



电子测量技术  
*Electronic Measurement Technology*  
ISSN 1002-7300, CN 11-2175/TN

## 《电子测量技术》网络首发论文

题目：基于 TAS-YOLO 的道路表面缺陷检测  
作者：李金涛，周兴林，尹雨飞，敖思铭  
网络首发日期：2024-09-29  
引用格式：李金涛，周兴林，尹雨飞，敖思铭. 基于 TAS-YOLO 的道路表面缺陷检测  
[J/OL]. 电子测量技术. <https://link.cnki.net/urlid/11.2175.TN.20240929.0939.005>



**网络首发：**在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

**出版确认：**纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

# 基于 TAS-YOLO 的道路表面缺陷检测

李金涛 周兴林 尹雨飞 敖思铭  
(武汉科技大学机械自动化学院, 湖北 武汉 430081)

**摘要：**针对路面小缺陷检测准确率低、漏检率和误检率高且均匀分布缺陷类型数据集难以采集问题，本文提出一种基于 YOLOv5s 的 TAS-YOLO 改进网络模型方法。首先，在预测结果阶段采用特定任务的上下文解耦头，通过分离分类和定位任务，增强定位检测框的精度；其次，通过 FPN 结构将 5 个尺度的特征图输入解耦头进行预测，增强小目标的多尺度特征信息；最后，使用滑动窗口损失函数优化 YOLOv5，提高难分类样本的检测准确率，且模型收敛效果更好。实验结果表明，TAS-YOLO 算法提升了各类缺陷的平均检测精度，mAP50 值达到 91.4%，FPS 值达到 126，较 YOLOv7l、YOLOv8s、YOLOv9c-gelan 和 Efficientdet 等主流检测算法提高了检测的精度和效率。

**关键词：**路面检测；上下文解耦；多尺度特征；损失函数

中图分类号：TP391.4；U418；TN207 文献标志码：A 国家标准学科分类代码：520.20

## Road surface defect detection based on TAS-YOLO

Li Jintao, Zhou Xinglin, Yin Yufei, Ao Siming

(School of Mechanical Automation, Wuhan University of Science and Technology, Wuhan 430081, China)

**Abstract:** This paper proposes an improved TAS-YOLO network model method based on YOLOv5s to address the issues of low accuracy, high missed and false detection rates, and difficulty in collecting uniformly distributed defect types datasets for detecting small road surface defects. Firstly, in the prediction result stage, a context decoupling head for a specific task is used to enhance the accuracy of the localization detection box by separating classification and localization tasks; Secondly, by using the FPN structure to input feature maps of 5 scales into the decoupling head for prediction, the multi-scale feature information of small targets is enhanced; Finally, use the Silde Loss function to optimize YOLOv5 and improve the detection accuracy of difficult to classify samples. The experimental results showed that TAS-YOLO algorithm improved the average detection accuracy of various defects, with mAP50 reaching 91.4% and FPS reaching 126, which improved the detection accuracy and efficiency compared with mainstream detection algorithms such as YOLOv7l, YOLOv8s, YOLOv9c-gelan and Efficientdet.

**Keywords:** pavement detection; context decoupling; multiscale features; loss functions

## 0 引言

随着中国国内公路网络结构的建设，道路养护工作从简单的抢修模式向全面养护模式过渡。已有研究表明，每投入 1 元进行道路维护，就相当于路面使用费用节省 3 元<sup>[1]</sup>。及时的检测维护可以有效防止损伤扩大，节约路面修复支出。传统方法通常耗时低效，妨碍交通且资源消耗大，因此研究人工智能对路面损伤检测有着重要的意义<sup>[2]</sup>。目前人工智能的检测路面缺陷方法主要为两种：一种是手动提取特征的机器学习方法，一种是基于机器视觉的深度学习方法。

在手动提取特征的机器学习方法中，Prasanna 等<sup>[3]</sup>提出一种名为 STRUM 裂缝检测算法的新算法，用于检测桥梁中的裂缝信息；Chen 等<sup>[4]</sup>将彩色裂缝图转化成灰度图，边缘化后用局部定向模式(LDP)算法进行特征提取后，采用 FHRCNN 用于路面裂缝检测；Sun 等<sup>[5]</sup>利用梯度直方图(HOG)提取裂缝特征，然后使用贝叶斯优化设置超参数后的 SVM 算法进行混凝土桥梁裂缝检测。综上所述，机器学习算法虽然可以考虑到缺陷外观的影响因素，但是对数据集和训练资源的要求较高，并且依赖研究人员对损伤特征的了解，因此手动提取特征的机器学习算法的通用性较差。

基于机器视觉的深度学习因其优于传统机器学习

和其他算法的检测能力成为备受关注的研究领域<sup>[1]</sup>。赵朗月等<sup>[6]</sup>阐述表面缺陷检测在工业发展的必要性,分析了传统图像处理方式、机器学习和深度学习检测表面缺陷的优缺点;高艺平等<sup>[7]</sup>对深度智能视觉表面缺陷检测方法中分类、定位、分割三种方法进行描述,总结叙述工业场景中表面缺陷检测应用的重要性。目前基于深度学习的目标检测算法分为两类,一种是基于目标区域的两阶段目标检测算法,例如 RCNN<sup>[8]</sup>、Fast RCNN<sup>[9]</sup>、Faster RCNN<sup>[10]</sup>、Mask RCNN<sup>[11]</sup>、Cascade RCNN<sup>[12]</sup>和 SPPNet<sup>[13]</sup>等,还有一种是基于回归的一阶段目标检测算法,如 SSD<sup>[14]</sup>、YOLO<sup>[15-17]</sup>、RetinaNet<sup>[18]</sup>和 EfficientDet<sup>[19]</sup>等,这其中的代表算法为 YOLO 系列。Ren 等<sup>[1]</sup>在 YOLOv5 的基础上使用广义特征金字塔网络结构(GFPN)替换 SPPF,添加检测层、替换损失函数和原解耦头,有效提升模型街景路面损伤检测精度;Zhao 等<sup>[20]</sup>为了准确检测钢材表面缺陷提出 RDD-YOLO 模型,利用 Res2Net 替换 YOLOv5 主干,设计双特征金字塔网络结构(DFPN)丰富特征信息,采用任务分离解耦头,提高了对钢材表面缺陷的检测精度;张艳君等<sup>[21]</sup>针对路面缺陷检测精度不高的问题,将重参数化、高效聚合、辅助训练等模块和卷积注意力机制添加在 YOLOv7 的网络中,提升检测精度;罗晖等<sup>[22]</sup>使用深度可分离卷积、改进损失函数和迁移学习的方式对 YOLOv4 网络改进,并进行数据增强,有效解决公路路面病害存在的类别多、尺度变化大及样本数据集小导致的病害检测困难的问题;程文党等<sup>[23]</sup>将 ECA 注意力机制融入 TransUNet 网络,有效的对沥青路面的病害进行检测;陈建瑜等<sup>[24]</sup>针对路面裂缝检测耗时耗力,成本高等问题,提出 YOLOv5-Crack 网络,该网络引入坐标注意力机制、ESPP 结构、GSConv 和 SIOU 损失函数,解决了路面裂缝检测的弊端;Roy 等<sup>[25]</sup>为提高复杂环境下路面损伤的检测精度,在 YOLOv5 中加入 DenseNet 模块、CBAM 注意力机制和特征融合层,使用 Swin-Transformer 检测头后,模型的通用性能提高,优于当前其他模型;孙建诚等<sup>[26]</sup>用 GhostConv 替换 YOLOv5 中的卷积,引入坐标注意力模块,提高在复杂情况下的路面裂缝检测准确率;景峰等<sup>[27]</sup>针对路面裂缝检测中漏检和定位不准的情况,在 YOLOv5 中使用任务空间分离检测头并引入注意力机制,有效改善问题;杨振等<sup>[28]</sup>提出一种名为 YOLO-EH 的特征网络增强算法,有效提高了同一路面上病害的识别率;王梦梦等<sup>[29]</sup>对 YOLOv5 改进,使用 CBAM 替换主干中的 C3,将损失函数替换为 EIoU,增加检测层,改进后的网络能有效检出路面损伤。上述检测路面缺陷方法已经取得进步,但存在一些问题需要改进:路面缺陷的数据集往往存在数量较少、易混淆和小尺寸样本,导致预测框的位置不准确且容易产生误检漏检情况,使得检测准确率不高。

针对上述问题,本文以 YOLOv5s 的主体框架为基础,采用 FPN 结构增加检测层提取更加丰富的特征信息,以加强小目标的特征信息;通过滑动窗口损失函数(SildeLoss)<sup>[30]</sup>优化模型,提高难分类样本的检测准确率,加速模型收敛速度;使用特定任务的上下文解耦头<sup>[31]</sup>(Task-Specific context decoupling head, TSCODE)来分离分类和定位任务,有助于预测目标框定位更加准确和稳定。

## 1 YOLOv5 算法理论

YOLOv5 作为一种主流目标检测算法,具有较快的检测速度和较高的检测准确率,工业应用表现良好并在路面损伤自动检测方面展示优势。YOLOv5 目前官方推出有 5 种大小的算法模型,其主要区别在于其中卷积和提取特征模块的重复次数。YOLOv5s 参数量较少且检测精度较高,为减少计算资源,本文选用 YOLOv5s 作为基线模型进行实验和下一步改进,下面介绍模型的主要组成部分。

该模型包括四个部分:输入(input)、主干(backbone)、颈部(neck)和头部(head)。输入端负责将图片自适应缩放至 640×640 大小,并进行锚框检测和图像增强以提升模型性能。主干模块采用 CSPDarknet,结合 Focus 和 CSP 结构,提高特征提取效率和准确性。颈部通过特征金字塔网络(FPN)、路径聚合网络(PAN)结构和快速空间金字塔池化 SPPF 进一步丰富特征图的语义信息,增强目标检测效果。头部使用 CIoU 损失函数和非极大值抑制算法,结合三个检测头输出结果,完成目标分类和定位。

## 2 TAS-YOLO 模型

针对路面缺陷的检测过程中存在病害样本数不平衡识别精度低、小目标识别难度大和漏检误检率高等问题,本文提出 TAS-YOLO 模型。首先,使用特定于任务的上下文解耦头替换 YOLOv5s 的检测头,使用更多的上下文语义信息加强头部的分类和定位功能;其次,采用 FPN 结构在网络的颈部提取特征图像,增加尺寸为 160×160 和 10×10 的特征图输入检测头,加强网络特征提取能力,丰富小目标的语义信息,提升小目标检测准确率;最后,使用 SildeLoss 优化对 YOLOv5 的损失函数,SildeLoss 可以自动学习正负样本的阈值,将更高的权重分配给困难分类的缺陷,用来提高由于样本数量不均导致的检测准确率低 YOLOv5s 网络结构和 TAS-YOLO 网络结构如图 1 所示。

### 2.1 改进解耦头

YOLOv5 网络中使用检测头进行分类和预测,输出检测结果。常用的检测头有 CoupledHead, DecoupledHead<sup>[1,20]</sup>, DynamicHead<sup>[32]</sup>和 TSCODE<sup>[31]</sup>等。CoupledHead 耦合头采用相同的特征图执行分类和定位任务,然而分类和定位所需要的信息侧重不同。DecoupledHead 解耦头在 CoupledHead 的基础上添加分类分支网络,将分类和定位任务分离,进一步提高了检测精度。然而该解耦头简单的将分类和定位的检测结果融合,难以适应复杂路面缺陷的检测任务;DynamicHead 检测头在耦合头的基础上添加了多种多头自注意力机制,达到提高精度的目的,但由于检测任务耦合,不能有效地解决分类和定位所需信息的冲突,检测精度较低。本文提出使用 TSCODE 解耦头,该解耦头针对不同任务的需要生成包含不同侧重信息的特征图来预测结果,为验证该解耦头的有效性,进行了对比实验,实验结果如表 1。



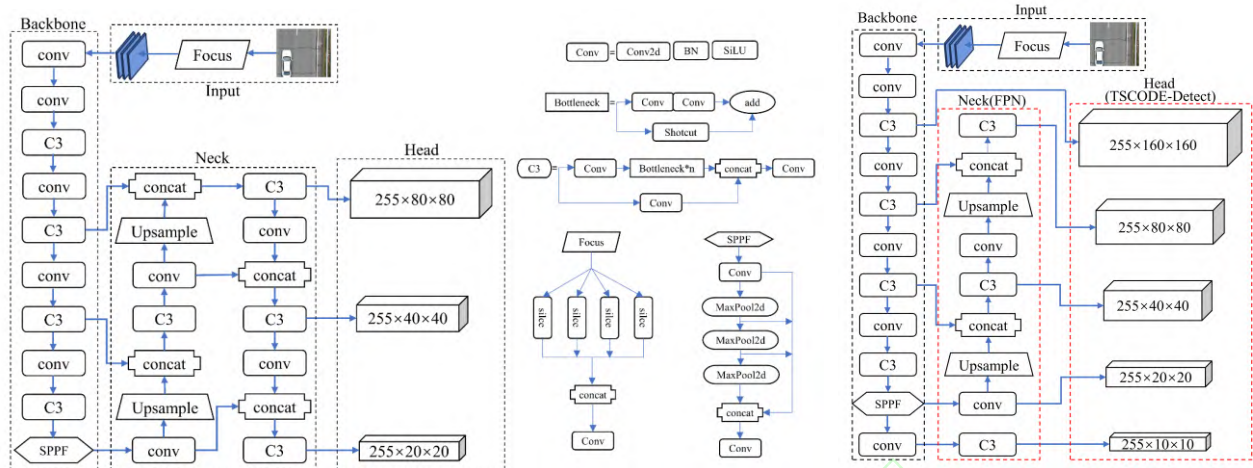


图1 YOLOv5s 网络结构(左)和 TAS-YOLO 网络结构(右)

Figure 1 YOLOv5s network structure (left) and TAS-YOLO network structure (right)

表 1 各检测头预测结果

Table 1 Prediction results of each decoupling head

| 检测头名称         | AP/% |      |      |      |      |      |
|---------------|------|------|------|------|------|------|
|               | D1   | D2   | D3   | D4   | D5   | D6   |
| DetectHead    | 94.0 | 87.0 | 89.2 | 71.5 | 97.4 | 92.7 |
| DecoupledHead | 93.2 | 85.2 | 89.3 | 70.3 | 96.2 | 92.7 |
| DynamicHead   | 79.6 | 82.3 | 84.2 | 53.4 | 95.2 | 86.1 |
| TSCODEHead    | 96.8 | 88.8 | 94.4 | 73.3 | 97.4 | 93.4 |

路面缺陷常存在缺陷样本数不平衡、小目标等问题,根据表1的结果对比可以看出原解耦头、DecoupledHead 和 DynamicHead 对于路面各类型复杂缺陷检测精确度均低于 TSCODEHead,均不能达到检测需求。因此使用 TSCODE 解耦头,该解耦头针对分类生成空间粗糙但语义强的特征编码,对于定位生成了包含更多边缘信息的高分辨率特征图,解决了多个任务需求和多种特征信息之间的冲突,来更好的提高检测精度。

图2为 TSCODE 解耦头的内部结构,该解耦头包含语义上下文编码(SCE)和细节保留编码(DPE)。具体细节如图2, SCE 模块对  $P_{l+1}$  层使用一个  $3 \times 3$  的卷积层进行2倍下采样然后与  $P_{l+1}$  层进行连接,生成用于分类的特征图。DPE 模块首先对  $P_l$  层进行2倍上采样与  $P_{l-1}$  层进行融合,然后对融合特征图使用一个  $3 \times 3$  的卷积层进行2倍下采样使其分辨率和  $P_l$  层一样,最后对  $P_{l+1}$  层特征进行2倍上采样,将经过处理的特征图进行融合得到用于定位的特征图。通过 SCE 和 DPE 两种结构将两个任务的语义信息上下文解耦,为分类带来更丰富的语义信息为定位带来更多的边缘信息,有效提升检测准确率。

该解耦头接收来自5个不同尺度的检测层预测结果, YOLOv5s 基础模型中有三个尺度的检测层,将输入图像经行32、16、8下采样,接着进行卷积、激活函数和特征连接得到  $80 \times 80$ 、 $40 \times 40$ 、 $20 \times 20$  三个尺寸的有效特征层,最后使用特征层对目标进行预测进行。本文采用 FPN 结构并增加检测层,以适应路面缺陷中大小分布不均匀的情况。改进后的网络将特征图通过快速特征金字塔池化后继续进行下采样,特征图继续缩小至

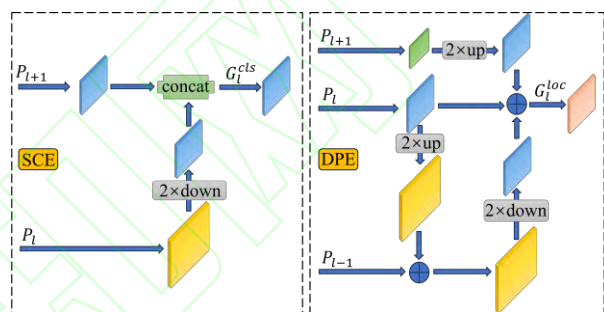


图2 语义上下文编码和细节保留编码

Figure 2 Semantic context encoding and detail retention encoding

$10 \times 10$ , 将骨干网络中第二层尺寸为  $160 \times 160$  的特征图输入头部进行预测,结构如图1右所示。这种新结构在减少网络深度的情况下,将检测层数量从3个增加到5个,提取更丰富的语义特征,提高小目标的检测精度。

为验证该结构的可行性,进行了对比实验,如表2所示分别使用①AFPN-YOLOv5、②GFPN-YOLOv5<sup>[1]</sup>、③DFPN-YOLOv5<sup>[20]</sup>和④DenseSPH-YOLOv5<sup>[25]</sup>网络结构对输入图像进行特征图的提取,4种网络结构是对YOLO的颈部进行改进,将提取出来的5个尺度的特征图输入 TSCODE 解耦头进行预测,与本文特征提取结构加 TSCODE 的预测结果进行对比。由表2可知,本文所提结构在平均精度和模型大小方面均优于其他4种结构,表明该结构具有良好的特征提取能力,且参数量较小。

表 2 各结构提取特征结果

Table 2 Feature extraction results of each structure

| 结构名称 | P%   | R%   | mAP50% | 参数量/M |
|------|------|------|--------|-------|
| ①    | 84.4 | 82.7 | 85.0   | 35.2  |
| ②    | 89.1 | 82.7 | 90.4   | 30.7  |
| ③    | 87.4 | 83.1 | 89.1   | 31.5  |
| ④    | 86.6 | 83.3 | 88.4   | 34.8  |
| ours | 89.1 | 84.6 | 90.7   | 18.8  |

## 2.2 改进损失函数

YOLOv5s模型的损失函数由三部分组成分别是置

信度损失  $L_{conf}$ 、分类损失  $L_{cls}$  和边界框回归损失  $L_{box}$ 。在基础YOLOv5网络中,置信度损失  $L_{conf}$  和  $L_{cls}$  分类损失是通过带有Sigmoid的二元交叉熵函数(BCEWithLogits Loss)计算得来的,三种损失具体计算公式如下。  
置信度损失:

$$L_{conf} = -\sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^0 [\hat{C}_i^j \ln C_i^j + (1 - \hat{C}_i^j) \ln(1 - \hat{C}_i^j)] - \lambda_n \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^n [\hat{C}_i^j \ln C_i^j + (1 - \hat{C}_i^j) \ln(1 - \hat{C}_i^j)] \quad (1)$$

式中:  $S^2$  为特征图的网格数;  $B$  为锚框数;  $\lambda_n$  为权重系数;  $I_{ij}^0$ 、 $I_{ij}^n$  中  $n$  为训练 epoch 的个数; 若第  $i$  个网格的第  $j$  个锚框处有目标, 则  $I_{ij}^0$ 、 $I_{ij}^n$  分别取 1 与 0, 若没有目标则  $I_{ij}^0$ 、 $I_{ij}^n$  分别取 0 与 1;  $\hat{C}_i^j$ 、 $C_i^j$  分别为预测框和实际框的置信度。分类损失:

$$L_{cls} = -\sum_{i=0}^{S^2} I_{ij}^0 \sum_{c \in \text{classes}} \left[ \hat{P}_i^j \ln P_i^j + (1 - \hat{P}_i^j) \ln(1 - \hat{P}_i^j) \right] \quad (2)$$

式中:  $\hat{P}_i^j$ 、 $P_i^j$  分别为预测框和实际框的类别概率。边界框回归损失:

$$L_{box} = 1 - R_{IoU}(M, N) + \frac{\rho^2(M_{ctr}, N_{ctr})}{m^2} + \alpha v \quad (3)$$

$$\alpha = \frac{v}{[1 - R_{IoU}(M, N)] + v} \quad (4)$$

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

式中:  $R_{IoU}(M, N)$  表示为预测框和实际框的IoU值;  $\rho^2(M_{ctr}, N_{ctr})$  表示预测框中心点和实际框中心点的欧氏距离;  $m$  表示预测框和实际框之间最小外接矩形的对角线的距离;  $w$ 、 $h$  分别取值为预测框的长和宽,  $w_{gt}$ 、 $h_{gt}$  分别为实际框的长和宽。

在路面缺陷数据集当中,有些缺陷的数据量小、缺陷特征易混淆,这些样本称为困难样本,如图3所示,D4类型仅占总缺陷数的2%;在基线网络实验中,D4类型缺陷的检测精度是最低的。

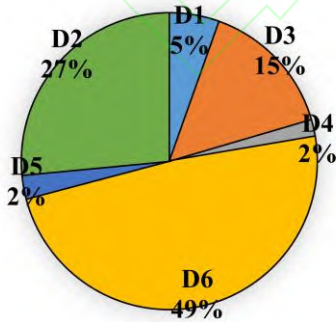


图3 数据集缺陷类型分布

Figure 3 Distribution of defect types in data set

为平衡数据集当中的困难样本和容易样本,提高检测精度,本文提出了滑动窗口损失函数(SlideLoss)<sup>[25]</sup>来优化二元交叉熵函数(BCEWithLogitsLoss)。该损失函数以所有边界框的IoU值的平均值作为阈值,小于阈值的被视为负样本,大于阈值的被视为正样本。该损失函数通

过调节损失值的权重分配,将注意力更多放在难分类、错误分类的样本。采用加权函数计算权重强调困难样本,加权函数为:

$$\gamma = \begin{cases} 1 & L \leq \mu - 0.1 \\ e^{1-\mu} & \mu - 0.1 < L < \mu \\ e^{1-L} & L \geq \mu \end{cases} \quad (6)$$

$$L' = \gamma L \quad (7)$$

式中  $L$  表示为置信度损失或分类损失,  $\gamma$  表示加权函数所计算出来的损失权重,  $L'$  表示加权得来的损失值。

### 3.实验结果与分析

#### 3.1 数据集来源以及训练环境配置

本文采用 UVA-PDD2023 数据集<sup>[33]</sup>,该数据集有 2440 张图片包含 10006 个实例标签,均使用无人机在垂直角度拍摄中国道路,包括不同建设标准下的高速公路、省道和县公路。缺陷主要包括纵向裂缝(Longitudinal crack)、横向裂缝(Transverse crack)、斜缝(Oblique crack)、鳄鱼裂缝(Alligator crack)、修补(Repair)和坑洞(Pothole)六种类型,各类型实例标签的个数如表3所示。将数据集按比例 9:1 随机划分为训练集和测试集。训练总共设置 300 个轮次,具体实验环境和参数设置如表4所示。

表3 数据集构成

Table 3 Data set composition

| 类标识符 | 类型名称                     | 实例个数 |
|------|--------------------------|------|
| D1   | 鳄鱼裂缝(Alligator crack)    | 547  |
| D2   | 纵向裂缝(Longitudinal crack) | 2661 |
| D3   | 斜缝(Oblique crack)        | 1506 |
| D4   | 坑洞(Pothole)              | 184  |
| D5   | 修补(Repair)               | 247  |
| D6   | 横向裂缝(Transverse crack)   | 4862 |

表4 实验环境及部分参数

Table 4 Experimental environment and some parameters

| 名称   | 配置参数  |
|------|---|
| CPU  | Intel(R) Core (TM) i5-12400F<br>@2.50GHz    |
| GPU  | NVIDIA GeForce RTX4060Ti desktop<br>16G     |
| RAM  | 32G   |
| 操作环境 | Windows 10                                  |
| 软件平台 | PyTorch 1.10.0<br>PyThon 3.8<br>CUDA 11.3.0 |

#### 3.2 评价指标

为了评估TAS-YOLO网络模型的检测效果,本文采用精确度P(Precision)、召回率R(Recall)、F1分数、平均精确度(Average Precision, AP)和多分类平均精确度均值(mean Average Precision, mAP)对模型进行评估。分析训练结果后的混淆矩阵可以得到真阳性(TP)、假阳性

(FP), 假阴性(FN)和真阴性(TN)的值, 在分类过程中,  $IoU \geq 0.5$  的为  $TP$ ,  $IoU \leq 0.5$  为  $FP$ 。精确度  $P$  的计算公式为:

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

召回率  $R$  的计算公式为:

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

$F1$  为  $P$  和  $R$  的算术平均值, 计算公式为:

$$F1 = 2 \times \left( \frac{PR}{P + R} \right) \quad (10)$$

$AP$  为  $P$ - $R$  曲线下包围的面积, 计算公式为:

$$AP = \int_0^1 P(R) dR \quad (11)$$

本文使用  $mAP50$  和  $mAP50-95$  作为评价指标,  $mAP50$  是指将  $IoU$  的阈值设置为 0.5, 计算出来每个类别  $AP$  的平均值;  $mAP50-95$  是指将  $IoU$  的阈值取为 0.5: 0.05: 0.95, 计算出来的平均  $mAP$  值。  $mAP50$  的计算公式如下:

$$mAP50 = \frac{\sum_{i=1}^N P(Y | IoU \geq 0.5)}{N} \quad (12)$$

式中:  $N$  为数据集样本的类别数。

### 3.3 消融实验及收敛性结果分析

为验证 TSCODE Structure 和 Silde Loss 对模型改进的有效性, 本文进行消融实验。表中 TSCODE Structure 表示 TSCODE 解耦头加上多尺度网络提取特征图结构, SildeLoss 表示对 YOLOv5 采用滑动窗口损失函数进行优化, 实验结果如表 5。

从表 5 可得, 在单独加入 TSCODE Structure 后, 对比原网络  $mAP50$  值提高 3.2%,  $mAP50-95$  值提高

3.5%, 表明分离分类和定位任务的解耦头有效提高了检测精度, 验证了特征提取结构能有效提高模型的多尺度特征提取和融合能力; 对 D1(鳄鱼裂缝)、D2(纵向裂缝)、D3(斜缝)、D4(坑洞)和 D6(横向裂缝)六个类型缺陷的平均检测精度均有提升, 表明该解耦头有效增强模型的定位和分类信息, 提高了模型整体性能; D3 类缺陷的平均检测精度提升 5.2%, 该类型缺陷中包含有小目标, 说明该结构有效的丰富了小目标的语义信息。单独使用 Silde 损失函数后, D4 缺陷的平均检测精度相较于原网络提高 7.2%, 这类缺陷的特点是数量较少且易混淆, 检测时难以精确识别。优化损失函数在不对样本数据进行数据增强的情况下, 有效改善了困难样本检测率低的情况。在加入两种改进策略后, D4 缺陷的平均检测精度进一步提高至 78.7%, 加入优化损失函数对比单独使用 TSCODE Structure, D1 类型的检测精度下降了 2.2%, 其余 5 类的检测精确度都有提升, 这对检测多种不同类型缺陷更有利, 表明优化损失函数进一步提高了模型检测复杂路面缺陷的能力。加入混合改进策略后, 各类型的平均检测精度都有提升; 对于模型的  $mAP50$  值和  $mAP50-95$  值达到最高, 相比较于基础模型分别提高 3.9%和 4.6%,  $F1$  分数提高了 3%。

图 4 为 TAS-YOLO 和 YOLOv5s 的检测框损失、分类损失和置信度损失的变化图。在分类损失曲线图中, 两个模型都具有较好的收敛速度且趋近于 0; 在检测框损失曲线中, 两个模型在 15 轮训练后开始收敛, 改进模型在 25 轮训练后, 损失值下降幅度比原模型大, 收敛速度加快, 最终收敛效果比原模型好; 在置信度损失曲线图中, 改进模型在 145 轮时收敛速度开始比原模型快, 最终损失值收敛到 0.00107。通过对比损失值的变化, 改进模型的总体收敛效果好于原模型, 说明改进模型的优越性, 能够很好应用于道路缺陷类型的检测中。

表 5 改进模型消融对比实验  
Table 5 Improved model ablation comparison experiment

| TSCODE<br>Structure | Silde<br>Loss | AP%  |      |      |      |      |      | mAP<br>50/% | mAP<br>50-95/% | F1<br>Score/% |
|---------------------|---------------|------|------|------|------|------|------|-------------|----------------|---------------|
|                     |               | D1   | D2   | D3   | D4   | D5   | D6   |             |                |               |
|                     |               | 94.0 | 87.0 | 89.2 | 71.5 | 97.4 | 92.7 | 87.5        | 55.7           | 85            |
| ✓                   |               | 96.8 | 88.8 | 94.4 | 73.3 | 97.4 | 93.4 | 90.7        | 59.2           | 87            |
|                     | ✓             | 96.0 | 88.6 | 89.2 | 77.8 | 97.4 | 92.6 | 90.3        | 57.6           | 83            |
| ✓                   | ✓             | 94.6 | 90.8 | 91.6 | 78.7 | 98.7 | 94.1 | 91.4        | 60.3           | 88            |

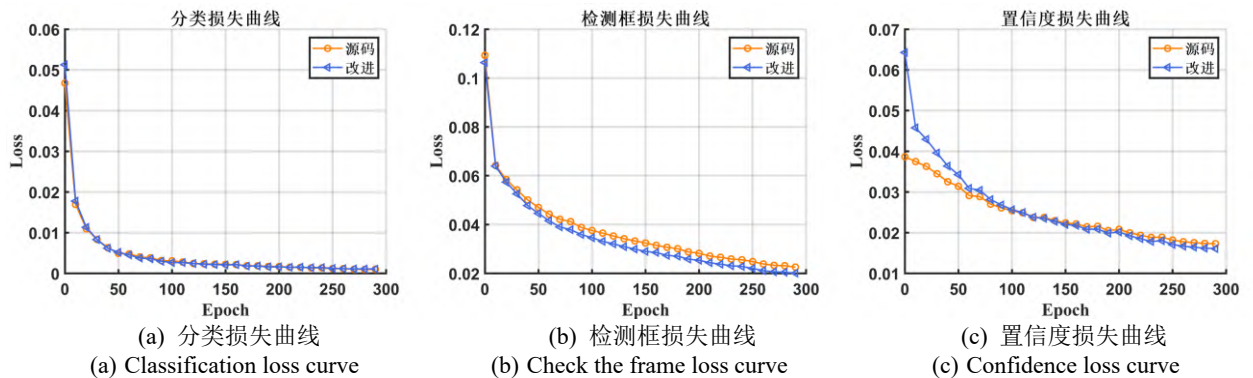


图 4 TAS-YOLO 和 YOLOv5s 损失曲线  
Figure. 4 Loss curves of TAS-YOLO and YOLOv5s



3.4 模型检测结果分析

为进一步验证 TAS-YOLO 网络模型的有效性，随机选择 3 张图片进行检测，对比检测结果图如图 5 所示。观察图(a1)、(a2)，原模型在检测时，漏检了 D5(修补)缺陷，改进后的模型成功检出，并且准确率到达 90%；对比图(b1)、(b2)，TAS-YOLO 改善了 D3(斜缝)缺陷的漏检情况且检测框包围目标的范围更准确，大幅提高 D6(横向裂缝)和 D2(纵向裂缝)的检测准确率，分别提升了 45%和 42%；对比(c1)、(c2)图，在原模型的检测图中出现对 D1(鳄鱼裂缝)和 D4(坑洞)误检情况，且两类缺陷检测框的重合，改进模型中准确检测出两类缺陷，并且由于使用特定任务的上下文解耦头，消除了重叠检测框，预测框更符合实际目标框。根据检测图的对

比结果分析，改进后的模型对漏检、误检的情况有明显改善，检出率更高，有效提高了检测准确率。

3.5 不同网络的对比实验分析

如表 6 所示 TAS-YOLO 网络模型和目前主流的目标检测算法进行对比，结果表明，相比较于其他算法，该算法提升了各类缺陷的平均检测精度，比 YOLOv8s 提高 12.8%；对比 YOLOv7l 算法提高 15.9%；对比 YOLOv9c-gelan 算法提高 6.1%；对比 Centernet 和 Efficientdet 算法，平均精度提高 26%和 41%。检测速度对比 YOLOv7l、YOLOv9c-gelan、Efficientdet 算法有明显提高，验证了 TAS-YOLO 网络模型具有更好的鲁棒性和检测性能。

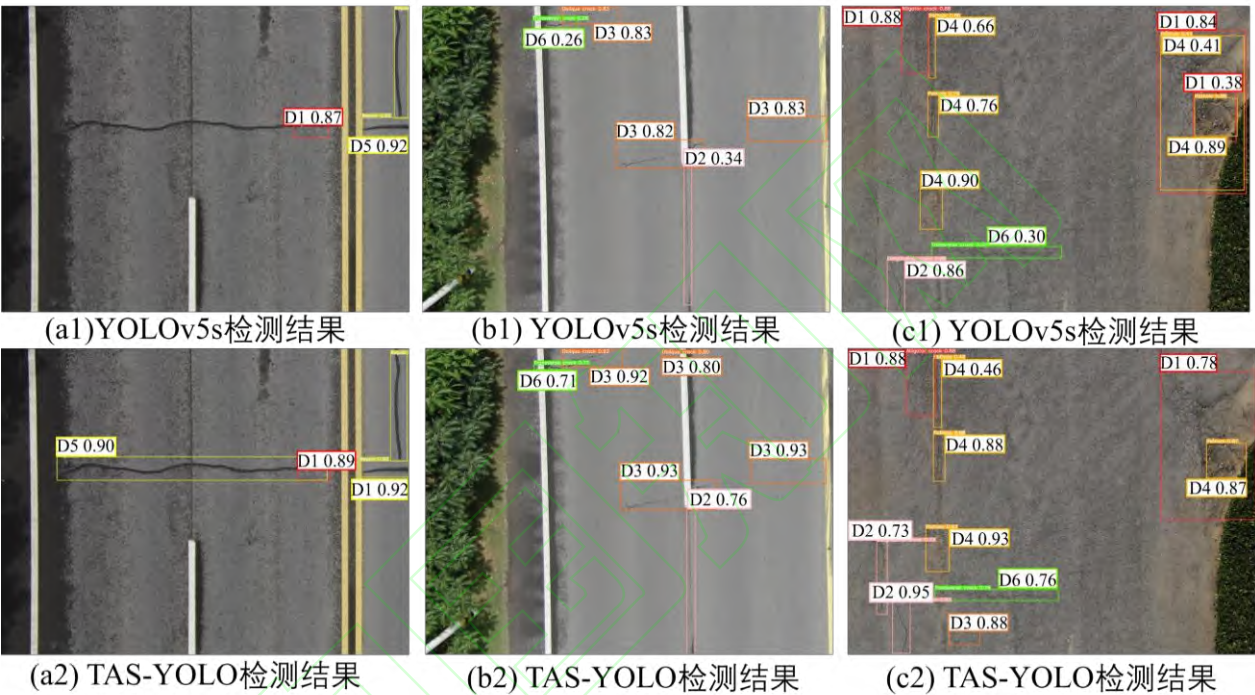


图 5 YOLOv5s 和 TAS-YOLO 检测结果  
Figure 5 Test results of YOLOv5s and TAS-YOLO

表 6 不同方法检测的实验结果对比

Table 6 Comparison of experimental results detected by different methods

| 模型名称          | AP%  |      |      |      |      |      | mAP  | F1      | FPS/        |
|---------------|------|------|------|------|------|------|------|---------|-------------|
|               | D1   | D2   | D3   | D4   | D5   | D6   | 50/% | Score/% | (frame·s-1) |
| YOLOv8s       | 73.9 | 87.8 | 77.1 | 62.0 | 85.0 | 85.5 | 78.6 | 77      | 95          |
| YOLOv7l       | 84.4 | 76.8 | 63.7 | 50.9 | 97.2 | 79.0 | 75.3 | 71      | 61          |
| YOLOv9c-gelan | 95.3 | 86.4 | 82.5 | 57.3 | 99.1 | 91.2 | 85.3 | 83      | 60          |
| Centernet     | 67.0 | 73.7 | 57.1 | 45.9 | 77.5 | 70.3 | 65.2 | 54      | 97          |
| Efficientdet  | 65.9 | 31.3 | 57.6 | 25.2 | 57.4 | 63.0 | 50.2 | 47      | 33          |
| Ours          | 94.6 | 90.8 | 91.6 | 78.7 | 98.7 | 94.1 | 91.4 | 88      | 126         |

4 结语

本文针对路面缺陷数据集建立一种 TAS-YOLO 网

络模型，该模型基于 YOLOv5s 优化了网络损失函数、增加了特征提取层，以及使用了特定于任务的上下文解耦头。实验结果表明，优化后的损失函数，收敛速度更快，对困难样本的平均识别精度大大提高；增加特征提

取层后,丰富了小目标的特征信息,提高了识别率;采用的特定于任务的上下文解耦头对比于原检测头,有效改善了目标框定位不准而导致的误检和漏检情况,提高了检测准确度。综上所述,本文所提出的 TAS-YOLO 网络模型能有效的实现路面缺陷的准确检测,对道路智慧养护决策提供数据参考。

#### 参考文献

- [1] REN M, ZHANG X, CHEN X, et al. YOLOv5s-M: A deep learning network model for road pavement damage detection from urban street-view imagery[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2023,120: 103335.
- [2] 张用川, 牟凤云, 陈建坤, 等. 基于改进 YOLOv5 算法的道路损伤检测[J]. *电子测量技术*, 2023, 46(4):161-168.  
ZHANG Y, MOU F, CHEN J, et al. Road Injury Detection based on improved YOLOv5 Algorithm [J]. *Electronic Measurement Technology*, 2019,46(4):161-168.
- [3] PRASANNA P, DANA K J, GUCUNSKI N, et al. Automated crack detection on concrete bridges[J]. *IEEE Transactions on automation science and engineering*, 2014, 13(2):591-599.
- [4] CHEN J, SU M, CAO R, et al. A self organizing map optimization based image recognition and processing model for bridge crack inspection[J]. *Automation in Construction*, 2017, 73:58-66.
- [5] SUN Z, CAETANO E, PEREIRA S, et al. Employing histogram of oriented gradient to enhance concrete crack detection performance with classification algorithm and Bayesian optimization[J]. *Engineering Failure Analysis*, 2023, 150:107351.
- [6] 赵朗月, 吴一全. 基于机器视觉的表面缺陷检测方法研究进展[J]. *仪器仪表学报*, 2023, 43(1):198-219.  
ZHAO L, WU Y. Research progress of surface defect detection methods based on machine vision [J]. *Chinese Journal of Scientific Instrument*, 2023, 43(1):198-219
- [7] 高艺平, 王浩, 李新宇, 等. 基于深度智能视觉的表面缺陷检测研究进展[J]. *工业工程*, 2024, 27(2): 27.  
GAO Y, WANG H, LI X, et al. Research progress of surface defect detection based on deep intelligent vision [J]. *Industrial Engineering Journal*, 2024, 27(2):27.
- [8] GIRSHICK R, DONAHUE J, DARRELL T, et al. R-ich feature hierarchies for accurate object detection and semantic segmentation[C]//*Proceedings of the IEEE conference on computer vision and pattern recognition*. 2014: 580-587.
- [9] GIRSHICK R. Fast R-CNN[C]//*IEEE International Conference on Computer Vision*, 2015:1440-1448.
- [10] HASCOET T, ZHANG Y, PERSCH A, et al. FasterRCNN monitoring of road damages: Competition and deployment[C]//*2020 IEEE International Conference on Big Data (Big Data)*. IEEE, 2020:5545-5552.
- [11] SINGH J, SHEKHAR S. Road damage detection and classification in smartphone captured images using mask R-CNN[J]. *arxiv preprint arxiv: 1811.04535*, 2018.
- [12] PEI Z, LIN R, ZHANG X, et al. CFM: A consistency filtering mechanism for road damage detection[C]//*2020 IEEE International Conference on Big Data (Big Data)*. IEEE, 2020:5584-5591.
- [13] HE K, ZHANG X, REN S, 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] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]//*14th European Conference on Computer Vision*, Amsterdam, The Netherlands, Springer, 2016:21-37.
- [15] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]//*IEEE Conference on Computer Vision and Pattern Recognition*, USA, 2016:779-788.
- [16] REDMON J, FARHADI A. Yolov3: An incremental improvement[J]. *ArXiv*, 2018(8):1804.02767.
- [17] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: Optimal speed and accuracy of object detection[J]. *arXiv preprint arXiv: 2004.10934*, 2020.
- [18] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//*Proceedings of the IEEE international conference on computer vision*. 2017:2980-2988.
- [19] NADDAF-SH S, NADDAF-SH M M, KASHANI A R, et al. An efficient and scalable deep learning approach for road damage detection[C]//*2020 IEEE International Conference on Big Data (Big Data)*. IEEE, 2020:5602-5608.
- [20] ZHAO C, SHU X, YAN X, et al. RDD-YOLO: A modified YOLO for detection of steel surface defects[J]. *Measurement*, 2023, 214:112776.



- [21] 张艳君, 沈平, 郭安辉. 融合 CBAM- YOLOv7 模型的路面缺陷智能检测方法研究[J]. 重庆理工大学学报(自然科学), 2023, 37(11):213-220.
- ZHANG Y, SHEN P, GUO A. Research on intelligent pavement defect detection method based on CBAM-YOLOv7 model [J]. Journal of Chongqing University of Technology (Natural Science), 2023, 37(11):213-220
- [22] 罗晖, 贾晨, 李健. 基于改进 YOLOv4 的公路路面病害检测算法[J]. 激光与光电子学进展, 2021, 58(14):1410025.
- LUO H, JIA C, LI J. Highway surface disease detection algorithm based on improved YOLOv4 [J]. Laser&Optoelectronics Progress, 2021, 58(14):1410025.
- [23] 程文党, 彭福刚, 刘文江, 等. 基于改进 TransUNet 的沥青路面病害检测研究[J]. 土木工程, 2024, 13(1):29-35.
- CHENG W, PENG F, LIU W, et al. Research on asphalt pavement disease detection based on improved TransUNet[J]. Hans Journal of Civil Engineering, 2024, 13(1):29-35.
- [24] 陈建瑜, 邹春龙, 王生怀, 等. 改进 YOLOv5 的路面缺陷快速检测方法研究[J]. 电子测量技术, 2023, 46(10):129-135.
- CHEN J, ZOU C, WANG S, et al. Research on Rapid Detection Method of Pavement Defects with improved YOLOv5[J]. Electronic Measurement Technology, 2019, 46(10):129-135.
- [25] ROY A M, BHADURI J. DenseSPH-YOLOv5: An automated damage detection model based on DenseNet and Swin-Transformer prediction head-enabled YOLOv5 with attention mechanism[J]. Advanced Engineering Informatics, 2023, 56:102007.
- [26] 孙建诚, 杨舒涵, 龚芳媛, 等. 基于改进 YOLOv5 的复杂背景下路面裂缝检测[J]. 中国科技论文, 2023, 18(7):779-785.
- SUN J, YANG S, GONG F, et al. Pavement crack detection under complex background based on improved YOLOv5[J]. Chinese Journal of Science and Technology, 2023, 18(7):779-785.
- [27] 景峰, 刘晓捷, 刘军, 等. 基于融合注意力和任务解耦的路面裂缝检测[J]. 计算机工程与设计, 2023, 44(5):1565-1571.
- JING F, LIU X, LIU J, et al. Pavement crack detection based on fusion attention and task decoupling[J]. Computer Engineering and Design, 2023, 44(5):1565-1571.
- [28] 杨振, 李林, 罗文婷, 等. 改进 YOLOv5 的沥青路面病害检测算法[J]. 计算机工程与设计, 2023, 44(11):3360-3372.
- YANG Z, LI L, LUO W, et al. Improved Asphalt pavement disease Detection Algorithm of YOLOv5[J]. Computer Engineering and Design, 2023, 44(11):3360-3372.
- [29] 王梦梦, 黄德启, 刘爽娜. 基于改进 YOLOv5s 的路面损伤检测[J]. 激光杂志, 2023, 44(5):66-71.
- WANG M, HUANG D, LIU S. Pavement damage detection based on improved YOLOv5s[J]. Laser Journal, 2023, 44(5):66-71.
- [30] YU Z, HUANG H, CHEN W, et al. Yolo-facev2: A scale and occlusion aware face detector[J]. arXiv preprint arXiv: 2208.02019, 2022.
- [31] ZHUANG J, QIN Z, YU H, et al. Task-specific context decoupling for object detection[J]. arXiv preprint arXiv: 2303.01047, 2023.
- [32] Dai X, Chen Y, Xiao B, et al. Dynamic head: Unifying object detection heads with attentions[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021:7373-7382.
- [33] YAN H, ZHANG J. UAV-PDD2023: A benchmark dataset for pavement distress detection based on UAV images[J]. Data in Brief, 2023, 51:109692.

#### 作者简介

李金涛(1999-), 男, 硕士研究生, 主要从事路面缺陷检测、深度学习等方面的研究等。

E-mail: 2510941573@qq.com

周兴林(1965-), 男, 教授, 主要研究领域为先进检测技术与应用, 交通信息与控制系统。

E-mail: zxl65@163.com