

# 基于 DCGAN 和改进 YOLOv5s 的 钢丝帘布缺陷检测方法 \*

黄鹏 蔡露 陈彬 周益航 易冬旺

(桂林理工大学机械与控制工程学院 桂林 541000)

**摘要:** 为解决钢丝帘布表面缺陷检测准确率低且泛化能力不强的问题,提出了一种基于 DCGAN 和改进 YOLOv5s 的缺陷检测方法。首先,通过调整 DCGAN 网络参数并优化超参数,使生成器能够生成具有丰富特征和清晰纹理的钢丝帘布缺陷图像,从而扩充数据集;其次,采用 K-means++ 算法对钢丝帘布缺陷数据重新聚类锚框,以获得更优的锚框参数,实现锚框与实际缺陷的精确匹配;然后,在 YOLOv5s 主干网络中的 C3 模块添加坐标注意力机制,以增强模型的特征提取能力和精确定位能力;最后,引入 MPDIoU 损失函数替换 YOLOv5s 原损失函数,进一步提高检测精度。实验结果表明,在实测钢丝帘布缺陷数据集上,采用 DCGAN 数据增强和改进后的 YOLOv5s 检测模型,缺陷检测平均精度提高了 6.6%,达到了 89.4%,并且检测准确率和召回率也有所提高。与其他主流检测模型相比,该模型不仅在检测速度上提高了约 30%,还保持较高的检测精度。在公开的 NEU-DET 数据集上,该模型的 mAP 值达到了 82.6%,较原始 YOLOv5s 模型提高了 3.8%。

**关键词:** 钢丝帘布缺陷检测;生成对抗网络;K-means++;注意力机制;MPDIoU 损失函数

**中图分类号:** TP391.4 **文献标识码:** A **国家标准学科分类代码:** 520.6040

## Defect detection method of steel cord based on DCGAN and improved YOLOv5s

Huang Peng Cai Lu Chen Bin Zhou Yihang Yi Dongwang

(School of Mechanical and Control Engineering, Guilin University of Technology, Guilin 541000, China)

**Abstract:** In order to solve the problems of low detection accuracy and weak generalization ability of steel cord surface defects, a steel cord defect detection method based on DCGAN and improved YOLOv5s was proposed. Firstly, by adjusting DCGAN network parameters and optimizing hyperparameters, the generator can generate steel cord defect images with rich features and clear texture, thus expanding the data set. Secondly, the K-Means++ algorithm is used to re-cluster the anchor frame to obtain better anchor frame parameters and achieve accurate matching between anchor frame and actual defects. Then, coordinate attention mechanism was added to C3 module of YOLOv5s backbone network to enhance the feature extraction capability and accurate localization capability of the model. Finally, MPDIoU loss function is introduced to replace YOLOv5s original loss function to further improve the detection accuracy. The experimental results show that on the measured steel cord defect data set, the average accuracy of defect detection is increased by 6.6%, reaching 89.4% by using the YOLOv5s detection model enhanced and improved by DCGAN data, and the detection accuracy and recall rate are also improved. Compared with other mainstream detection models, this model not only improves the detection speed by about 30%, but also maintains high detection accuracy. On the publicly available NEU-DET dataset, the mAP value of this model reaches 82.6%, which is 3.8% higher than that of the original YOLOv5s model.

**Keywords:** steel cord defect detection; generate adversarial network; K-means ++; attention mechanism; MPDIoU loss function

## 0 引言

钢丝帘布是轮胎内部结构的主体,它是带束层和冠带

层的关键组成部分。作为轮胎的主要受力部件,钢丝帘布承受着内部气压、外部载荷和冲击等。因此,轮胎制造企业十分关注钢丝帘布的表面缺陷,以确保轮胎的寿命和安全

收稿日期:2023-10-09

\* 基金项目:国家自然科学基金(72001054)、广西科技计划项目(桂科 AB22035041)、桂林市科学研究与技术开发计划项目(20210217-14)、桂林理工大学科研启动基金(GUTQDJ20160140)项目资助

性能,并及时发现和控制缺陷。但是,现有视觉缺陷检测系统存在大量的误报和漏报现象,仍然无法满足生产需要。钢丝帘布生产线是 24 小时不间断运行的,检测系统误报就会导致生产线停机,这极大的影响了生产效率;而漏报现象就会导致存在缺陷的钢丝帘布进入下一步的生产环节中,从而影响产品质量。因此,对于检测系统的改进和完善是十分必要的,以提高检测的准确率和效率。误报和漏报现象的原因主要有以下 3 点:首先,不同的工厂环境、光线条件以及产品本身反光率等因素导致最终成像质量有所差异。其次,钢丝帘布本身缺陷的大小和位置可能存在差异。此外,检测模型的准确率是基于大量缺陷样本进行训练而建立的,但实际上缺陷样本较少,训练的模型很难达到预期的缺陷检测效果。

为实现钢丝帘布缺陷的准确检测,将缺陷检测与深度学习相结合是一种常见的方法<sup>[1-3]</sup>。近年来,基于深度学习的缺陷检测方法在很多场景下得到广泛的应用,主流缺陷检测算法有区域卷积神经网络<sup>[4]</sup>(region-based convolutional neural networks, R-CNN)系列、单步多框检测器<sup>[5]</sup>(single shot multibox detector, SSD)以及 YOLO<sup>[6]</sup>(you only look once, YOLO)系列。YOLO 算法作为一阶段网络的代表,具有检测速度快,检测精度高等优势,从而成为当前目标检测研究领域的热门方法。曾秀云等<sup>[7]</sup>通过对 Faster R-CNN 中区域建议网络做出改进,并在主干网络中融合特征金字塔结构,从而实现对棉布包装的缺陷的检测,但模型检测速度较低,无法满足实时性检测。阎馨等<sup>[8]</sup>采用 Transformer 多头注意力机制模块改进 SSD 网络实现对钢材表面的缺陷检测,虽然模型整体检测精度有所提升,但不同类别缺陷检测精度差异较大。黄连等<sup>[9]</sup>提出了一种基于奇异值分解的橡胶密封圈表面缺陷检测方法,通过提取缺陷特征信息实现对缺陷大小及位置检测,但该算法稳定性受环境影响较大,不利于工业应用。朱焕宇等<sup>[10]</sup>通过改进 YOLOv5 实现对胎体的钢丝帘线缺陷检测,通过在网络结构中引入注意力机制和深度可分离卷积模块,从而提高模型检测准确率,但检测帧率有所降低。王素珍等<sup>[11]</sup>提出了一种改进的 YOLOv4-Tiny 绝缘子缺陷检测算法,在网络中引入轻量级的注意力机制,同时利用 K-means++ 算法重新对绝缘子的先验锚框聚类,并优化损失函数,从而提高了模型的检测精度,但模型体积仍然较大,不便于工业应用中的部署。

基于深度学习的缺陷检测方法,缺陷样本的匮乏限制了检测的准确率和泛化性能。神经网络强大的特征提取能力以及高水平的缺陷检测能力,需要建立在丰富的数据集的基础之上。为解决因小样本而导致的检测准确率低以及泛化能力不强的问题,采用数据增强的方式就是对已有的钢丝帘布缺陷样本数据进行扩充。数据增强有很多种方式,传统的方式有仿射变换、信息删除、图像融合等方法,钢丝帘布缺陷较为特殊,缺陷的大小、形状、以及位置多样,因

此简单的采用平移、旋转、调整对比度的数据增强方式不适合扩充该缺陷样本的多样性。生成对抗网络<sup>[12]</sup>(generative adversarial network, GAN)在 2014 年被 GOODFELLOW 提出,不同于传统的数据增强,生成对抗网络数据增强通过学习真实的缺陷样本分布,由一组随机噪声生成新的缺陷图像,从而解决了小样本数据增强的问题,GAN 作为一种流行的数据增强方式被广泛应用。GAN 存在训练不稳定、梯度消失、模式崩溃的问题,因此出现了大量的学者对原始 GAN 做出了不同的改进,其中 AlecRadfor 在 2015 年提出深度卷积生成对抗神经网络<sup>[13]</sup>(deep convolution generative adversarial networks, DCGAN),开创性的将卷积神经网络(convolutional neural networks, CNN)与 GAN 相结合,并提出一系列措施从而提高了网络训练的稳定性,使得生成样本的质量得到提升。大量的实践<sup>[14-16]</sup>证明了 DCGAN 对提升检测精度的有效性。然而,由于原始网络生成的图像尺寸较小,难以实现对钢丝帘布缺陷的精细标注。此外,原始网络生成图像的质量较差,缺陷纹理不清晰,反而降低了钢丝帘布缺陷检测的准确率。针对这些问题,因此需要对原始 DCGAN 网络进行改进,以确保对钢丝帘布缺陷数据集的有效增强,并提高最终检测模型的准确率。

为了实现钢丝帘布缺陷的准确检测,提出一种基于 DCGAN 和改进的 YOLOv5s 的钢丝帘布缺陷检测方法。首先,根据待检测缺陷在原始图像中的分辨率与其在检测模型中能够准确检测的最低分辨率,计算出原始图像和生成图像分辨率的比例关系,并据此设置 DCGAN 的生成图像分辨率的大小,在此基础之上,通过优化 DCGAN 网络结构参数使得对抗网络能生成钢丝帘布缺陷图像。同时,以 YOLOv5s-6.2 版本为基础框架,将原来 YOLOv5s 模型中锚框算法由 K-means<sup>[17]</sup>算法改进为 Kmeans++<sup>[18]</sup>聚类算法,实现对钢丝帘布数据集锚框重新聚类,从而提高锚框的尺寸匹配度,并且在模型主干网络中的 C3 结构添加坐标注意力机制<sup>[19]</sup>(coordinate attention, CA),提高模型对缺陷特征提取能力,最后引入基于最小点距离的边界框相似度比较度量<sup>[20]</sup>(minimum point distance based IoU, MPDIoU)作损失函数,简化两个边界框的相似性比较方式,加速模型收敛并进一步提升模型的缺陷检测准确率。

## 1 钢丝帘布缺陷检测总体设计

### 1.1 图像采集系统结构设计

某钢丝帘布生产线的输送带速度最大为 260 m/min,相机能够在 600 mm 左右的工作距离完整采集到视野范围约为 700 mm 的钢丝帘布图像,要求可自动识别的最小缺陷尺寸为 5×5 mm。为了实现钢丝帘布的双面的缺陷准确检测,设计的钢丝帘布图像采集系统整体结构,通过图像采集模块获取钢丝帘布的缺陷原始图像,图像采集模块的硬件结构主要由线阵相机、镜头、光源、工控机及相机支架等组成,图像采集系统结构如图 1 所示。

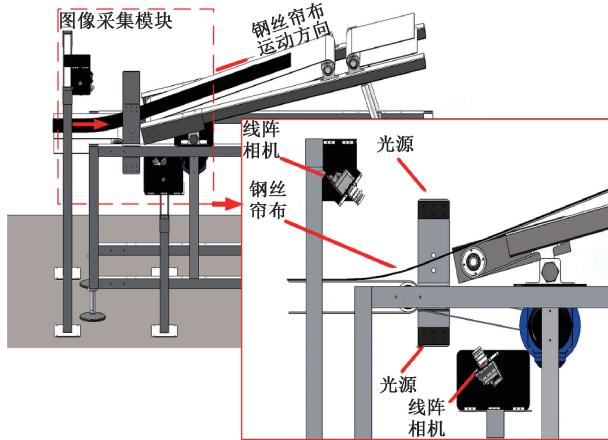


图 1 图像采集系统整体示意图

图像采集系统设计是钢丝帘布缺陷检测系统设计的重要环节,本文根据检测需求和图像采集系统的结构设计了一套满足检测要求的成像硬件。其中相机选用海康威视的 MV-CL042-91CC 型号的工业相机,镜头选用海康威视的 MVL-AF2524M-M42 型号的 FA 镜头,光源最终选用信诺诚电子科技有限公司的 SEENO-SL81238 型号的 LED 白色条形光源,工控机型号为研华 IPC-940。图像采集装置安装在两个输送带之间空隙处,并上下两侧均安装一个线阵相机,钢丝帘布跟随输送带运动,在经过检测位置时,首先由条形光源提供充足的光信号,相机实时捕捉钢丝帘布表面的光信号,并将其转化为电信号,再进一步转化为数字信号,实现对较长钢丝帘布表面的连续性拍照,从而完成了钢丝帘布图像的采集工作。

## 1.2 缺陷检测流程及关键算法

为了实现钢丝帘布的缺陷的检测,首先通过图像采集系统采集到钢丝帘布原始图像(图像中红框标注了需要检测的缺陷,而实际图像中并没有进行该区域的标注)。采集到的钢丝帘布原始图像进行相应的预处理并保留钢丝帘布缺陷部分,同时对图像进行滤波处理,减少噪声对图像特征提取的影响。在此基础上,利用钢丝帘布缺陷检测模型(算法)对图像进行检测,以判断钢丝帘布是否具有缺陷以及缺陷位置。钢丝帘布缺陷大小、位置以及形状具有多样性,这也是钢丝帘布缺陷准确检测的主要难题。为保证在复杂的生产现场中实现钢丝帘布缺陷的准确检测,首先对采集的钢丝帘布缺陷数据进行数据增强,同时选择以 YOLOv5s 模型为基础并对其进行改进从而提升检测精度,缺陷检测工作流程如图 2 所示。

## 2 基于 DCGAN 的钢丝帘布缺陷数据增强

在深度学习中,充足的训练数据可以大大提高检测模型的性能。为了提高钢丝帘布缺陷检测的准确率,本文采用卷积生成对抗网络生成钢丝帘布缺陷图像,并将其加入原始数据集中以扩充数据集,有效的解决训练样本不足的

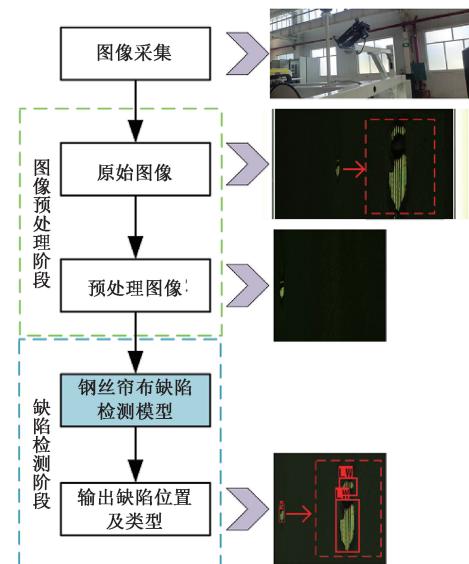


图 2 缺陷检测工作流程图

问题,从而提高 YOLOv5s 模型检测准确率及其泛化性能。

### 2.1 深度卷积生成对抗网络原理

深度卷积生成对抗网络是在 2015 年由 Alec Radford 等提出的一种模型,该模型以初始 GAN 为基础,通过结合卷积神经网络进行图像生成,并且提出了相应的一系列限制来提高网络的稳定性。

GAN 网络基本结构框图如图 3 所示,该网络由两个模型构成:一个是生成器模型(generative model, G),另一个是判别器模型(discriminative model, D),生成模型通过学习真实样本数据分布,将输入噪声(Noise)信号转化为图像数据。判别器模型的主要任务是判断给定的图像是真实数据(Real data)还是生成数据(Fake data),并输出相应的概率,这两个模型的网络参数会通过更新来不断优化。GAN 网络的训练目标是为了使生成模型和判别模型双方达到博奔论中的纳什均衡点(nash equilibrium, NE),此时生成模型的生成数据分布将与真实数据的分布相同。

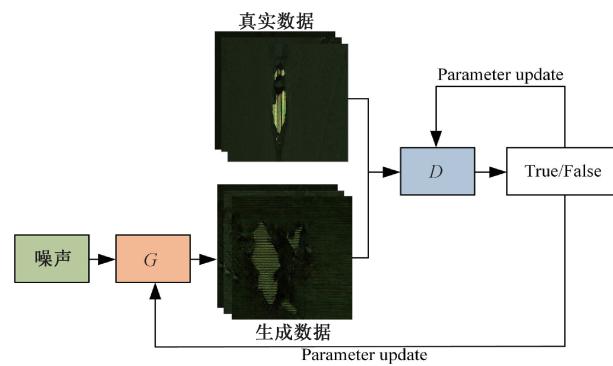


图 3 生成对抗网络结构示意图

生成器模型试图欺骗判别器模型,判别器模型则努力不被生成器模型欺骗。这两者模型经过多次交替优化训

练,使得两种模型的能力都能得到提升,最终得到效果提升很好的生成模型,使得这个生成模型所生成的钢丝帘布缺陷图像能达到真假难分的地步。GAN的目标函数如下所示。

$$\min_{G} \max_{D} V(D, G) = E_{x \sim p_{data}(x)} [\log D(x)] + E_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (1)$$

其中,  $V(D, G)$  表示目标函数,  $P_{data}(x)$  代表真实样本的分布,  $P_z(z)$  代表噪音样本的分布,  $G(z)$  代表生成器的输出,  $D(x)$  代表判别器对真实样本的结果,  $D(G(z))$  代表判别器对生成样本的结果。

从上式可以看出,通过  $G$  与  $D$  的对抗训练,  $G$  的训练目的是为了使得函数取得最小值,目的是要让生成样本更加接近于真实样本,即使得  $D(G(z)) = 1$ , 让  $E_{z \sim p_z(z)} [\log(1 - D(G(z)))]$  获得最小值,使  $G$  生成的样本更加接近于真实样本。反之  $D$  的训练目标是为了使得函数取得最大值,目的是让判别器区分真实样本和生成样本,即使得  $D(x) = 1, D(G(z)) = 0$ , 让  $E_{x \sim p_{data}(x)} [\log D(x)]$  与  $E_{z \sim p_z(z)} [\log(1 - D(G(z)))]$  均获得最大值,使  $D$  更好的区别真实样本与生成样本。

相比于传统GAN,DCGAN拥有多项优势:首先,DCGAN生成器和判别器去掉了卷积神经网络中的池化层,生成器使用转置卷积进行上采样,同时采用卷积层替换了所有的全连接层,从而转化为全卷积神经网络。其次,DCGAN生成器模型的输出层使用Tanh激活函数,其余层使用Relu激活函数,判别器模型输出层使用Sigmoid函数,其余层采用LeakyReLU激活函数。最后,DCGAN还使用了批量归一化(batch normalization, BN)应用在判别器和生成器的每一层中,从而加速收敛和减缓了过拟合。这些改变均使得DCGAN的训练更加稳定。综上所述,DCGAN可以更好的学习到钢丝帘布缺陷图像的特征,并生成更加清晰,更加真实的缺陷纹理。

## 2.2 DCGAN网络结构参数设置

在缺陷检测中,生成的钢丝帘布缺陷图像被用于提升检测准确率,为了将这些缺陷图像用于扩充YOLOv5s训练集,生成的缺陷样本必须具备适当的尺寸,并且需要具有较高图像质量,否则过小或质量较低的图像反而会影响检测的准确率。因此,针对钢丝帘布缺陷小样本问题,通过设置生成图像分辨率的大小、增加网络通道数并优化相关超参数从而生成符合要求的高质量钢丝帘布缺陷图像,用于提高YOLOv5s的检测准确率。

在DCGAN网络中,生成图像分辨率选择至关重要。过高的分辨率图像会导致网络训练缓慢,甚至会因计算机内存不足而无法进行训练;而分辨率过低会导致生成图像缺陷纹理不清晰,无法实现对钢丝帘布缺陷数据集的有效增强,最终影响检测模型对较小的缺陷识别。在生成对抗网络中,生成的缺陷图像与原始缺陷图像存在着对应关系,网络的输入是原始图像压缩后的图

像,其中原始图像中的缺陷也会等比例的缩小。而输出的生成图像与输入图像具有相同尺寸,并且具有与输入图像相似的数据分布。因此,通过多次的实验,得出了计算对抗生成网络生成图像分辨率大小的经验公式可表示为:

$$L \geq \frac{W \times N}{M} \quad (2)$$

其中,  $L$  代表生成图像的分辨率大小,  $W$  为原始缺陷图像分辨率的大小,  $M$  为原始图像中可识别最小尺寸缺陷的分辨率,  $N$  为该缺陷满足检测模型准确检测的最低分辨率。

例如,采集的原始缺陷图像分辨率为 2 000 pixel  $\times$  2 000 pixel,在该图像中,尺寸大小为 5 mm  $\times$  5 mm 的缺陷所占像素约为 2 500 个。检测模型检测一个尺寸大小 5 mm  $\times$  5 mm 的钢丝帘布缺陷需要达到 225 个像素才能较为准确的识别,则需要保证压缩后的图像分辨率不应低于 600 pixel  $\times$  600 pixel。通过对比不同分辨率图像缺陷的检测效果,检测效果对比如图 4 所示,当钢丝帘布缺陷图像的分辨率设置为 640 pixel  $\times$  640 pixel 时,缺陷纹理较为清晰且缺陷识别准确率明显较高。对于钢丝帘布帘线露出缺陷(leaky wire-steel, LW)而言,低分辨率的图像往往难以准确识别小缺陷。并且缺陷识别的置信度会随图像像素变低而下降,例如分辨率为 160 pixel  $\times$  160 pixel 的图像缺陷检测置信度为 0.63,而在分辨率为 640 pixel  $\times$  640 pixel 的图像中该缺陷检测置信度为 0.85。本例说明了该方法具有一定的可靠性,使用这种方法可以确保生成的钢丝帘布缺陷图像具有足够的纹理信息,并能够满足检测模型对较小缺陷的识别,并且在类似的数据增强中的对图像分辨率选择有一定的参考意义。因此,本文最终设置生成图像分辨率为 640 pixel  $\times$  640 pixel。

为了保证生成具有较高分辨率的钢丝帘布缺陷图像,通过增加网络通道数,提高网络对图像特征提取能力。生成器的详细参数如表 1 所示,200 维度的噪声输入经过线性层并且通过 View 函数转化为(40, 40, 2 048)的向量,随后,该向量经过批量归一化和 ReLU 函数激活,再通过反卷积层进行处理。反卷积核大小为 4  $\times$  4,步长为 2,前三层反卷积逐渐减小输出通道的维数至 256,每一层反卷积之后都进行了批量归一化和 ReLU 函数激活处理,经过最后的反卷积层,生成器以 Tanh 函数激活输出 3 通道钢丝帘布缺陷图像。

判别器的结构如表 2 所示,该结构包含一个图像输入层,4 个卷积层和一个全连接层,卷积核大小为 4  $\times$  4,步长为 2,钢丝帘布缺陷图像首先通过输入层输入到判别器中,经过 4 层的卷积层后通道数变为 2 048,每一层卷积层之后都进行了批量归一化和 LeakyReLU 函数激活处理,最后通过一个全连接层通过 Sigmoid 函数激活最终输出图像的真假信息。

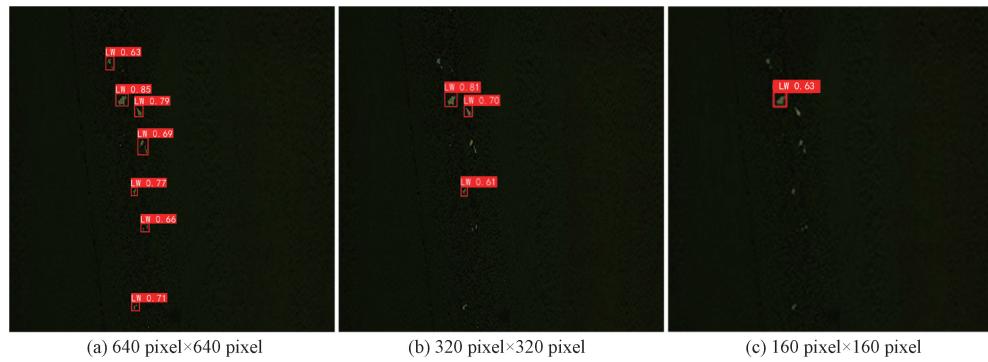


图 4 不同分辨率的缺陷检测对比图

表 1 生成器结构参数

操作类型	核尺寸	输出尺寸	输出通道数	激活函数
输入层	无	40×40	2 048	ReLU
反卷积	4×4	80×80	1 024	ReLU
反卷积	4×4	160×160	512	ReLU
反卷积	4×4	320×320	256	ReLU
反卷积	4×4	640×640	3	Tanh

表 2 判别器结构参数

操作类型	核尺寸	输出尺寸	输出通道数	激活函数
输入层	无	640×640	3	无
卷积层	4×4	320×320	256	LeakyReLU
卷积层	4×4	160×160	512	LeakyReLU
卷积层	4×4	80×80	1 024	LeakyReLU
卷积层	4×4	40×40	2 048	LeakyTanh
全连接层	无	1	无	Sigmoid

### 3 基于改进 YOLOv5s 的钢丝帘布缺陷检测

#### 3.1 YOLOv5s 网络模型介绍

YOLO 作为目前杰出的检测算法之一,在工业检测上的应用较为广泛<sup>[21-23]</sup>。YOLOv5 根据网络深度和宽度可将其划分了 5 个不同的版本,YOLOv5s 作为其中量级较小的一个版本,具有权重文件小、检测速度快、网络参数量较小等优势,YOLOv5s 的网络结构包含输入端(Input)、主干网络(Backbone)、颈部网络(Neck)和预测端(Prediction)。

输入端主要是用于原始钢丝帘布缺陷图像的数据集的进行 Mosaic 数据增强的策略、使用 K-means 聚类算法进行锚框计算;主干网络则是由 Conv、C3、和 SPPF 模块构成。Conv 模块是由卷积层、批量归一化层及激活函数组成,用于网络的非线性表达能力,在 6.2 版本中骨干网络在第一个 Conv 模块采用卷积核为 6,步长为 2 的替换 5.0 版本中的 Focus 模块,简化了网络结构,C3 模块采用多个

残差网络,增加了网络深度,并且可以减少模型计算量和提高推理速度。SPPF 模块通过 Concat 融合不同特征图,以实现目标的多尺度特征融合,从而扩大感受野;颈部网络的主要作用是对主干网络提取到的特征进行多次的融合,得到更加丰富的特征信息,并将这些特征信息传递到预测端;预测端是用于检测目标的位置和类别,网络最终输出不同尺度的特征图以适应不同大小的缺陷,同时在预测端应用完全交并比(complete intersection over union, CIoU)用于评估预测框与真实框的重合度,为边界框提供优化移动方向,最后采用非极大值抑制(non-maximum suppression, NMS)筛选多余的检测框,最终得到图像中缺陷检测框。

#### 3.2 改进的 YOLOv5s 模型

为了保证 YOLOv5s 模型在具有一定检测速度的前提下还能进一步提升钢丝帘布缺陷检测的准确率,首先,将输入端所采用的 K-means 算法改进为 K-means++聚类算法,从而生成更优的锚框尺寸参数;其次,将主干网络中 C3 模块中添加 CA 注意力机制,在此将该模块称为 C3CA 模块,用以增强主干网络对钢丝帘布缺陷特征提取能力以及对缺陷位置捕获能力,进一步提高缺陷识别准确率;最后,采用 MPDIoU 损失函数替代原 YOLOv5s 模型中预测端的 CIoU 损失函数,从而提高了检测框的定位精度,改进后的 YOLOv5s 网络结构如图 5 所示。

#### 3.3 K-means++算法提高锚框尺寸匹配度

YOLOv5 算法在训练和预测过程需要提前设置锚框的大小,根据预定义的锚框大小进行位置和类别判定,YOLOv5 默认的锚框是采用 K-means 算法在 COCO 数据上聚类得到的。而 K-means 算法在聚类之前首先需要随机初始化  $K$  聚类中心,随机性可能会导致初始聚类中心与实际最佳聚类中心位置有所差异,并且该方法不稳定,需要多次的聚类才能得到较稳定的聚类效果。同时,由于钢丝帘布缺陷数据集与 COCO 数据集的检测对象差异巨大,因此采用默认的锚框大小也是不适合本数据集的。

因此,本文采用了 K-means++算法对钢丝帘布缺陷进行重新聚类分析,该方法降低了 K-means 算法初始点可

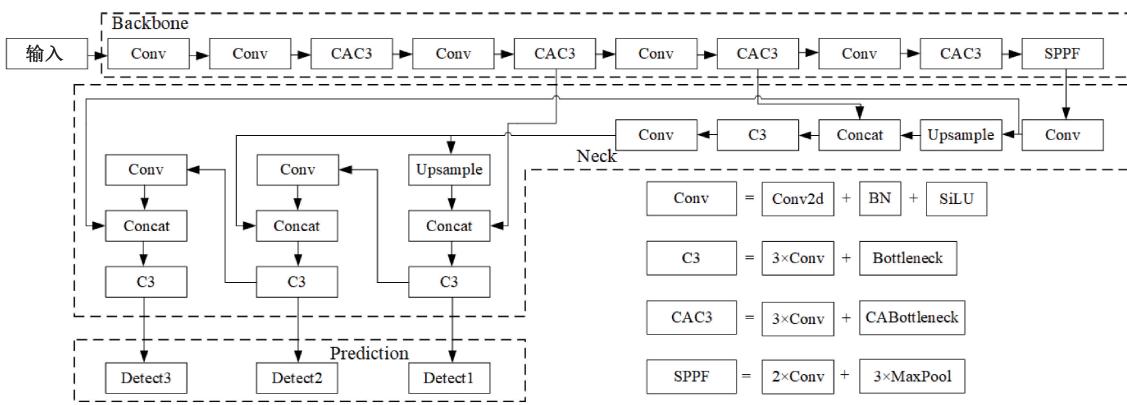


图5 改进后的YOLOv5s网络结构图

能会造成聚类偏差的问题,有效地弥补了原始算法的缺陷。通过使用该方法可获得更好的聚类效果,得到更优的锚框参数。K-means++算法的具体步骤如下:

1)从数据集中随机选取一个点作为初始的聚类中心  $u_1$ 。

2)计算每个数据样本与初始聚类中心的距离,用  $D(x)$  表示,并计算每个样本被选择为下一个聚类中心的概率  $P(x)$  通轮盘法选出下一个聚类中心点。

$$P(x) = \frac{D(x)^2}{\sum_{x \in X} D(x)^2} \quad (3)$$

3)重复步骤 2) 找出  $k$  个聚类中心点。

4)分别计算每个数据样本  $x_i$  到  $k$  个聚类中心的距离,并将该样本分配至距离最小的聚类中心所对应的类中,并针对每个类别  $C_i$ ,重新计算该类别的聚类中心  $u_i$ 。

$$u_i = \frac{1}{|C_i|} \sum_{x \in C_i} x \quad (4)$$

5)重复步骤 4) 直到聚类中心位置不再变化,最终得到  $k$  个聚类中心。

为了使得锚框与YOLOv5的多尺度检测头相匹配,将  $k$  值设置为 9,迭代 2 000 次后,最终生成 9 种尺度的锚框,锚框尺度分别为 [16, 23], [32, 35], [35, 95], [48, 419], [77, 73], [123, 29], [108, 534], [177, 154], [405, 72]。

### 3.4 主干网络的改进

注意力机制在目标检测算法领域有大量的应用<sup>[24-25]</sup>,通常情况下,大多数注意力机制采用加深网络深度的方法,这种方式会增加模型的计算量,可能会影响到模型的推理速度。在钢丝帘布缺陷的大小、形状、以及位置具有多样性,并且较小的缺陷像素较少,在训练过程中网络不易捕捉这类缺陷的特征,为了增强YOLOv5s网络对缺陷的特征提取能力,实现对缺陷更加精准的检测,本文在传统主干网络C3模块中添加了坐标注意力机制。

CA注意力机制模块弥补了挤压和激励注意力机制(squeeze-and-excitation, SE)在编码过程中只考虑通道信息而忽略位置信息的重要性不足之处,同时也弥补了卷

积注意力机制<sup>[27]</sup>(convolutional block attention module, CBAM)采用卷积计算空间注意力来利用位置信息时,只能捕获局部信息,而无法捕获长距离的依赖关系的劣势。CA注意力机制不仅能捕获跨通道的信息,还能捕获方向感知和位置感知的信息,这能帮助模型更加精准地定位和识别感兴趣的目标;该模块沿着水平和垂直的空间方向编码特征图的通道信息,可以使得 CA 注意力机制能够在通道内捕获长距离依赖关系,还能有助于保存精确的位置信息;并且该模块灵活且轻量,可以灵活的插入到移动网络中,比如 MobileNetV2、MobileNeXt 和 EfficientNet,并且没有过大的计算开销。

CA 注意力机制可以分为 Coordinate 信息嵌入和 Coordinate Attention 生成 2 个步骤。信息嵌入部分首先对输入特征图  $X$  分别采用  $(H, 1)$  和  $(1, W)$  的池化核对特征图的 X 和 Y 轴方向进行一维平均池化(adaptive average pooling),产生  $C \times H \times 1$  和  $C \times 1 \times W$  的两个形状的特征图,分别由式(5)与(6)可以得到。

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i < W} x_c(h, i) \quad (5)$$

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} x_c(j, w) \quad (6)$$

Coordinate Attention 生成部分将上述两个特征图按空间维度进行拼接,拼接成  $(C/r) \times 1 \times (W+H)$  形状的特征图,其中  $r$  用代表卷积的下采样比例,再经过 F1 卷积变换函数( $1 \times 1$  卷积)和非线性激活函数产生中间特征图  $f$ 。然后  $f$  将按通道维度拆分成两个张量  $f^h$  和  $f^w$ ,使得大小分别为  $(C/r) \times H \times 1$  和  $(C/r) \times 1 \times W$ ,接着将  $f^h$  和  $f^w$  分别进行  $1 \times 1$  卷积进行通道变换,使得大小变为  $C \times H \times 1$  和  $C \times 1 \times W$ ,并且经过 Sigmoid 激活函数,最终得到两个空间方向上的  $g^h$  和  $g^w$  坐标注意力,最后将  $g^h$  和  $g^w$  与原输入进行相乘,得到与输入相同形状的输出。CA 注意力机制模块结构如图 6 所示。

为了更好的利用坐标注意力机制的优势,本文将 CA 模块嵌入 C3 网络内部设计出 CAC3 模块,CAC3 模块结构如图 7 所示。

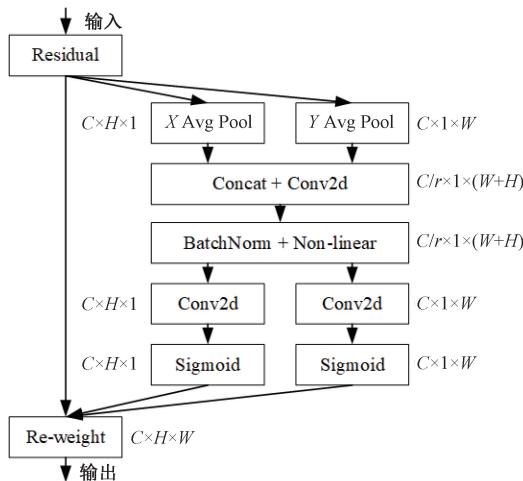


图 6 CA 注意力机制结构图

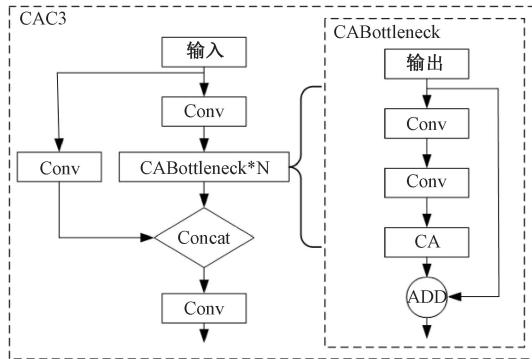


图 7 CAC3 模块结构图

### 3.5 损失函数的改进

在原 YOLOv5s 网络中,损失函数所采用的是 CIoU 损失函数,损失函数的表达如式(7)所示,在该公式中,  $A$ ,  $B$  代表了两个框,  $IoU$  表示真实框和预测框的交并比,  $A_{cr}$  和  $B_{cr}$  代表了  $A$  和  $B$  的中心点,  $c$  表示包围预测框和真实框最小矩形框的最短对角线的距离,  $v$  描述长宽比一致性的参数,  $\alpha$  是表示 trade-off 参数,其中  $v$  和  $\alpha$  的定义如式(8)和(9),其中,  $w^{gt}$  和  $h^{gt}$  表示真实框的宽度与高度,而  $w$  和  $h$  表示预测框的宽度与高度。

$$L_{\text{CIoU}} = 1 - IoU(A, B) + \frac{p^2(A_{cr}, B_{cr})}{c^2} + \alpha v \quad (7)$$

$$v = \frac{4}{\pi^2} \left( \arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (8)$$

$$\alpha = \frac{v}{(1 - IoU) + v} \quad (9)$$

虽然 CIoU 相对大多数的损失函数而言,考虑了检测框尺度、锚框的长度和宽度的比例、以及中心距离,使得检测框的回归损失函数更加稳定,不会像 IoU 损失函数在训练过程中可能现发散等问题。但是从公式中可以看出,当预测框和真实框的长宽比相同时,此时  $v=0$ ,长宽比的惩罚恒为 0,但此时可能存在长度和宽度的值不同,这将限制

收敛速度和精度。并且  $w$  和  $h$  相对于  $v$  的梯度是一对相反数,  $w$  和  $h$  不能同时增大或者减小。

为了解决上述问题,本文引入 MPDIoU 损失函数,该损失函数包含了现有损失函数中考虑的所有相关因素,如锚框的重叠或非重叠区域、中心点距离以及长宽比偏差。MPDIoU 损失函数原理如图 8 所示,通过直接最小化预测框和真实框的左上和右下的距离来简化计算过程,使得预测框更加接近于真实框,从而实现准确高效的边界框回归。MPDIoU 计算如下:

$$d_1^2 = (x_1^{prd} - x_1^{gt})^2 + (y_1^{prd} - y_1^{gt})^2 \quad (10)$$

$$d_2^2 = (x_2^{prd} - x_2^{gt})^2 + (y_2^{prd} - y_2^{gt})^2 \quad (11)$$

$$MPDIoU = \frac{A \cap B}{A \cup B} - \frac{d_1^2}{w^2 + h^2} - \frac{d_2^2}{w^2 + h^2} \quad (12)$$

其中,  $(x_1^{prd}, y_1^{prd})$  和  $(x_2^{prd}, y_2^{prd})$  表示预测框的左上角和右下角的坐标,  $(x_1^{gt}, y_1^{gt})$  和  $(x_2^{gt}, y_2^{gt})$  表示真实框的左上角和右下角的坐标,  $w$  和  $h$  表示输入图像的宽度和高度。

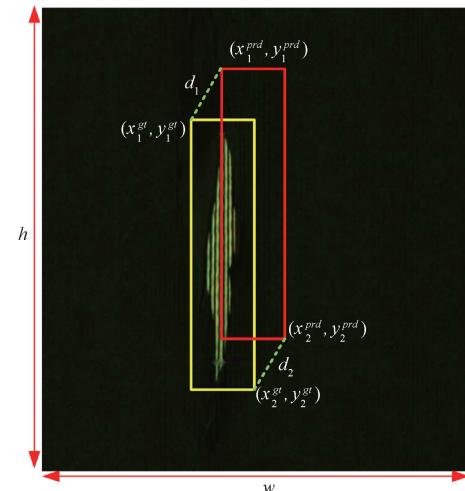


图 8 MPDIoU 原理示意图

MPDIoU 的损失函数通过预测框和真实框的最小距离来判断两个框的相似性,其中最小距离是由  $d_1$  和  $d_2$  所确定,因此该损失函数可以适应重叠或不重叠的边界框回归。在训练阶段,通过最小化损失函数迫使模型预测边界框不断的去接近其真实边界框。最终 MPDIoU 的损失函数定义如下:

$$L_{\text{MPDIoU}} = 1 - MPDIoU \quad (13)$$

通过最小化损失函数  $L_{\text{MPDIoU}}$  来使模型的预测框接近真实框,即是使得 MPDIoU 获得最大值,仅需  $d_1$  和  $d_2$  最小即可。该损失函数简化了计算过程,改善了缺陷检测训练效果,提升模型收敛速度及回归精度,从而提高了缺陷检测的效率和准确率,有助于提高检测模型的整体性能。

## 4 实验结果与分析

### 4.1 钢丝帘布缺陷数据集的制作

原始的钢丝帘布缺陷图像是在实际生产现场所采集的。对缺陷图像进行相应的预处理,使用图像分割技术获

取缺陷目标区域,制作成符合生成对抗网络和目标检测网络训练的数据集。通过筛选和预处理后,获得漏钢丝缺陷图像总共640张,图像分辨率为2000 pixel×2000 pixel。

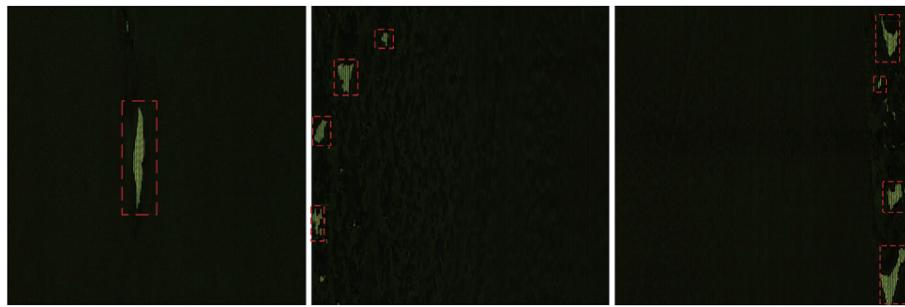


图9 预处理后的缺陷图像(缺陷使用了虚线框进行标注)

为了训练检测模型,使用了Labelimg工具对缺陷原始数据和生成数据集进行锚框标注,生成了xml类型的标准文件,以便进行网络框架的训练。并且将原始缺陷数据按照8:1:1划分为训练集,验证集和测试集,测试集图像仅用于检验最终缺陷检测模型检测的准确率。为了保证实验最终的测试集一致性,本文使用DCGAN网络生成钢丝帘布缺陷图像,对其进行标注后仅添加到训练集。综合考虑原始缺陷图像的数量,最终确定添加到训练集的生成图像数量为300张。

#### 4.2 DCGAN 对抗网络的训练

本实验将漏钢丝缺陷样本图像传入深度卷积生成对抗网络中进行训练。实验的硬件环境如表3所示。

表3 实验环境配置

平台	配置详情
CPU	Inter® Xeon® Gold5218CPU
GPU	NVIDIA RTX3090
内存	32 GB
显存	24 GB GDDR6X
环境配置	Python3.9.16 cuda11.7.1 pytorch1.13.1

在对抗网络的训练中,学习率作为网络重要超参数之一,它控制着网络梯度更新的量级和模型学习的进展速度。过小的学习率会导致网络收敛缓慢且容易陷入局部最优解,从而无法找到全局最优解;而过大的学习率则会使网络无法收敛至最优解。为了优化目标函数并在接近最优解时实现更好的性能,设置了学习率 $lr=0.002$ ,并且使用了余弦退火(cosine annealing)的学习率衰减。这种方法可以在梯度下降算法中动态地调整学习率,提高模型的训练效果。本实验将迭代次数(epoch)设置为600次,对于小样本的钢丝帘布缺陷图像,考虑输入单张图像的分辨率为640 pixel×640 pixel, batch\_size设置为8。实验所采用的梯度优化算法是Adam优化算法,其中一阶矩优化指数

预处理后的部分钢丝帘布缺陷图像如图9所示,虚线框标注的区域是需要进行检测的缺陷,而在原始图像中并没有该区域的标注。

衰减率为 $\text{beta1}=0.5$ ,二阶矩优化指数衰减率为 $\text{beta2}=0.999$ 。

#### 4.3 DCGAN 实验结果

为了更好地展示不同迭代次数下生成器生成的缺陷图像效果,本实验在每100轮迭代时保存了生成器生成的图像。图10展示了不同迭代次数下生成器网络生成的钢丝帘布缺陷图像的局部效果。从图中可以看出,在迭代100次时,生成器网络已经具备了生成缺陷图像的基本能力,但是缺陷纹理不够真实,生成图像过于模糊。在迭代300次时,生成器网络的能力得到了显著提升,基本掌握了缺陷纹理,但生成图像的质量还不够稳定。在迭代600次时,网络生成的图像已经非常稳定,生成图像的质量与迭代500次时相差不大,而且生成的图像缺陷具有多样性,缺陷纹理清晰,能够满足图像标注和目标检测的要求。

本文对真实图像和生成图像的局部缺陷图像进行了对比分析,结果如表4所示,需要说明的是,原始图像分辨率为2000 pixel×2000 pixel,而对抗网络生成的图像分辨率为640 pixel×640 pixel,因此生成的图像清晰度相对较低。然而,生成的图像与原始图像在纹理上具有很高的相似度,而且对抗网络生成器生成的图像种类丰富,对图像增强有较好的效果,可以用于扩充钢丝帘布缺陷样本数据集。

#### 4.4 YOLOv5 网络的训练

训练YOLOv5网络实验的硬件环境同上表3一致,将标注完成的数据集输入网络进行训练,本文输入网络的图片大小均设置为 $640\times640$ , epoch设置为300轮,设置batch\_size为16,初始学习率为0.01,并使用余弦退火的学习率调度策略。梯度优化算法采用SGD优化算法,动量参数为0.937,权重衰减参数为0.0005。

#### 4.5 评价指标

本文选择了准确率(precision,P)、召回率(recall,R)、平均精度(average precision,AP)和均值平均精度(mean average precision,mAP)作为模型检测精度的评价指标,同时也使用模型的参数量(parameters)作为模型大小的评价

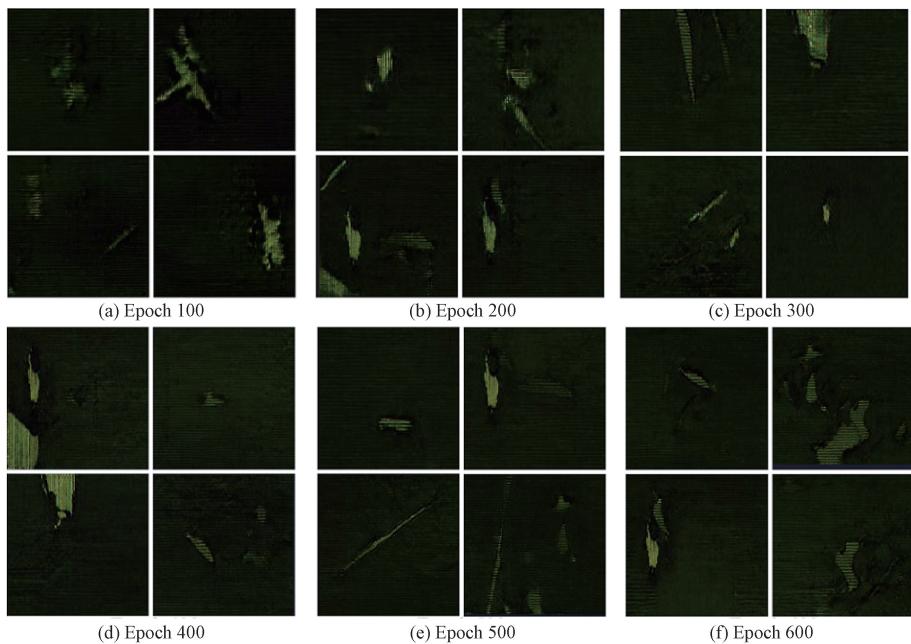


图 10 生成器生成图像

表 4 原始图像与生成图像对比

类型	漏钢丝缺陷(LW)		
原始图像			
生成图像			

指标,检测帧数(frames per second, FPS)作为衡量网络模型处理图像的速度评价指标。各项指标计算公式如下所示:

$$P = \frac{TP}{FP + TP} \quad (14)$$

$$R = \frac{TP}{TP + FN} \quad (15)$$

$$AP = \int_0^1 P(R) dR \quad (16)$$

$$mAP = \frac{1}{n} \sum_{i=1}^n AP_i \quad (17)$$

$$Parameters = H_K \times W_K \times C_{in} \times C_{out} \quad (18)$$

其中,  $TP$  表示被模型预测为正类的正样本,  $FP$  表示被模型预测为正类的负样本,  $FN$  表示被模型预测为负类的正样本,  $H_K$  代表卷积核的高,  $W_K$  代表卷积核的宽,  $C_{in}$  代表输入的通道数,  $C_{out}$  代表输出的通道数。

#### 4.6 实验结果对比分析

为了验证改进算法对原算法优化的有效性,本文使用相同的数据集,以 YOLOv5s 为基础网络,改进不同模块分别进行实验,实验结果如表 5 所示。相比使用原始缺陷数据训练的检测模型而言,通过添加由 DCGAN 模型生成的缺陷图像,YOLOv5s 检测模型的检测准确率  $P$  值提高了 1.5%,  $R$  值提高了 2% 且  $AP$  值提高了 0.9%; 使用 K-means++聚类算法优化 Anchor, 检测模型的  $P$ 、 $R$ 、 $AP$  分别提高了 2%、3.2% 和 2.8%, 说明使用聚类后的 Anchor 对缺陷检测的精度有所提高; 使用 CA 注意力机制模块改进原始 YOLOv5s 的主干网络,对钢丝帘布缺陷检测的准确率提升较为明显,  $P$  值提高了 3.6% 并且  $R$  值达到 80.7%,  $AP$  值也提升了 3%, 虽然该模块会稍微增加一部分的计算量,但并不影响整体的推理速度。值得注意的是,使用本文的 MPDIoU 损失函数替换原先的 CIoU 损失函数,虽然  $P$  值和  $AP$  值都有上升,但是  $R$  值略微有所下降,该损失函数对缺陷定位更加精确和严格,说明该损失函数具有提升回归精度的作用。在 DCGAN 数据增强的基础上增加 kmeans++算法生成锚框能够较为明显的提高检测的召回率,  $R$  值提高了 5.1%, 在此基础上增加 CA 注意力机制再一次提高了  $P$  值,  $P$  值达到 87.9%, 充分说明 CA 注意力机制可加强网络对钢丝帘布缺陷特征提取能力,并有效的提高检测模型的准确率。最后,将所有改进点一同加入网络训练,平均检测准确率可达到最大值 89.4%, 此外,相比原 YOLOv5s 模型准确率和查全率最终分别提升 3.8% 及 6.4%。通过该实验,证明了本文的 DCGAN 数据增强及改进的 YOLOv5s 模型能有效的提升钢丝帘布缺陷检测精度。

表5 消融实验结果对比

DCGAN数据增强	K-means++	CA注意力机制	MPDIoU	P/%	R/%	AP/%
—	—	—	—	82.9	77.3	82.8
✓	—	—	—	84.4	79.3	83.7
—	✓	—	—	84.9	80.5	85.6
—	—	✓	—	86.5	80.7	85.8
—	—	—	✓	85.9	74.8	82.9
✓	✓	—	—	82.4	82.4	84.1
✓	✓	✓	—	<b>87.9</b>	82.6	88.7
✓	✓	✓	✓	86.7	<b>83.7</b>	<b>89.4</b>

注:表中✓代表了应用了此项技术,加粗数据为最优值

为了满足工业现场钢丝帘布缺陷检测的实时性和准确性,并进一步验证本文提出的改进检测模型在钢丝帘布缺陷检测方面优于其他先进检测算法,使用同样的原始数据集,分别对主流的单阶段检测模型SSD, YOLOv4, YOLOv7进行训练和测试,并对其模型参数量和检测速度做出比较,实验结果如表6所示,通过比较分析可以得出,使用改进后的YOLOv5s检测模型AP值比YOLOv7检测模型高出3.3%,且准确率和查全率分别高出3.2%和7.1%。同时本文模型的参数量(Parameters)远低于YOLOv7,对设备计算性能要求较低,模型FPS高达90.65,具有更快的检测速度,可以较好的适用于高速生产线上钢丝帘布的缺陷的快速检测。

表6 不同模型性能比较

检测模型	P/%	R/%	AP/%	Parameters/M	FPS
SSD	83.2	68.2	79.8	23.75	67.65
YOLOv4	74.6	69.0	79.1	63.94	36.25
YOLOv7	83.5	76.6	86.1	37.19	50.71
本文方法	<b>86.7</b>	<b>83.7</b>	<b>89.4</b>	<b>7.04</b>	<b>90.65</b>

改进前后的YOLOv5s模型在钢丝帘布缺陷检测数据集上检测结果进行对比,如图11(a)和(b)所示,改进后的检测模型对较小的缺陷可有效的进行识别,而原YOLOv5s存在漏检的情况,并且改进后模型的预测框具有更高的置信度,相比之下,使用本文的方法可以有效的提高钢丝帘布的缺陷检测精度。

#### 4.7 公开缺陷数据集检测结果

为了进一步验证改进的YOLOv5s模型对其他缺陷检测的有效性,本文使用了东北大学发布的热轧带钢表面缺陷数据集(NEU-DET)进行了进一步测试,该数据集包含了裂纹(Crazing)、夹杂(Inclusion)、麻点(Pitted Surface)、斑块(Patches)、压入氧化铁皮(Rolled-in Scale)、划痕(Scratches)这6中缺陷,数据集总共1800张图像,图像分辨率为200×200。按照8:1:1划分训练集训练集、验证集、测试集。在不使用DCGAN数据增强的情况下,分别采用原YOLOv5s方法和本文改进的YOLOv5s方法进行

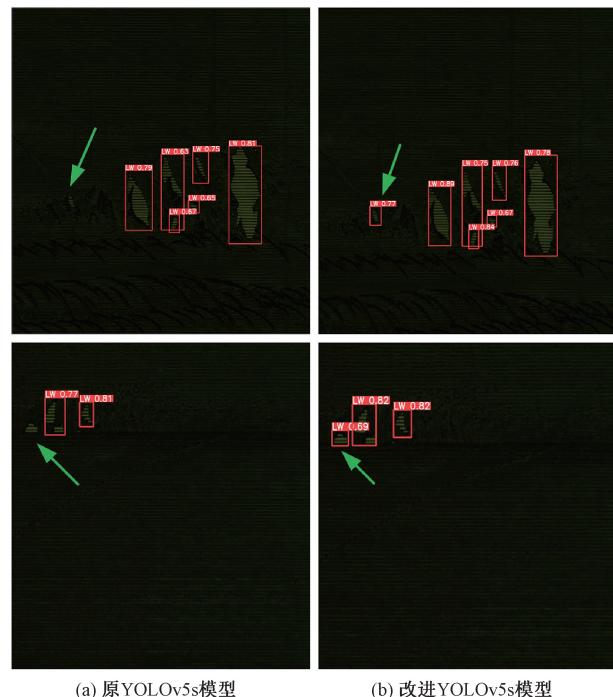
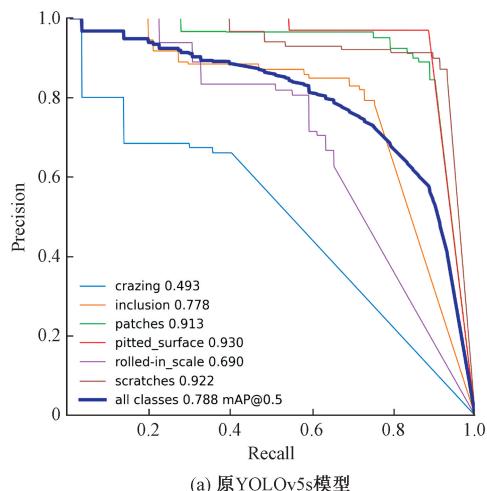


图11 改进前后YOLOv5s模型的检测结果对比图

训练,并在测试集上对模型检测效果进行对比,PR曲线对比如图12(a)和(b)所示,改进后的模型在该数据集各项



(a) 原YOLOv5s模型

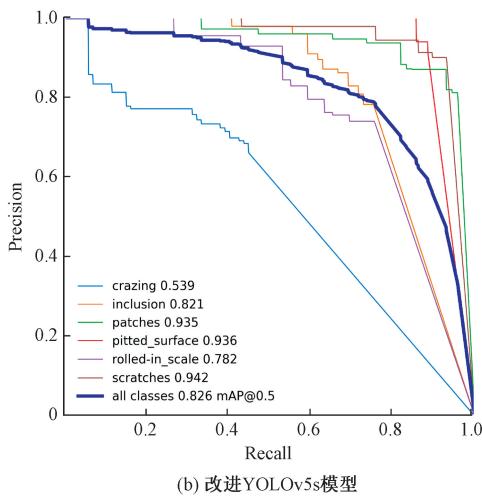


图 12 热轧钢表面缺陷检测结果 PR 曲线对比图

缺陷上的检测精度都有一定的提升,综合均值平均精度(mAP)从 78.8%提升至 82.6%。

## 5 结 论

针对钢丝帘布缺陷小样本而导致缺陷检测准确率低且泛化能力不强的问题。提出了一种基于 DCGAN 和改进的 YOLOv5s 模型的缺陷检测方法,首先,使用 DCGAN 网络对采集到的钢丝帘布缺陷数据进行数据增强,实现了对钢丝帘布数据的有效扩充;此外,使用 K-means++聚类算法生成更优的锚框参数,同时引入 CA 注意力机制改进 YOLOv5s 主干网络结构,增强网络对缺陷的特征提取能力;最后,采用 MPDIoU 的损失函数进一步精确检测框的位置,最终提高钢丝帘布缺陷的检测精度。

针对某生产线上的钢丝帘布实际缺陷数据,使用 DCGAN 数据增强和改进的 YOLOv5s 模型后,YOLOv5s 模型的平均精度提高了 6.6% 最终达到 89.4%,并且推理速度快,FPS 可达到 90 左右。在公开的热轧钢带表面缺陷数据集对改进后的 YOLOv5s 模型做了进一步测试,各类缺陷的检测精度均有提高,综合 mAP 值提高了 3.8%。实验结果充分体现了本文方法的有效性和可行性。

本文方法可以为同类缺陷检测问题提供思路和方法。同时针对少数未能检测出的钢丝帘布缺陷进行了分析,可能是由于该类缺陷样本数量过少,检测模型无法有效的捕捉该类缺陷特征。因此,在之后的工作中将会获取在不同工况下,更加全面的缺陷特征图像,并对模型做出进一步的改进和适配,从而进一步提高缺陷检测精度。

## 参考文献

- [1] BHATT P M, MALHAN R K, RAJENDRAN P, et al. Image-based surface defect detection using deep learning: A review [J]. Journal of Computing and Information Science in Engineering, 2021, 21(4):801-816.
- [2] 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,2022,43(1):198-219.
- [3] 杨永跃,夏远超. PCB 缺陷检测深度学习算法的精度改进[J].电子测量与仪器学报,2023,37(5):11-19.
- [4] GIRSHICK R, DONAHUE J, DARRELL T, et al. Region-based convolutional networks for accurate object detection and segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 38(1):142-158.
- [5] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]. European Conference on Computer Vision. Springer, Cham, 2016:21-37.
- [6] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 779-788.
- [7] 曾秀云,陆华才,吕禾丰.基于改进 Faster R-CNN 的棉布包装缺陷检测的方法研究[J].电子测量与仪器学报,2022,36(4):179-186.
- [8] 阎馨,杨月川,屠乃威.基于改进 SSD 的钢材表面缺陷检测[J].现代制造工程,2023(5):112-120.
- [9] 黄连,刘晓军,雷自力,等.基于奇异值分解的橡胶密封圈表面缺陷检测方法[J].润滑与密封,2021,46(11):84-88.
- [10] 朱焕宇,王明泉,李磊磊,等.胎体钢丝帘线缺陷的 CD-YOLO 分割算法应用[J].国外电子测量技术,2023,42(7):175-180.
- [11] 王素珍,许浩,邵明伟,等.基于改进 YOLOv4-Tiny 算法的绝缘子缺陷检测[J].国外电子测量技术,2022,41(9):155-162.
- [12] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks[J]. Communications of the ACM, 2020, 63(11): 139-144.
- [13] RADFORD A, METZ L, CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial networks [J]. ArXiv Preprint, 2015, ArXiv: 1511.06434.
- [14] WU Q F, CHEN Y P, MENG J. DCGAN-based data augmentation for tomato leaf disease identification[J]. IEEE Access, 2020, 8: 98716-98728.
- [15] 甘岚,沈鸿飞,王璐,等.基于改进 DCGAN 的数据增强方法[J].计算机应用,2021,41(5):1305-1313.
- [16] ZHANG L J, DUAN L X, HONG X C, et al. Imbalanced data enhancement method based on improved DCGAN and its application[J]. Journal of Intelligent & Fuzzy Systems, 2021, 41(2): 3485-98.

[17] LIKAS A, VLASSIS N, J VERBEEK J. The global k-means clustering algorithm[J]. Pattern Recognition, 2003, 36(2):451-461.

[18] 胡欣,周运强,肖剑,等.基于改进 YOLOv5 的螺纹钢表面缺陷检测[J].图学学报,2023, 44(3): 427-437.

[19] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13713-13722.

[20] MA S L, XU Y. MPDIoU: A loss for efficient and accurate bounding box regression[J]. ArXiv Preprint, 2023, ArXiv:2307.07662.

[21] YUE X, WANG Q, HE L, et al. Research on tiny target detection technology of fabric defects based on improved Yolo[J]. Applied Sciences, 2022, 12(13): 6823-6839.

[22] 彭道刚,潘俊臻,王丹豪,等.基于改进 YOLO v5 的电厂管道油液泄漏检测[J].电子测量与仪器学报,2022, 36(12):200-209.

[23] 霍超,谷晓钢,黄玲琴,等.基于 YOLO v4 优化的航拍绝缘子缺陷图像检测模型[J].电子测量技术,2023, 46(9):175-181.

[24] LUO X, NI Q, TAO R, et al. A lightweight detector based on attention mechanism for fabric defect detection[J]. IEEE Access, 2023, 11: 33554-33569.

[25] WANG X, GAO J S, HOU B J, et al. A lightweight modified YOLOX network using coordinate attention mechanism for PCB surface defect detection[J]. IEEE Sensors Journal, 2022, 22(21):20910-20920.

[26] LI W, LIU K, ZHANG L, et al. Object detection based on an adaptive attention mechanism [J]. Scientific Reports, 2020, 10(1):11307.

[27] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]. Proceedings of the European Conference on Computer Vision(ECCV), 2018:3-19.

### 作者简介

黄鹏(通信作者),博士,副教授,主要研究方向为机器视觉、缺陷检测、新型光学器件。

E-mail:2981020035@qq.com

蔡露,硕士研究生,主要研究方向为机器视觉、缺陷检测、深度学习。

E-mail:1131702069@qq.com

陈彬,硕士研究生,主要研究方向为光学测量原理,机器视觉。

E-mail:2085406797@qq.com

周益航,硕士研究生,主要研究方向为机器视觉、视觉测量。

E-mail:1667624249@qq.com

易冬旺,本科,高级工程师,主要研究方向为橡胶机械结构设计。

E-mail:409590512@qq.com