

# 基于机器学习的金属软管缺陷检测系统<sup>\*</sup>

倪洪启<sup>1</sup> 李鑫宇<sup>1</sup> 戴文博<sup>2</sup> 李宝立<sup>1</sup>

(1. 沈阳化工大学机械与动力工程学院 沈阳 110142; 2. 秦皇岛北方管业有限公司 秦皇岛 066004)

**摘要:** 为了实现工业上对金属软管缺陷部分的自动检测,提出一种基于深度学习的缺陷检测方法,首先利用相机采集金属软管缺陷部分的图像并将采集图像中的缺陷特征部分进行分类与标定,金属软管外表面缺陷可分为断丝、散丝、叠丝3种并制作出对应的自制数据集;其次对YOLOv5s网络进行改进,通过在YOLOv5s中的主干网络中添加SimAM注意力机制;然后利用EIoU损失函数替换初始网络所采用的IoU损失函数;最后对YOLOv5s中的金字塔池化层进行改进,采用SimSPPF模块替换SPPF模块。利用改进后的算法对金属软管缺陷数据集进行训练,改进后的算法相较于初始YOLOv5s网络的平均精度mAP提升了1.5%,特征复杂且小目标的漏检误检情况有了明显改善。

**关键词:** 缺陷检测;YOLOv5s;注意力机制;损失函数

中图分类号: TP277; TN06 文献标识码: A 国家标准学科分类代码: 520.6040

## Machine learning based defect detection system for metal hoses

Ni Hongqi<sup>1</sup> Li Xinyu<sup>1</sup> Dai Wenbo<sup>2</sup> Li Baoli<sup>1</sup>

(1. School of Mechanical and Power Engineering, Shenyang University of Chemical Technology, Shenyang 110142, China;

2. Qinhuangdao North Pipe Industry Co., Ltd., Qinhuangdao 066004, China)

**Abstract:** In order to achieve automatic detection of defects in metal hoses in industry, a deep learning based defect detection method is proposed. Firstly, a camera is used to capture images of defects in metal hoses, and the defect feature parts in the collected images are classified and calibrated. The surface defects of metal hoses can be divided into three types: broken wire, loose wire, and stacked wire, and corresponding self-made datasets are created; Secondly, the YOLOv5s network is improved by adding SimAM attention mechanism to the backbone network of YOLOv5s; Then use the EIoU loss function to replace the IoU loss function used by the initial network; Finally, the pyramid pooling layer in YOLOv5s was improved by replacing the SPPF module with the SimSPPF module. The improved algorithm was used to train the dataset of metal hose defects. Compared with the initial YOLOv5s network, the average accuracy mAP of the improved algorithm increased by 1.5%, and the missed and false detections of complex features and small targets were significantly improved.

**Keywords:** defect detection; YOLOv5s; attention mechanism; loss function

## 0 引言

金属软管常被应用在石油化工领域,是一种十分重要的承压部件,其不仅能够吸收设备振动所产生的大部分能量,还能够对管道或设备的连接进行位移补偿<sup>[1]</sup>。

工程上对于金属软管的缺陷检测主要依靠人工肉眼进行抽样检测,由于金属编织软管的缺陷特征较为复杂,且不同的员工其技能水平也不相同,所以在进行人工抽样检测的过程中不仅会浪费大量时间,往往还会出现不同程度的漏检或误检的情况<sup>[2-4]</sup>。

金属软管外表面编织网的表面特征十分的复杂且传统视觉检测方法对背景环境以及光照要求都较为严格,并不适用于检测缺陷种类较多且特征较为复杂的金属编织软管<sup>[5]</sup>。所以,考虑采用基于深度学习的计算机视觉检测方法对金属编织软管进行缺陷识别与检验。

基于深度学习的计算机视觉检测方法不需要提前进行人工设定,能够在缺陷检测的过程中自动提取出所要检测的缺陷特征。常用的深度学习方法大致有以下6种:卷积神经网络(convolutional neural network, CNN)、单阶段检测器(single shot multibox detector, SSD)、特征金字塔网络

(feature pyramid networks, FPN) 以及另一种单阶段目标检测算法 (you only look once, YOLO) 等<sup>[6-12]</sup>。其中, YOLOv5 是 YOLO 算法中最为稳定且最为完善的算法, 同时也是在工程检测以及目标识别当中使用最多的算法之一, YOLOv5 具有检测速度快, 目标识别精准且所占内存较小等多个优点, 而在移动端设备当中 YOLOv5s 是检测速度最快且检测精度较高的算法<sup>[13]</sup>。

## 1 YOLOv5s 网络结构

YOLOv5 是基于 YOLOv4 基础上进行改进的一种目标检测算法, 相较于前几代 YOLO 算法, YOLOv5 具有更快的检测速度、更高的检测精度且更方便使用。YOLOv5 一共具有 5 种检测框架分别为: YOLOv5n、YOLOv5s、YOLOv5m、YOLOv5l 以及 YOLOv5x, 按照顺序其网络深度、宽度以及参数量都是在逐步增加的。YOLOv5s 计算量适中且具有较高的精度以及检测速度, 更适用于采用移动设备进行模型的训练与缺陷检测<sup>[14]</sup>。

YOLOv5s 模型主要是由输入端 (Input)、主干网络 (Backbone)、特征融合模块 (Neck) 以及目标检测模块 (Head)4 部分所组成, 其模型结构如图 1 所示。

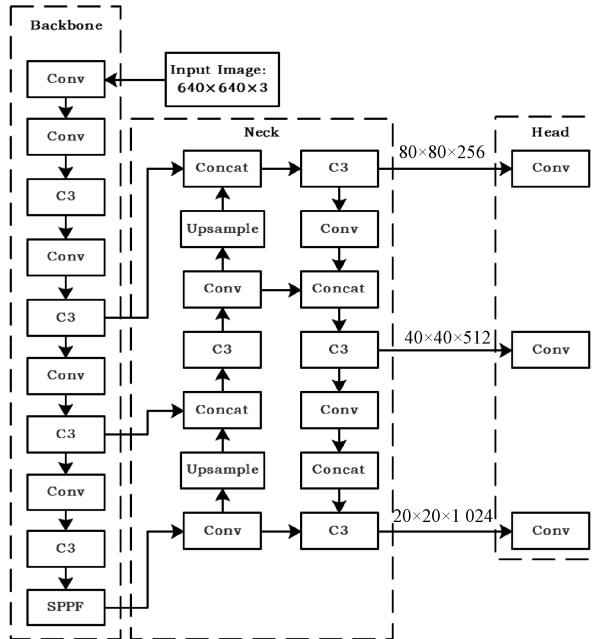


图 1 YOLOv5s 模型结构

在 YOLOv5 中输入端 (Input) 会将图片进行自适应缩放, 将所有的图像均缩放为像素大小  $640 \times 640$  的标准图像, 然后利用 Mosaic 数据增强方法, 在数据集中随机选择 4 张训练图像进行拼接剪裁生成一个新的图像, 并使其拼接后的图像大小同样为  $640 \times 640$ , 以此来提高模型的性能以及泛化能力进而提高后续检测的准确度。

Backbone 骨干网络是一个深度卷积神经网络, 其主要作用是将原始输入图像转化为多层特征图像以便于后续检

测提取图像特征, 提高检测精确度, 其中主要包含: C3、Conv、快速金字塔池化结构模块 (spatial pyramid pooling-fast, SPPF) 3 个模块。

Neck 特征金字塔的主要作用是在目标检测过程中提供多尺度特征表达, 提高模型对小目标或远距离目标的检测能力, 对 Backbone 骨干网络处理后的多特征图像进行上采样或者下采样处理, 将不同层次的特征图进行融合, 提高模型的性能以及鲁棒性。

Head 主要作用是对 Neck 中特征金字塔处理后的图像进行目标边界框位置以及类别概率的预测, 将检测到的特征利用不同颜色的方框进行标注并显示其标注类别与相似概率。

## 2 YOLOv5s 网络改进及训练

### 2.1 添加注意力机制

针对金属软管缺陷特征复杂且尺寸较小、分布密集等特点, 致使检测时容易产生漏检或误检的情况。因此为了提高模型检测精度, 考虑在 backbone 主干网络末端处添加无参数 3D 注意力机制 (SimAM)。

通道注意力模块只能计算一维的注意力权值, 而空间注意力模块只能计算二维的注意力权值, SimAM 注意力机制将通道注意力与空间注意力相结合, 构成三维注意力机制且能够直接计算出三维注意力权值, 在尽可能的保证计算量不变的情况下提高准确率, SimAM 注意力机制原理如图 2 所示。

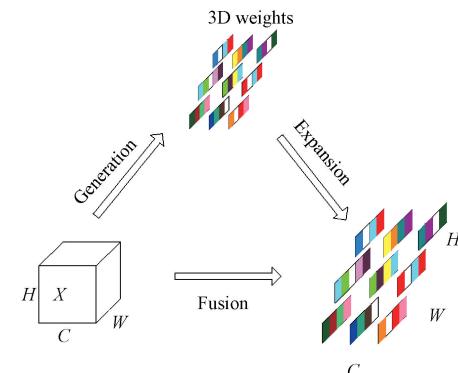


图 2 SimAM 注意力机制原理图

该注意力机制的灵感来源于神经科学理论, 在信息传递的过程当中, 携带更多信息的神经元的放电模式与周围携带较少信息的神经元的放电模式不同, 携带较多信息的神经元还会对周围携带较少信息的神经元产生空间抑制<sup>[15]</sup>。将该理论与能量函数相结合, 便能够直接从当前的神经元中推断出图像中所标记的重要部分。

神经元能量函数为:

$$e_i(w_i, b_i, y_i, x_i) = (y_i - \hat{y}_i)^2 + \frac{1}{M-1} \sum_{i=1}^{M-1} (y_0 - \hat{x}_i)^2 \quad (1)$$

式中:  $e_t$  为神经元能量函数;  $y_0$  为其他神经元,  $y_0 = -1$ ;  $y_t$  为目标神经元,  $y_t = 1$ ;  $x_i$  为其余神经元;  $\hat{t}$  为目标神经元线性变化所得值;  $\hat{x}_i$  为其他神经元线性变化所得值;  $w_t$  为权值;  $b_t$  为偏置。

上述公式中,  $\hat{t}$ 、 $\hat{x}_i$  是线性变化所得, 所以目标神经元与其他神经元的线性变化公式为:

$$\hat{x}_i = w_t t + b_t \quad (2)$$

$$\hat{t} = w_t x_i + b_t \quad (3)$$

式中:  $x_i$  为输入特征  $X$  的单个通道中的其他神经元;  $t$  为输入特征  $X$  的单个通道中的目标神经元。

通过将  $y_0 = -1$ 、 $y_t = 1$  以及式(2)和(3)带入到式(1)中并加入正则项  $\lambda w_t^2$  后对  $w_t$  和  $b_t$  求偏导解得:

$$\mu_t = \frac{1}{M-1} \sum_{i=1}^{M-1} x_i \quad (4)$$

$$\sigma_t^2 = \frac{1}{M-1} \sum_{i=1}^{M-1} (x_i - \mu_t)^2 \quad (5)$$

$$w_t = -\frac{2(t - \mu_t)}{(t - \mu_t)^2 + 2\sigma_t^2 + 2\lambda} \quad (6)$$

$$b_t = -\frac{1}{2}(t + \mu_t)w_t \quad (7)$$

式中:  $\mu_t$  为除  $t$  以外所有神经元的均值;  $\sigma_t$  为除  $t$  以外所有神经元的方差;  $\lambda$  为常量。

由上述公式能够得到其最小能量为:

$$e_t^* = \frac{4(\sigma^2 + \lambda)}{(t - \hat{\mu})^2 + 2\sigma^2 + 2\lambda} \quad (8)$$

由式(8)可以看出, 能量越小, 神经元  $t$  与其余神经元的差别越大, 越能凸显出神经元  $t$  的重要性, 由此再对神经元所对应的特征进行增强处理, 便能够使得该注意力机制提高特征识别的准确率。

神经元对应特征增强处理公式为:

$$\tilde{X} = \text{sigmoid}\left(\frac{1}{E}\right) \times X \quad (9)$$

式中:  $X$  为输入的特征张量;  $E$  为所有最小能量跨通道和空间维度之和;  $\text{sigmoid}$  为固定函数。

## 2.2 更换空间金字塔池化层

YOLOv5s 主干网络中金字塔池化层默认使用的是 SPPF 模块, 该模块是基于 YOLOv3 中默认采用的金字塔池化结构模块(spatial pyramid pooling, SPP)改进所得<sup>[16]</sup>, 其主要是对不同尺度的特征图进行池化操作, 通过设置大小不同的池化窗口, 对图像中不同尺度的空间信息进行捕捉, 以此来丰富图像特征的表达, 同时还能够避免对图像进行剪裁以及缩放等操作后造成图像失真的问题<sup>[17]</sup>, SPPF 模块原理如图 3 所示。

为了加快学习速度并提高小目标以及复杂特征识别的准确性, 对其金字塔池化层进行改进, 采用轻型快速金字塔池化结构模块(simplified spatial pyramid pooling, SimSPPF)

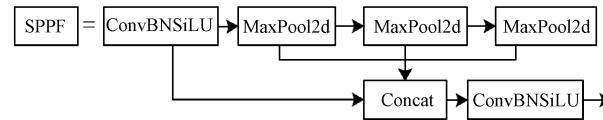


图 3 SPPF 模块原理

替换 SPPF 模块并以此来提取图像的多尺度特征, 其中, SimSPPF 模块原理如图 4 所示。

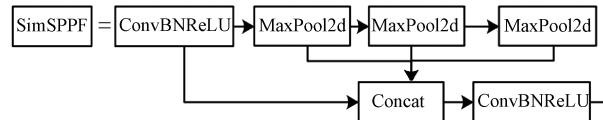


图 4 SimSPPF 模块原理

SimSPPF 是根据 SPPF 作为基础改进而来的金字塔池化层模块, 其结构大部分与 SPPF 模块相同, 均是将 3 个尺寸大小分别为:  $5 \times 5$ ,  $9 \times 9$  以及  $13 \times 13$  的 3 个池化层进行串联, 而相较于 SPPF 模块的不同之处在于, SimSPPF 模块将 SPPF 中的 SiLU 激活函数替换为 ReLU, 减少模型的计算量, 使其模型运算速度更快, 加快模型的学习效率。同时 SimSPPF 模块与 SPPF 模块中的 Conv 模块还会分出一部分直接与 Concat 模块相连接, 以此来增强图像特征的识别能力, 进而保持模型学习的准确率。因为 SimSPPF 是以 SPPF 作为基础所改进的轻量化模型, 所以模型学习的速度会有所提升, 准确率 P 及 mAP 值照或许不会有较大变化, 但相较于 YOLOv3 中所采用的 SPP 模块, 其准确率 P 以及 mAP 值会有较大的提升。

## 2.3 更换损失函数

在 YOLOv5 中损失函数是用来计算预测边界框与真实边界框之间的重叠程度, 通过损失函数值可以直观的看出模型对于目标特征位置预测的准确程度, 同时损失函数也常用于非极大值抑制(non-maximum suppression, NMS), 选择出相似度最高最具代表性的边界框进行显示, 以此来提高目标特征检测的准确率。

为了提高金属编织软管缺陷模型训练过程中对于目标缺陷位置识别的准确度, 将 YOLOv5s 网络默认使用的损失函数(IoU)替换为 EIoU, 该损失函数是在 CIoU 的基础上进行改进, CIoU 不仅考虑到检测框与目标框的重叠面积, 同时还考虑到了 IoU 为考虑到的中心点距离以及宽高比, 除长宽比外, EIoU 所考虑的部分与 CIoU 是相同的<sup>[18]</sup>。

在考虑宽高比时, EIoU 会使目标盒与锚盒的宽高之差变为最小, 使得其收敛速度更快, 特征识别更准确, EIoU 计算公式如下:

$$\text{IoU} = \frac{A \cap B}{A \cup B} \quad (10)$$

$$\text{EIoU} = \text{IoU} - \frac{\rho^2(b, b^{gt})}{c^2} - \frac{\rho^2(w, w^{gt})}{c_w^2} - \frac{\rho^2(h, h^{gt})}{c_h^2} \quad (11)$$

式中:  $IoU$  为预测框与真实框间的交并比;  $A \cap B$  为区域 A 与区域 B 的交集;  $A \cup B$  为区域 A 与区域 B 的并集;  $c_w$  为代表覆盖两个盒的最小外接框的宽度;  $c_h$  为代表覆盖两个盒的最小外接框的高度;  $\rho$  为代表两个区域内中心点之间的欧式距离;  $b$  为预测框中心点;  $b^{gt}$  为真实框的中心点;  $c$  为预测框与真实框重合区域的对角线间距。

## 2.4 金属编织软管数据集

采用自制的金属软管缺陷数据集进行训练, 将常见的金属软管缺陷分为 3 种分别为: 断丝如图 5(a)所示、叠丝如图 5(b)所示以及散丝如图 5(c)所示。

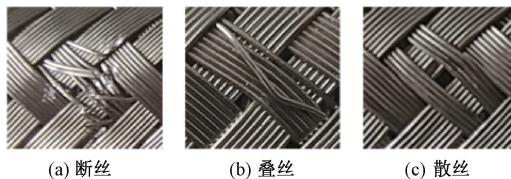


图 5 金属软管缺陷

数据集通过 CCD 工业相机在自然环境中以及实验室中进行采集, 数据集中的缺陷部分主要通过人工破坏、长久使用磨损以及运输过程中的一些磕碰等不确定因素所造成, 多种不同形状的破坏提高数据集的可信度。在对数据集进行采集的过程中需要尽可能清晰的拍摄出金属软管中的缺陷部分, 以便于在训练过程中提高缺陷的识别能力与缺陷识别的准确度。

初始数据集含有 593 张带有缺陷的金属软管图像, 对数据集内的图片进行亮度上的改变、对比度改变以及对图片显示方向进行改变, 然后对扩充后的数据集进行筛选, 去除极为模糊不具有训练价值的图像, 经过扩充后参与模型训练的数据集达到了 1 852 张, 其中选取 1 052 张数据集作为训练集, 400 张测试集, 400 张验证集, 其中包含叠丝缺陷的图片有 1 620 张、散丝缺陷的图片有 844 张、叠丝缺陷的图片 793 张, 在这些图像中, 有些金属软管图像包含多种缺陷, 而有些金属软管只包含其中一种缺陷。在 LabelImg 中将 1 852 张数据集中的断丝、叠丝、散丝缺陷以及管体部分用最小外接矩形进行框选标注, 对管体部分进行标注的主要目的是为了在检测时能够更加直观且清晰的观察到复杂背景中金属软管的所在位置。

## 2.5 设备信息及模型评估

模型训练所采用的计算机硬件及开发环境名称如表 1 所示。

使用改进后的 YOLOv5s 网络训练模型耗时大约为 7.5 h, 得到金属软管缺陷模型文件, 其大小为 15.3 MB。将训练成功的模型文件放入到相对应的文件夹中并对其模型进行评估, 综合考虑其准确率  $P$ 、召回率  $R$ 、均值平均精度  $mAP$  等结果参数来评价模型的性能。

准确率  $P$  代表在所有的样本中, 预测正确的概率, 准确率  $P$  越高代表模型评估预测结果越准确; 召回率  $R$  代表

表 1 计算机硬件及开发环境

配置名称	参数
计算机系统	Windows11 64 位
GPU	Intel i5-12500H
CPU	NVIDIA GeForce RTX 3050
内存大小	16 GB
开发语言	Python3.8
深度学习框架	PyTorch

所有预测出的正样本中实际为正样本的概率, 召回率  $R$  越高代表预测结果越全面; 均值平均精度  $mAP$  代表衡量模型在每个类别上的检验效果好与坏的平均值,  $mAP$  越高, 代表模型对于每个类别的检验效果越好。准确率  $P$ 、召回率  $R$ 、均值平均精度  $mAP$  公式如下:

$$P = \frac{TP}{TP + FP} \quad (12)$$

$$R = \frac{TP}{TP + FN} \quad (13)$$

$$AP = \sum_{i=1}^{n-1} (r_{i+1} - r_i) P(r_i + 1) \quad (14)$$

$$mAP = \frac{\sum_{i=1}^k AP_i}{k} \quad (15)$$

式中:  $TP$  为缺陷目标被成功识别的数量;  $FP$  为将非缺陷目标识别为缺陷目标的数量;  $FN$  为漏检数量;  $AP$  为平均精准度;  $r_i$  为每次迭代所对应的召回率;  $K$  为迭代次数。

## 3 金属软管缺陷检测结果与分析

### 3.1 消融实验

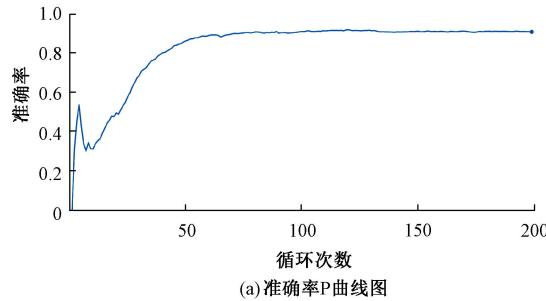
为验证改进模块对模型性能所产生的影响, 将 3 种模块拆开放入到 YOLOv5s 初始网络当中, 以原始 YOLOv5s 网络训练结果作为基础进行实验对比, 观察其训练后的结果并进行对比, 其消融实验结果如表 2 所示。

当改进模型训练成功后, 可以看出当模型训练开始收敛时, 总体准确率  $P$  如图 6(a)所示始终稳定在 90% 左右, 召回率  $R$  如图 6(b)所示稳定在 80% 以上,  $mAP$  值如图 6(c)所示稳定在 85% 左右。

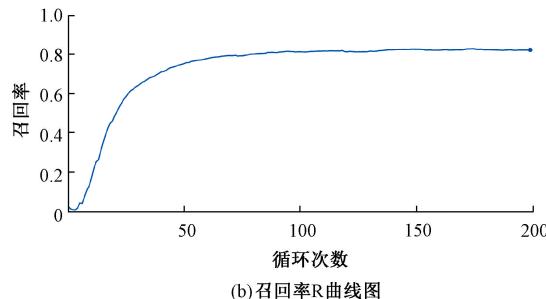
通过将上述表格中的数据进行比对可知, 当采用不同算法的情况下, 各个算法的正确率、召回率以及  $mAP$  值均会产生不同的变化, 其总体  $mAP$  值仍然是改进后的 YOLOv5s 网络最高, 其值为 87.8%, 对比初始网络的总体  $mAP$  值提高了 1.5%, 这也意味着改进后的模型对于缺陷特征的识别最为准确。改进后的 YOLOv5s 模型其断丝缺陷的  $mAP$  值最高, 相较于其余模型  $mAP$  值最高提高了 2.0%; 叠丝缺陷的  $mAP$  值相较于只采用了 SimSPPF 的模型降低了 0.4%, 而对比其余模型结果均为提高且最高提高了 2.7%; 散丝缺陷的  $mAP$  值最高且相较于其余模型

表 2 消融实验结果表

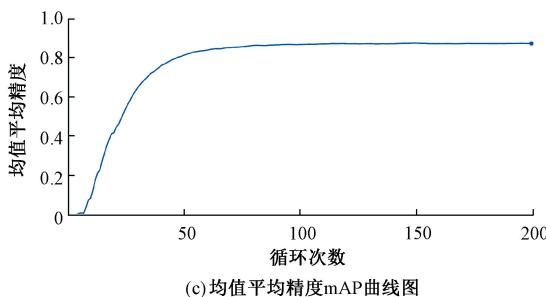
改进模块			断丝 (Fracture)			叠丝 (Wire stacking)			散丝 (Wire dispersion)			管体 (Tubebody)			总体 mAP
SimAM	SimSPPF	EIoU	P	R	mAP	P	R	mAP	P	R	mAP	P	R	mAP	
×	×	×	97.5	90.8	92.7	77.2	69.1	77.4	84.6	64.6	75.8	98.1	99.2	99.4	86.3
×	×	✓	96.7	90.6	93.0	77.0	69.3	76.6	87.8	61.1	74.4	99.4	99	99.4	85.9
×	✓	×	96.6	91.7	93.8	82.9	70.5	79.4	89.8	65.4	75.3	99.6	99.2	99.5	87.0
✓	×	×	96.0	90.5	92.8	79.1	74.1	78.9	86.7	65.3	75.7	99.0	99.6	99.4	86.7
✓	✓	×	96.6	90.6	93.5	75.5	72.3	76.3	84.5	59.3	72.2	99.8	99.2	99.4	85.4
✓	✓	✓	95.9	92.4	94.7	75.4	75.2	79.0	88.9	71.7	78.2	98.7	99.0	99.5	87.8



(a) 准确率P曲线图



(b) 召回率R曲线图



(c) 均值平均精度mAP曲线图

图 6 训练结果曲线

最高提高了 6%；管体特征的  $mAP$  值则是与 SimSPPF 模型相互持平，对比其余模型最高仅提高了 0.1%。

### 3.2 对比实验

为了能够说明改进后模型的优越性且更好的分析本文算法的性能，将该模型与目前使用较多的深度学习网络所训练出的模型结果进行对比分析，具体模型数据如表 3 所示。

通过与目前使用较多的主流算法进行对比分析可知，

表 3 对比实验结果表

模型	模型大小/MB	mAP/%	FPS/(f·s <sup>-1</sup> )
SSD	92.1	74.92	45.38
Faster-RCNN	108	86.41	21.57
YOLOv4	244	46.65	25.08
YOLOv5s	13.7	86.30	95.24
本算法	15.3	87.8	104.17

本算法所得到的模型所占内存只比 YOLOv5s 初始网络模型增大 2.4 MB，此内存大小差距可忽略不计，相比所占内存最大的 YOLOv4 模型其内存减小了 228.74 MB；对比其余几种算法，本算法的 FPS 值最高，达到了 104.17 f/s，相较于 FPS 最低的 Faste-RCNN 模型，FPS 值提高了 82.6 f/s；经过改进后的模型其对金属软管缺陷识别的准确的也最高，相较于准确度最低的 YOLOv4 模型，本算法所训练的模型对于缺陷识别的准确度提高了 41.15%。综上所述，本算法的检测速度最快且缺陷识别的准确度也最高。

### 3.3 结果分析

利用训练好的模型对带有缺陷的金属软管图片进行检验与验证，其断丝缺陷检测效果如图 7(a)所示、散丝检测效果图如图 7(b)所示、叠丝缺陷检测效果如图 7(c)所示，同时还对包含多种缺陷的单根金属软管进行检测，其检测效果图如图 7(d)所示。通过图片可以看出，该训练模型能够清晰的检验并标定出与训练集种类相同但形状不同的缺陷。

通过对包含各个特征的 120 张金属软管图片进行检验可以得知，断丝的检验正确率为 97.5%，散丝的检验正确率为 95.8%，叠丝的检验正确率为 96.7%，管体特征检验正确率为 98.3%。整体检验效果比较理想，通过对检验的图像进行分析可以得知，断丝、叠丝以及散丝被误检的主要原因是因为其拍摄距离较远导致像素不清晰，且金属软管表面特征较为复杂容易与复杂的背景环境混淆。检验结果如表 4 所示。



图 7 缺陷检测效果图

表 4 金属软管特征检验结果

特征种类	缺陷数量	误检原因及数量		准确率/%
		漏检数量	错检数量	
断丝	120	1	2	97.5
散丝	120	4	1	95.8
叠丝	120	4	0	96.7
管体	120	1	1	98.3

#### 4 结 论

本文在 YOLOv5s 算法中添加 SimAM 三维注意力机制来提高模型训练的准确率, 将初始 IoU 损失函数修改为 EIoU 损失函数以此来提高模型缺陷位置识别的准确程度, 同时将初始金字塔池化层 SPPF 修改为 SimSPPF, 在保持模型准确度的同时提高模型的训练速度减少计算量。

采用改进 YOLOv5s 算法训练金属编织软管缺陷模型并利用该模型对断丝、叠丝、散丝 3 个缺陷以及管体本身

进行实时检测,该检测方法不仅能在复杂背景中准确的识别并标定出管体本身以及管体表面所含有的缺陷,同时还能够显示出其特征名称方便观察记录。经过计算可知,模型的准确率高达 87.8%,而模型总体大小为 15.3 MB,FPS 为 104.17 f/s,该模型检测速率、模型大小以及对于特征检测的准确率满足实时检测要求且对硬件设备要求不高。

## 参考文献

- [1] 苏文娟,毛小虎,王志成,等.金属软管的标准比对和质量提升研究[J].压力容器,2020,37(9):53-60,72.
- [2] 王威,张庆.机器视觉在机械设计制造及其自动化中的应用研究综述[J].内燃机与配件,2023(18):114-116.
- [3] 张涛,刘玉婷,杨亚宁,等.基于机器视觉的表面缺陷检测研究综述[J].科学技术与工程,2020,20(35):14366-14376.
- [4] 李衍照,于镭,田金文.基于改进 YOLOv5 的金属焊缝缺陷检测[J].电子测量技术,2022,45(19):70-75.
- [5] 戈明辉,张俊,陆慧娟.基于机器视觉的食品外包装缺陷检测算法研究进展[J].食品与机械,2023,39(9):95-102,116.
- [6] LI K, WAN G, CHENG G, et al. Object detection in optical remote sensing images: A survey and a new benchmark[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 159: 296-307.
- [7] 卢宏涛,张秦川.深度卷积神经网络在计算机视觉中的应用研究综述[J].数据采集与处理,2016,31(1):1-17.
- [8] HE K, GKIOXARI G, DOLLÁR P, et al. Mask r-cnn [C]. Proceedings of the IEEE International Conference on Computer Vision, 2017: 2961-2969.
- [9] 汤文虎,吴龙,黎尧,等.基于改进 Faster RCNN 的钢线圈头部小目标检测算法[J].现代制造工程,2023,(8):127-133,147.
- [10] 周昕,骆莉莎,张瑞军,等.基于 SSD 算法的老旧建筑墙面裂缝识别[J].科技创新与应用,2023,13(28):36-39.
- [11] 黄强,潘晴,田妮莉.一种用于目标检测的位置加权特征金字塔网络[J].计算机仿真,2023,40(6):192-196.
- [12] 郑新月,任俊超.基于 BERT-FNN 的意图识别分类[J].计算机与现代化,2021(7):71-76,88.
- [13] 王建,徐闯.基于 YOLOv5s 的脐橙虫害检测研究[J].工业控制计算机,2023,36(7):105-106,109.
- [14] 吴祥祥,谈莉斌,徐向荣.基于 YOLOv5s 的丁腈手套缺陷检测方法[J].制造业自动化,2023,45(9):1-4,10.
- [15] XIE J, WU Z, ZHU R, et al. Melanoma detection based on swin transformer and SimAM [C]. 2021 IEEE 5th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), IEEE, 2021, 5: 1517-1521.
- [16] LANG X, REN Z, WAN D, et al. MR-YOLO: An improved YOLOv5 network for detecting magnetic ring surface defects[J]. Sensors, 2022, 22(24): 9897.
- [17] 何铁军,李华恩.基于改进 YOLOv5 的路面病害检测模型[J].土木工程学报,2024,57(2):96-106.
- [18] ZHANG Y F, REN W, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression[J]. Neurocomputing, 2022, 506: 146-157.

## 作者简介

倪洪启,副教授,主要研究方向为机械设计制造及其自动化研究。

E-mail: nihongqi@163.com

李鑫宇(通信作者),硕士研究生,主要研究方向为机器视觉。

E-mail: 18704146487@163.com