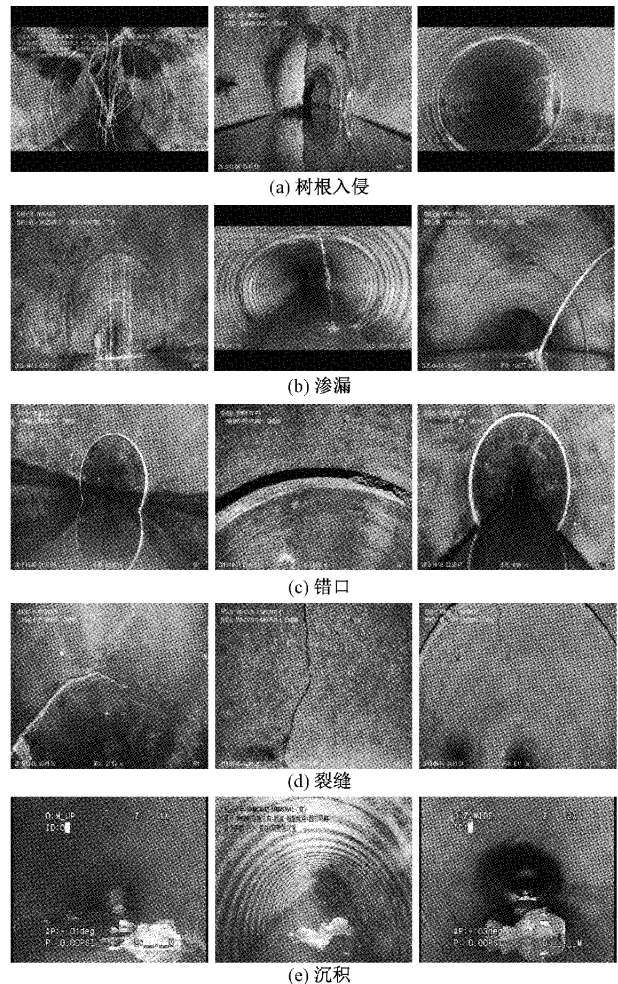


图10 数据集部分图像

3 实验结果分析

3.1 评价指标

为了定量的评价本文模型的性能,本文采用精确率(Precision),召回率(Recall),平均精度(average precision,

AP)以及平均精度均值(mean average precision, mAP)这四种常见的目标检测评价指标,各种评价指标公式如下:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (10)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (11)$$

$$\text{AP} = \int_0^1 p(r) dr \quad (12)$$

$$\text{mAP} = \frac{\sum AP}{N_{cls}} \quad (13)$$

其中, TP 表示正样本被模型正确预测为正样本的数量, FP 表示负样本被模型错误预测正样本的数量, FN 表示正样本被模型预测为负样本的数量, N_{cls} 表示所有缺陷的种类。

3.2 网络训练参数设置

将数据集输入到改进的 YOLOX 算法中,采用迁移学习的思想,导入 YOLOX 在 COCO 数据集上的预训练权重参数,冻结层的训练 epoch 轮次为 50 次,解冻层训练为 300 轮次,具体参数如表 2 所示。

表 2 训练参数设置

训练阶段	epoch	学习率	优化器	学习率调整策略
冻结阶段	50	0.000 625	SGD	余弦退火 衰减策略
解冻阶段	300	0.000 282	SGD	余弦退火 衰减策略

改进前后 YOLOX 算法的 loss 曲线如图 11 所示。从图中可知在 50 个 epoch 的时候模型开始解冻,loss 会有一定的升高,随着训练的进行,loss 逐渐下降并趋于稳定,并且从两图可看出,改进后的 YOLOX 算法在训练时收敛的速度更快,并且稳定性更好,loss 值也更小,说明改进的模型训练的效果要更好。

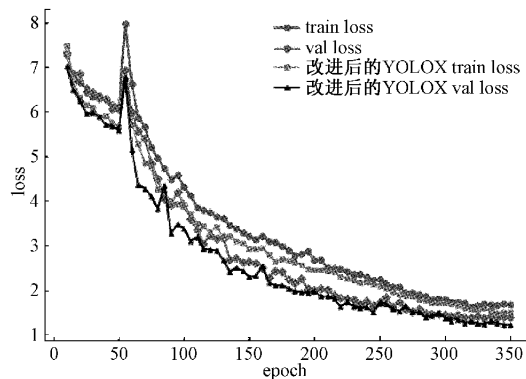


图 11 改进前后 YOLOX 算法的 loss 曲线图

3.3 实验结果

本文算法检测结果如表 3 所示,本文模型对 5 种缺陷

类型检测识别的精确率为:沉积为 98.53%,树根入侵为 98.46%,渗漏为 98.92%,裂缝为 97.15%,错口为 98.15%,mAP 为 68.76%,达到了较高的识别效果。图 12 展示了本文模型对部分样本的检测识别效果。

表 3 本文算法实验结果

评价指标	沉积	树根入侵	渗漏	裂缝	错口
Precision	98.53	98.46	98.92	97.15	98.15
Recall	60.27	68.75	70.81	73.25	70.57
AP	60.06	67.82	70.95	73.10	71.85
mAP	68.76				

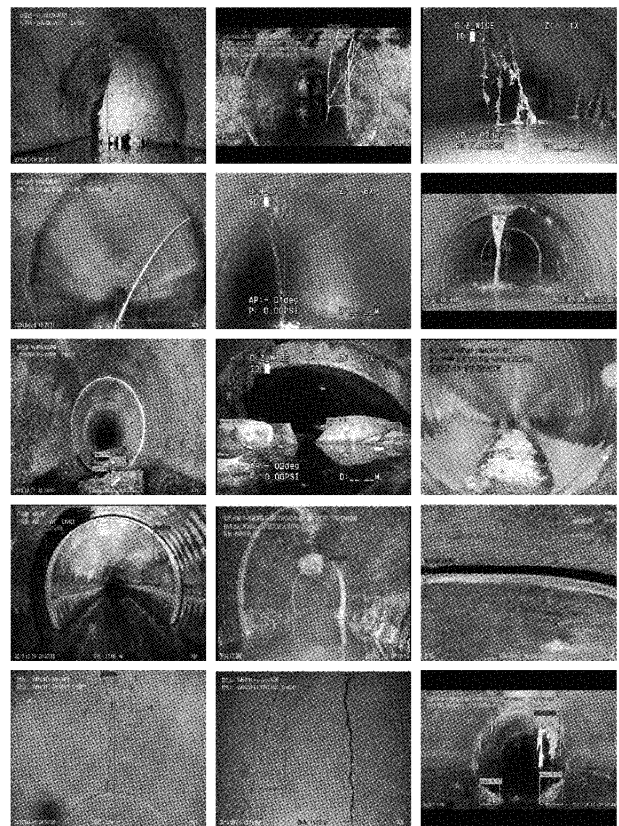


图 12 检测识别部分结果图

3.4 消融实验分析

为了验证本文模型改进方案的有效性,进行 4 组消融实验,实验结果如表所示。方案 a 为原 YOLOX 算法;方案 b 为添加基于改进的空洞空间金字塔模块;方案 c 为添加基于改进的空洞空间金字塔模块和基于权重的特征融合模块;方案 d 为在方案 c 的基础上将边界损失函数 IOU 替换成 GIoU^[25],方案 e 为在方案 c 的基础上将边界损失函数 IOU 替换成 CIoU。实验结果如表 4 所示。从实验结果来看,改进后的 YOLOX 算法,mAP 提升了 0.98%,证明了这两个模块的有效性。采用 GIoU 作为边界损失

函数相比 IOU, mAP 提高了 0.19%, 而采用 CIOU 作为边界损失函数相比 IOU, mAP 提升了 0.64%, 表明将中心点距离和纵横比加入到损失函数中, 能够明显的改善模型检测精度。

表 4 消融实验对比结果 %

方案	AP					mAP
	沉积	树根入侵	渗漏	裂缝	错口	
a	58.26	66.10	69.39	71.63	70.31	67.14
b	58.72	66.43	69.72	71.84	70.61	67.46
c	59.42	67.17	70.38	72.45	71.23	68.12
d	59.58	67.42	70.54	72.63	71.37	68.31
e	60.06	67.82	70.95	73.10	71.85	68.76

3.5 不同算法检测能力对比

为了进一步验证本文算法的高效性, 本文将其与经典的 4 种目标检测算法 Faster R-CNN, YOLOV3 和 YOLOX 在本文的数据集上进行对比, 得到的实验结果如表 5 所示, 从表中可以看出, YOLOv3 的检测效果最差, 相比 YOLOX 算法来说, YOLOX 算法采用了 Mosaic 数据增强的方式, 丰富了地下排水管道缺陷的背景和小目标。而相比两阶段算法 Faster R-CNN, YOLOX 算法通过采用 anchor free 机制以及先进标签匹配策略, 解决了 Faster R-CNN 预设的 anchor 尺寸大小在地下排水管道缺陷检测中, 由于 5 种缺陷尺寸和比例具有多样性, 对检测精度带来一定干扰。而本文所提算法在对 YOLOX 算法进行一定改进的程度上, 使得检测精度得到进一步的提高。

表 5 经典目标检测算法实验对比结果 %

检测算法	AP					mAP
	沉积	树根入侵	渗漏	裂缝	错口	
Faster R-CNN	56.38	64.71	67.28	69.32	66.42	64.82
YOLOV3	54.34	60.27	64.28	67.45	64.62	62.19
YOLOX	58.26	66.10	69.39	71.63	70.31	67.14
本文算法	60.06	67.82	70.95	73.10	71.85	68.76

4 结 论

本文针对地下排水管道缺陷检测识别任务提出了一种改进的 YOLOX 算法, 实现了污水管道缺陷的自动检测识别。为了提高 YOLOX 网络特征融合层的性能, 通过改进的空洞空间金字塔模块以及设计了一种基于权重的特征融合模块对特征融合层进行改进。并且由于地下排水管道缺陷数据样本有限, 本文采用了 GAN 网络对裂缝和错口两种缺陷类型图片进行生成以及各种数据增强方法对数据集进行扩充。实验结果表明, 本文模型在对地下排

水管道缺陷进行检测具有较高的识别能力, 最终 mAP 达到 68.76%。与专业检测人员判别相比, 可以提高检测效率并且降低检测人员由于主观性在判别时造成的影响。然而, 地下排水管道缺陷种类高达 16 种, 本文只对常见的沉积、渗漏、树根入侵、裂缝以及错口这 5 种缺陷类型进行了检测, 同时并不能确定缺陷的在管道中的具体位置信息, 因此, 在今后的研究中, 需要进一步扩大对缺陷种类的检测, 获取到缺陷的具体位置信息, 更好地辅助地下排水管道检测人员进行识别, 实现地下排水管道缺陷检测完全自动化。

参考文献

- [1] 安关峰.《城镇排水管道检测与评估技术规程》CJJ181-2012 实施指南[M]. 北京: 中国建筑工业出版社, 2013.
- [2] SUNIL K, PAUL W. Segmentation of buried concrete pipe images[J]. Automation in Construction, 2006, 15(1): 47-57.
- [3] SU T C, YANG M D. Application of morphological segmentation to leaking defect detection in sewer pipelines[J]. Sensors, 2014, 14(5): 8686-8704.
- [4] HAWARI A, ALAMIN M, ALKADOOR F, et al. Automated defect detection tool for closed circuit television (cctv) inspected sewer pipelines [J]. Automation in Construction, 2018, 89: 99-109.
- [5] 化嫣然, 张卓, 龙赛, 等. 基于改进 YOLO 算法的遥感图像目标检测[J]. 电子测量技术, 2020, 43(24): 87-92.
- [6] KUMAR S S, ABRAHAM D M, JAHANSHAHI M R, et al. Automated defect classification in sewer closed circuit television inspections using deep convolutional neural networks [J]. Automation in Construction, 2018, 91: 273-283.
- [7] WANG M Z, LUO H, CHENG J C P. Towards an automated condition assessment framework of underground sewer pipes based on closed-circuit television (CCTV) images [J]. Tunnelling and Underground Space Technology, 2021, 110: 103840.
- [8] LI D W, XIE Q, YU Z C, et al. Sewer pipedefect detection via deep learning with local and global feature fusion [J]. Automation in Construction, 2021, 129: 103823.
- [9] ZHENG G, LIU S T, WANG F, et al. YOLOX: Exceeding YOLO series in 2021[J]. ArXiv Preprint, 2021, ArXiv: 2107. 08430.
- [10] 程焕新, 蒋泽芹, 程力, 等. 基于改进 YOLOX-S 的安全帽反光衣检测算法[J]. 电子测量技术, 2022, 45(6): 130-135.
- [11] REDMON J, FARHADI A. YOLOv3: An incremental

- improvement [J]. ArXiv Preprint, 2018, ArXiv: 1804.02767.
- [12] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection [J]. ArXiv Preprint, 2020, ArXiv: 2004.10934.
- [13] HE K M, ZHANG X Y, REN S Q, et al. Spatial Pyramid pooling in deep convolutional networks for visual recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37 (9): 1904-1916.
- [14] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 936-944.
- [15] 李晖晖, 周康鹏, 韩太初. 基于 CReLU 和 FPN 改进的 SSD 舰船目标检测 [J]. 仪器仪表学报, 2020, 41 (4): 183-190.
- [16] CHEN L C, PAPANDEOU G, KOKKINOS I, et al. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40 (4): 834-848.
- [17] HU J, LI S, SUN G. Squeeze-and-excitation networks [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [18] ZHANG Q L, YANG Y B. SA-Net: Shuffle attention for deep convolutional neural networks [C]. IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP), 2021: 2235-2239.
- [19] ZHENG Z H, WANG P, LIU W, et al. Distance-IOU loss: faster and better learning for bounding box regression [C]. IAAAI Conference on Artificial Intelligence, 2020: 12993-13000.
- [20] GOODFELLOW I J, ABADIE J P, MIRZA M, et al. Generative adversarial networks [C]. Proceedings of the 27th International Conference on Neural Information Processing Systems, 2014: 2672-2680.
- [21] 谭明奎, 许守恺, 张书海, 等. 深度对抗视觉生成综述 [J]. 中国图象图形学报, 2021, 26 (12): 2751-2766.
- [22] KARRAS T, LAINE S, AITTLA M, et al. Analyzing and improving the image quality of style GAN [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 8107-8116.
- [23] YU F, ZHANG Y, SONG S, et al. LSUN: Construction of a large-scale image dataset using deep learning with humans in the loop [J]. Computer Vision and Pattern Recognition, 2015, DOI: 10.48550/ARXIV.1506.03365.
- [24] HEUSEL M, RAMSAUER H, UNTERTHINER T, et al. GANs trained by a two time-scale update rule converge to a local nash equilibrium [C]. Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS), 2017: 6626-6637.
- [25] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized intersection over union: A metric and a loss for bounding box regression [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019: 658-666.

作者简介

陆绮荣, 博士, 教授, 主要研究方向为信息检测与传输、图像处理等。

E-mail: 330588615@qq.com

丁昕, 硕士生, 主要研究方向为图像处理。

E-mail: 1040489861@qq.com

梁雅雯, 硕士生, 主要研究方向为图像处理。

E-mail: 1213446325@qq.com