

# 基于改进 YOLOv7 的 PDC 钻头复合片检测<sup>①</sup>



陈琳国<sup>1</sup>, 熊 凌<sup>1</sup>, 代啟亮<sup>1</sup>, 王冬梅<sup>1</sup>, 李姝凡<sup>2</sup>

<sup>1</sup>(武汉大学 信息科学与工程学院, 武汉 430080)

<sup>2</sup>(中石化江钻石油机械有限公司, 武汉 430080)

通信作者: 熊 凌, E-mail: [xiongling@wust.edu.cn](mailto:xiongling@wust.edu.cn)

**摘 要:** 复合片是 PDC 钻头的核心切削单元, 复合片自动检测技术是复合片自动修复技术的基础. 本文提出了一种基于改进 YOLOv7 的 PDC 钻头复合片检测方法, 在 YOLOv7 的基础上, 用深度可分离卷积替换了常规卷积, 减少了参数量和运算成本; 引入了 SimAM 注意力机制, 不需要额外的参数便可以从神经元中推导出 3D 注意力权重, 而且还能提高卷积神经网络的表达能力; 用 SPPFCSPC 替换了 SPPCSPC, 在保证感受野不变的同时获得了速度的提升; 采用 K-means++ 算法聚类先验框, 使用启发式算法定位出缺损的复合片. 实验结果表明, 本文算法较原 YOLOv7 模型 *mAP* 提高了 2.75%, 参数量减少了约 80%, 推理速度提高了 9.12 f/s, 且较其他算法也有较大优势, 可实现复合片检测的工业应用.

**关键词:** PDC 钻头复合片; YOLOv7; 深度可分离卷积; SimAM 注意力机制; SPPFCSPC; K-means++

引用格式: 陈琳国, 熊凌, 代啟亮, 王冬梅, 李姝凡. 基于改进 YOLOv7 的 PDC 钻头复合片检测. 计算机系统应用, 2024, 33(2): 216–223. <http://www.c-s-a.org.cn/1003-3254/9427.html>

## PDC Drill Bit Composite Piece Detection Based on Improved YOLOv7

CHEN Lin-Guo<sup>1</sup>, XIONG Ling<sup>1</sup>, DAI Qi-Liang<sup>1</sup>, WANG Dong-Mei<sup>1</sup>, LI Shu-Fan<sup>2</sup>

<sup>1</sup>(School of Information Science and Engineering, Wuhan University of Science and Technology, Wuhan 430080, China)

<sup>2</sup>(Sinopec Jiang Diamond Oil Machinery Co. Ltd., Wuhan 430080, China)

**Abstract:** The composite piece is the core cutting unit of the PDC drill bit, and its automatic detection technology is the basis of the automatic repair technology of the composite piece. This paper proposes a PDC drill bit composite piece detection method based on the improved YOLOv7. Based on YOLOv7, the conventional convolution is replaced with depth-separable convolution, which reduces the amount of parameters and computing cost. As the SimAM attention mechanism is introduced, the method can derive 3D attention weights from neurons without additional parameters and also improve the expressive ability of convolutional neural networks. SPPCSPC is replaced with SPPFCSPC, which improves the speed while ensuring that the receptive field remains unchanged. The priori frames of K-means++ algorithm clusters are adopted and a heuristic algorithm is applied to locate defective composite pieces. Experimental results show that compared with the original YOLOv7 model, the *mAP* of the proposed algorithm is increased by 2.75%, the number of parameters reduced by about 80%, and the inference speed increased by 9.12 f/s. It also has greater advantages than other algorithms and can realize industrial applications of composite piece detection.

**Key words:** PDC bit composite piece; YOLOv7; depth separable convolution; SimAM attention mechanism; SPPFCSPC; K-means++

① 基金项目: 国家自然科学基金 (62173261); 湖北省重点研发计划 (2020BAB021)

收稿时间: 2023-07-28; 修改时间: 2023-09-01, 2023-09-26; 采用时间: 2023-10-25; csa 在线出版时间: 2023-12-25

CNKI 网络首发时间: 2023-12-27

复合片是 PDC 钻头的核心切削单元. 钝钻头的修复很大程度上降低了生产成本, 并且修复后的 PDC 钻头仍然具有不低于出厂效果的良好应用性能<sup>[1]</sup>. 为了实现 PDC 钻头复合片的自动修复, 首先需要在刀翼中检测出单个的复合片, 而当前修复钻头依靠人工判断失效情况, 效率较低. 本文提出了一种基于改进 YOLOv7 的 PDC 钻头复合片检测方法, 为后续的自动修复提供了有效的技术手段.

基于深度学习的目标检测方法研究始于 2014 年, Girshick 等<sup>[2]</sup>提出了第 1 个基于卷积神经网络 (CNN) 的目标检测方法 R-CNN, 该方法先在图像中提取候选区域, 然后对每个候选区域进行 CNN 特征提取和分类, 最后进行边框回归得到检测框. 2015 年在 R-CNN 的基础上, 相继诞生了 Fast R-CNN<sup>[3]</sup>和 Faster R-CNN<sup>[4]</sup>, 使得检测速度进一步提升. 同年, YOLO 的开篇之作 YOLOv1<sup>[5]</sup>诞生, 它通过将图像分割成网格来检测图片, 每个网格输出多个边界框和类别概率, 然后通过非极大值抑制过滤检测框. 2018 年的 YOLOv3<sup>[6]</sup>是 YOLO 系列的里程碑之作, 它采用了更深的网络, 并使用了残差结构来提高检测性能, 此外, 特征金字塔和多尺度检测的加入则是让它有了更高的检测精度. 2020 年的 YOLOv4<sup>[7]</sup>采用了更快的网络结构和更多的技术来提高检测性能. 同年的 YOLOv5<sup>[8]</sup>使用了更小、更快的网络结构, 并使用了自适应锚框来提高检测性能. YOLOv7<sup>[9]</sup>于 2022 年提出, 它通过增加训练成本来提升精度.

2020 年, Ashok 等<sup>[10]</sup>首次利用图像处理技术和深度学习对 PDC 钻头切削齿的损伤程度进行量化, 通过 YOLOv3 自动检测识别金刚石复合片; Ashok 等<sup>[11]</sup>开发了一种软件算法, 使用 YOLOv3 自动检测在钻井现场拍摄到的钻头图像中的复合片, 提高了对于复合片小目标的检测精度; Lu 等<sup>[12]</sup>建立了复合片的 Faster R-CNN 检测模型并进行训练, 提高了在复杂光环境下的检测性能.

目前, 在 PDC 钻头复合片检测中仍然存在如复合片缺损形态多变、刀齿背景对比度较低以及周边环境的干扰等问题, 这可能导致复合片定位失败, 从而影响正常的修复进程. 针对上述问题, 本文借助 YOLOv7 网络结构, 从提升网络的特征提取能力及速度, 解决背景杂乱、光照不均等问题对网络进行改进, 提高复合片检测的整体性能.

## 1 改进的 YOLOv7 检测算法

在 YOLOv7 的基础上, 为解决复合片检测过程中存在的检测速度及模型大小问题, 加入了深度可分离卷积和 SPPFCSPC, 减少了模型参数, 提高了检测速度. 针对复合片与刀齿背景颜色近似, 特征复杂等问题, 引入了 SimAM 注意力机制, 它可以重点关注复合片区域的信息, 提高模型对关键特征的识别能力. 而对于复合片目标框大小与传统目标框大小存在差异, 通过使用 K-means++ 重新聚类目标框, 可以加速网络收敛, 提高检测精度. 启发式算法可以在复合片缺失的情况下定位出缺失位置, 降低漏检率.

网络使用固定大小的图片训练, 用深度可分离卷积替换了常规卷积, 用 SPPFCSPC 替换 SPPCSPC, 加入了 SimAM 注意力模块, 具体的网络结构如图 1 所示.

### 1.1 深度可分离卷积

在改进的复合片检测网络中, 构建了 DSConv+BN 层+SiLU 激活函数的 DSBS 模块, 其中的 DSConv 为深度可分离卷积<sup>[13]</sup>, 深度可分离卷积的加入在保证精度的同时减少了参数量和运算成本. 深度可分离卷积与常规卷积的卷积效果是类似的, 但计算量却成倍减少.

深度可分离卷积主要分为两个过程: 逐通道卷积和逐点卷积. 在逐通道卷积中, 一个卷积核负责一个通道, 一个通道只被一个卷积核进行卷积, 经过逐通道卷积产生的特征图数量和输入图片的通道数保持一致. 逐点卷积与常规卷积基本类似, 它的卷积核大小为  $1 \times 1 \times M$  ( $M$  为上一层的通道数), 相当于是将输入在深度方向进行加权相加, 生成新的特征图, 其中特征图的数量由卷积核的数量决定. 深度可分离卷积的示意图如图 2 所示.

### 1.2 SimAM 注意力模块

空间注意力的结构如图 3 所示, 这些注意力模块一般会被插入到网络块中, 对先前层的输出在通道或空间维度进行细化操作, 生成一维或二维权重, 并对不同的神经元区别对待, 以起到改进前面网络层输出的作用.

为了更好地实现注意力机制, 应该根据每个神经元的重要性单独分配权重, SimAM 注意力<sup>[14]</sup>便可以很好地实现这些功能. 在神经科学中, 信息量不同的神经元与周围神经元的放电模式往往是不相同的, 而且在激活神经元的过程中周围的神经元也会受到不同程度

的抑制, 所以便需要赋予这些神经元更大的权值以抵消抑制作用. 在 SimAM 注意力机制中, 为了找到这些重要的神经元, 通过度量它们之间的线性可分性定义

了如下的能量函数:

$$e_t = (y_t - \hat{t})^2 + \frac{1}{M-1} \sum_{i=1}^{M-1} (y_0 - \hat{x}_i)^2 \quad (1)$$

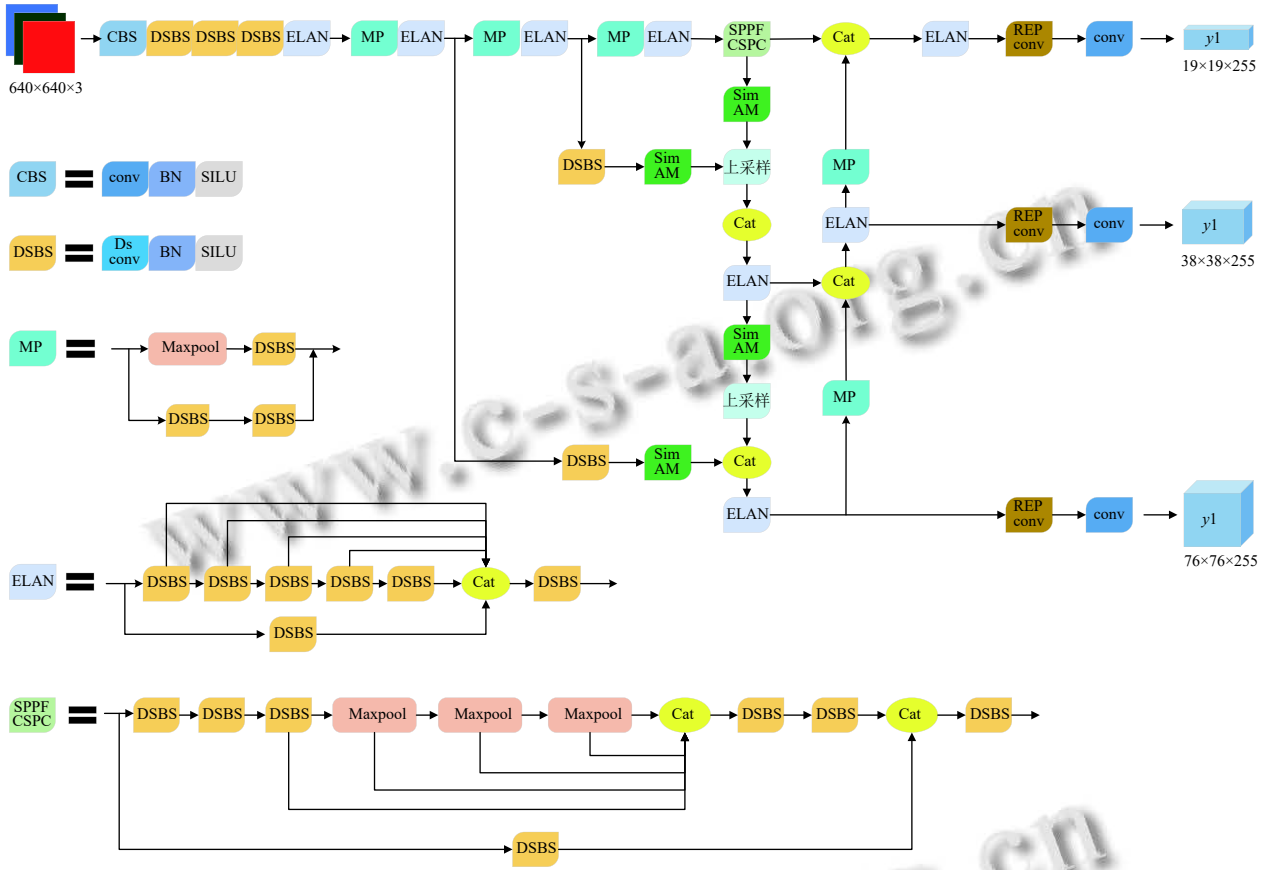


图1 改进的 YOLOv7 网络结构

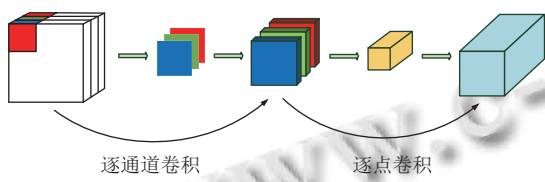


图2 深度可分离卷积示意图

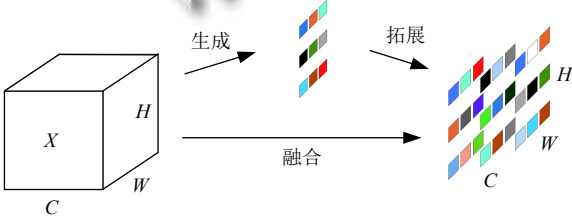


图3 空间注意力机制结构图

对式 (1) 最小化, 训练同一通道内某神经元与其他神经元的线性可分性, 并采用二值化标签, 添加正则项, 新的能量函数如下:

$$e_t = \frac{1}{M-1} \sum_{i=1}^{M-1} (-1 - (w_t x_i + b_t))^2 + (1 - (w_t t + b_t))^2 + \lambda w_t^2 \quad (2)$$

$$w_t = -\frac{2(t - \mu_t)}{(t - \mu_t)^2 + 2\sigma_t^2 + 2\lambda} \quad (3)$$

$$b_t = -\frac{1}{2}(t + \mu_t)w_t \quad (4)$$

其中,

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

$$\sigma_t^2 = \frac{1}{M-1} \sum_{i=1}^{M-1} (x_i - \mu_t)^2 \quad (6)$$

因此, 最小能量公式可以简化为:

$$e_t^* = \frac{4(\hat{\sigma}^2 + \lambda)}{(t - \hat{\mu})^2 + 2\hat{\sigma}^2 + 2\lambda} \quad (7)$$

通过能量函数可以看出, 神经元 $t$ 与周边神经元的区别 $\hat{\mu}$ 越大, 它的能量就越低, 重要性就越高, 所以便可以通过 $1/e^*$ 度量神经元的重要性. 按照注意力机制的定义, 需要对特征进行增强处理, 于是便得到:

$$\tilde{X} = \text{Sigmoid}\left(\frac{1}{E}\right) \odot X \quad (8)$$

相比 SE, ECA 等注意力机制, SimAM 注意力机制不仅性能更好, 拥有更强的鲁棒性, 而且不引入额外的参数, 保证了网络的运行速度. 它可以忽略掉不重要的背景信息, 快速定位出复合片的位置, 同时还会不断地调整权重, 对于不同复合片或不同环境条件下的复合

片, 都能够以高精度完成检测任务.

### 1.3 SPPFCSPC

在网络结构中, SPP 模块<sup>[15]</sup>的作用是通过最大化增大感受野, 以适应不同分辨率的图像. 而 CSP 模块则是将特征分为两部分, 其中一部分进行常规处理, 另一部分进行 SPP 结构的处理, 最后把这两个部分合二为一, 这样能减少一半的计算量, 使得速度更快, 精度更高. YOLOv7 中的 SPPCSPC 则是将 SPP 和 CSP 结构结合了起来. 本文在 SPPCSPC 的基础上, 借鉴 SPPF 的思想优化得到 SPPFCSPC, 在保证感受野不变的同时获得了速度的提升. SPPFCSPC 的结构如图 4 所示.

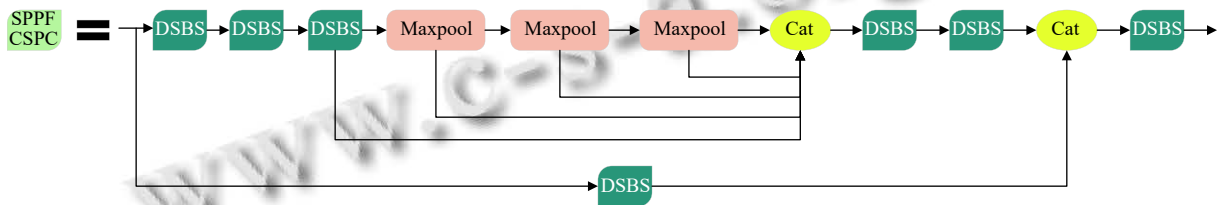


图 4 SPPFCSPC 结构图

在 SPPF 结构中, 首先对输入的特征图进行不同大小的池化操作, 然后对池化结果通过卷积进行融合. SPPF 结构可以自适应地融合不同尺度的特征信息, 以提高对不同形式缺损复合片的检测能力和检测速度.

### 1.4 K-means++

在复合片的检测任务中, 因为复合片的大小始终在一个固定的维度之间, 而 YOLO 网络的预设锚框则是适用于 COCO 数据集, 对于本文所使用的钻头数据集存在较大偏差, 不利于网络收敛. 在 YOLOv7 网络中, 使用了 K-means 算法默认对原始锚框进行聚类, 并在训练过程中自行调整锚框, 但是在聚类的过程中, K-means 在正式聚类之前首先要完成对 9 个簇中心的初始化, 这就使得网络的收敛情况严重依赖于簇中心的初始化状况, 如果初始化存在偏差将直接影响到网络收敛, 采用 K-means++ 算法便可以有效缓解这一问题, 从而在一定程度上提高复合片的检测精度和效果. 原始锚框、K-means 聚类锚框和 K-means++ 聚类锚框的对比如表 1 所示. 相比 K-means 算法, K-means++ 算法优化了第 1 步选择 9 个簇中心的过程, 使得随机选取的中心点不再只是趋于局部最优解, 而是寻找全局最优解. 在随机初始化 9 个簇中心的时候, 通过让各个簇的中心点之间尽可能远来保证它们不会位于同一个

簇内, 这会比 K-means 算法产生更好的结果. K-means++ 算法的具体步骤如下.

表 1 原始锚框、K-means 聚类锚框和 K-means++ 聚类锚框对比

锚框	大	中	小
原始锚框	10, 13, 16, 30, 30, 61	33, 23, 59, 62, 45	119, 116, 90, 156, 198, 373, 326
K-means 聚类锚框	22, 51, 27, 34, 41, 37	47, 62, 61, 83, 67, 49	86, 110, 87, 64, 113, 85
K-means++ 聚类锚框	17, 23, 23, 35, 30, 20	33, 44, 36, 28, 42, 55	49, 33, 60, 44, 66, 68

- (1) 从输入的点集中随机选取一个点作为聚类中心;
- (2) 对于点集中的每个点  $p$ , 计算它与最近的聚类中心之间的距离  $D(p)$ ;
- (3) 选择一个  $D(p)$  最大的点作为新的聚类中心;
- (4) 重复步骤 (2) 和步骤 (3) 直到 9 个聚类中心全部选出;
- (5) 用选出的 9 个聚类中心运行 K-means 算法, 选出最终的锚框.

最终的 K-means++ 聚类结果如图 5 所示.

### 1.5 启发式算法

为了解决缺失或者漏检复合片的问题, 本文提出一种启发式算法, 步骤如下.

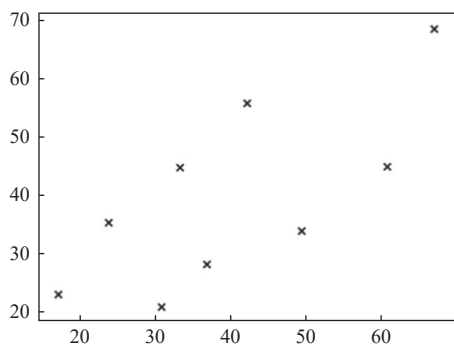


图5 K-means++聚类结果

(1) 获取每一个检测框的中心点坐标, 计算公式为:

$$center_x = top + \frac{bottom - top}{2} \quad (9)$$

$$center_y = left + \frac{right - left}{2} \quad (10)$$

其中,  $center_x$  为中心点的  $x$  轴坐标,  $center_y$  为中心点的  $y$  轴坐标,  $(top, left)$ ,  $(bottom, right)$  分别为检测框的左上角, 右下角坐标。

(2) 确定一个初始点  $P_1$ , 一般选取两端的点 (根据  $x$  坐标或者  $y$  坐标最大值选取初始点);

(3) 计算点  $P_1$  与其他点的距离, 选出最短距离的点  $P_2$  并连接;

(4) 计算点  $P_2$  与其他点 (除了已经筛选过的点) 的距离, 选出最短距离的点  $P_3$  并连接;

(5) 循环操作直到遍历完所有的点;

(6) 计算每两个相邻点之间的距离, 剔除最大值后求取平均值;

(7) 若某两个点之间的距离超过平均值一定范围 (根据钻头类型选取范围区间), 则判断此处存在缺损情况, 用红色线段表示。

启发式算法结果如图6所示, 图6中的红色线段处即为检测出的缺损位置。

## 2 实验与分析

本文的所有模型训练均是在 PyTorch 框架下完成的, 硬件环境为: CPU 为 AMD Ryzen 7 4800H, GPU 为 NVIDIA RTX 2060. 软件环境为 Windows 10、Python 3.7、PyTorch 1.7.0、cuda 10.1、cudnn 7.6.5.

### 2.1 数据集

实验的数据集来自生产现场的 PDC 钻头, 为了解决数据集过小的问题, 增强模型的泛化能力, 于是对数

据集进行了扩充. 本文采用亮度增强, 对比度增强、水平翻转、随机旋转、随机裁剪、添加噪声、颜色空间变换、图像模糊 8 种扩充方式, 扩充后共计 1800 张图像, 部分 PDC 钻头数据集扩充结果如图7所示。



图6 启发式算法结果图

### 2.2 评价指标

为了验证模型的有效性, 选取了平均精度均值 ( $mAP$ ), 召回率 ( $recall, R$ ), 准确率 ( $precision, P$ ), 参数量, 计算量和推理速度对网络的性能进行了测试。

在所有评价指标中,  $R$  是召回率, 其公式如下所示:

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

其中,  $TP$  是被正确划分为正样本, 在目标检测中指  $IoU > 0.5$  的目标框数量,  $FN$  是被错误划分为负样本, 在目标检测中指没有被检测到的目标框数量. 所以  $R$  指在所有真实目标中, 模型检测正确的目标框的占比。

$P$  是准确率, 公式如下:

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

其中,  $FP$  是被错误划分为正样本, 在目标检测中指  $IoU \leq 0.5$  的目标框数量. 所以  $P$  指在所有检测结果中, 模型检测正确的目标框的占比。

$mAP$  是平均精度均值, 指所有种类的  $AP$  均值, 在本文背景下指复合片检测中  $P-R$  (precision-recall) 曲线包围的面积. 此外,  $IoU$  是指所有检测出来的复合片中检测框与真实框的交并比, 公式为:

$$IoU = \frac{\text{检测框} \cap \text{真实框}}{\text{检测框} \cup \text{真实框}} \quad (13)$$

检出率指检测出来的复合片占有所有复合片的占比。

参数量指网络结构中所有参数的数量总和,用来衡量模型的空间复杂度.计算量指模型推断时需要的计算

次数,用来衡量模型的时间复杂度.推理速度指1s预测的图片数量,是单张图片预测时间的倒数.

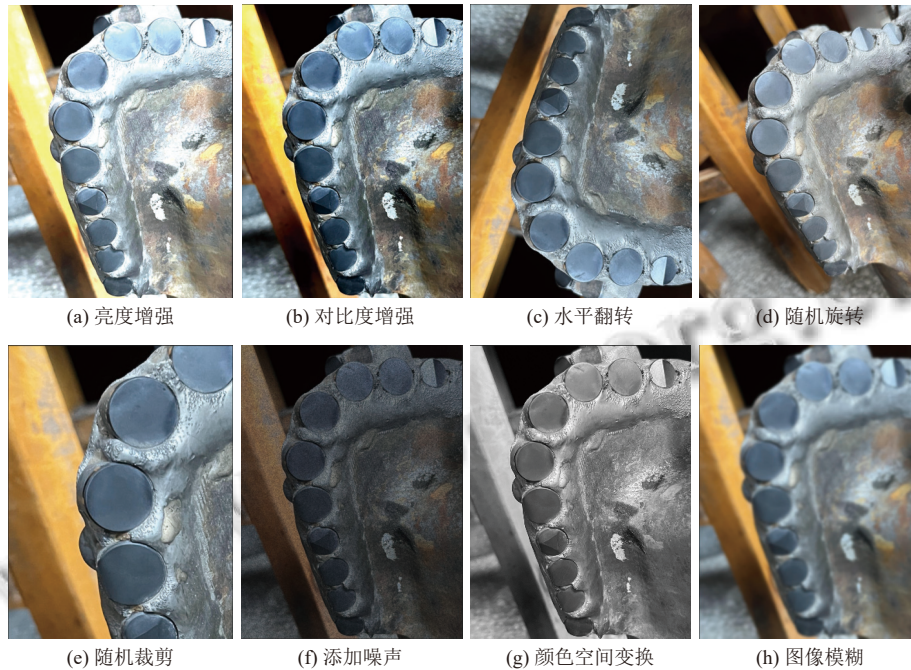


图7 部分数据扩充结果

### 2.3 消融实验

为了更加全面综合地分析改进YOLOv7网络中各个模块对复合片检测的性能提升,本文设计了消融实验.首先,将各个模块引入YOLOv5,验证新加入模块在不同的网络模型中仍然具有良好的性能,实验结果如表2所示.其次,将深度可分离卷积,SimAM注意力机制,SPPFCSPC,K-means++聚类分别加入到YOLOv7中,与原网络作为对照,具体的实验内容与结果如表3所示.

表2 新增模块在YOLOv5中的性能表现

模型	<i>mAP</i> (%)	参数量	推理速度 (f/s)
YOLOv5	80.12	$7.02 \times 10^6$	66.31
YOLOv5+新模块	83.78	$1.10 \times 10^7$	40.47

表3 消融实验对比

深度可分离卷积	SimAM注意力	SPPFCSPC	K-means++	<i>mAP</i> (%)	参数量	推理速度 (f/s)
—	—	—	—	83.81	$3.72 \times 10^7$	56.35
√	—	—	—	82.62	$7.49 \times 10^6$	60.42
—	√	—	—	85.77	$3.72 \times 10^7$	55.83
—	—	√	—	84.39	$3.72 \times 10^7$	62.65
—	—	—	√	85.45	$3.72 \times 10^7$	56.49
√	√	√	√	86.56	$7.49 \times 10^6$	65.47

从表2可以看出,新增模块在YOLOv5中也具有良好的性能表现,因为SPPFCSPC在原YOLOv5网络中并不存在,所以该模块的加入导致网络参数增加,推理速度也降低了,但是*mAP*却有3.66%的增长,提升了网络的检测精度.

从表3可以看出,各个新增模块对网络的性能都有不同程度的提升.深度可分离卷积的加入虽然小幅降低了网络的*mAP*,但却减少了80%的参数,降低了模型的大小.SimAM注意力机制使得网络的*mAP*提高了1.96%.SPPFCSPC的引入使得网络的推理速度提高了6.3 f/s,并且精度也有提升.K-means++聚类后的锚框因为更适合复合片大小,所以对精度的提升有着积极的影响,相比K-means聚类*mAP*提高了1.64%.由此可见,各个改进模块的加入均不同程度地提高了网络的性能,取得了良好的检测效果.

为了验证启发式算法的有效性,在相同环境下做了对比实验,实验结果如表4所示.

表4 启发式算法实验对比

算法	检出率 (%)
无启发式算法	86.42
加入启发式算法	94.72

通过实验可以看出,启发式算法的加入可以提高检出率,解决漏检或缺失的问题。

### 2.4 性能对比

本文共进行了5组对比实验,训练时的 batch\_size 为 8,使用 SGD 优化器优化网络参数,没有设置预训练权重,从 0 开始训练.各网络的检测结果如图 8

所示。

从图 8 可以看出,本文方法较 YOLOv7 能够捕捉到更多的暗部细节,可以检测出底部不明显的复合片,同时目标框更接近复合片大小,不存在过多冗余或者出界的情况,且明显优于其他网络模型.提取出的单个复合片如图 9 所示。

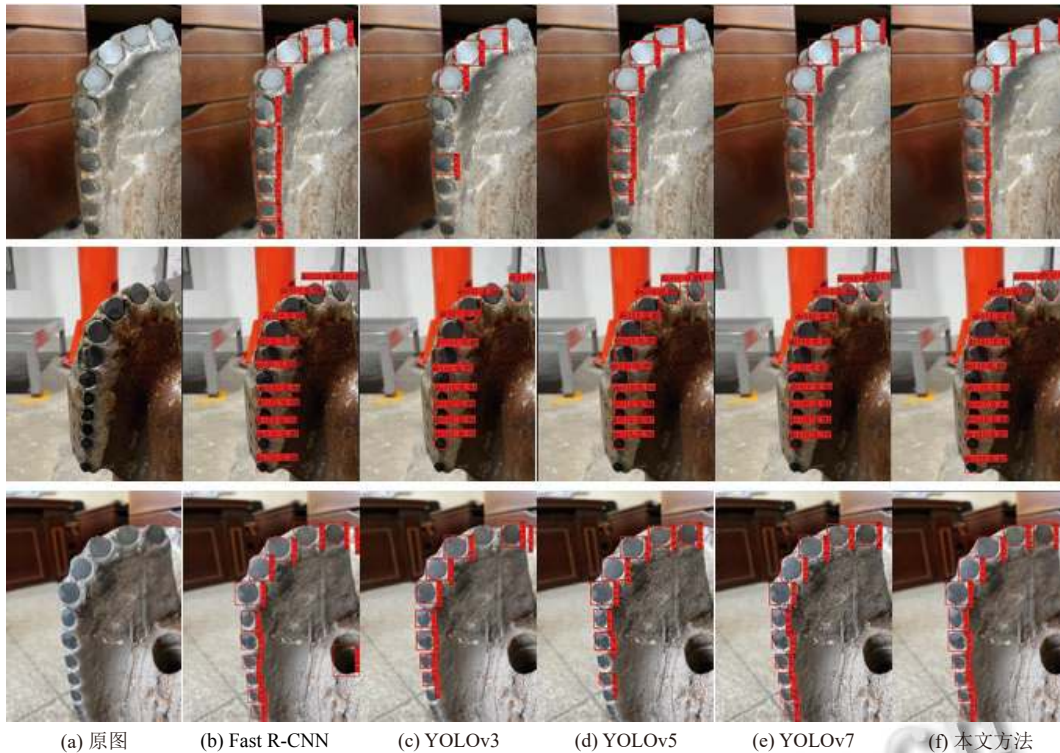


图 8 复合片检测结果对比图

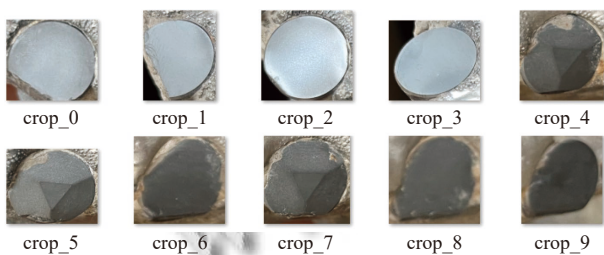


图 9 单个复合片提取图

从图 9 可以看出,提取出的单个复合片定位准确,不明显的边缘区域也能准确识别出来,方便后续的修复操作.实验的各项性能指标对比如表 5 所示。

从表 5 可以看出,本文方法的各项指标较 YOLOv7 均有提升, *mAP* 和推理速度各提高 2.75% 和 9.12 f/s,同时参数数量和计算量也分别降低了 80% 和 90.26 GFLOPs,参数数量的降低可以简化模型,使整个模型获得更好的可解释性,更有利于模型的推广和使用。

表 5 复合片检测不同网络检测结果对比

指标	Fast R-CNN	YOLOv3	YOLOv4	YOLOv5	YOLOv7	本文方法
<i>mAP</i> (%)	78.32	75.99	78.45	80.12	83.81	86.56
<i>R</i> (%)	83.95	68.41	70.13	82.30	81.55	83.67
<i>P</i> (%)	56.52	73.49	77.35	77.92	82.63	83.45
<i>IoU</i> (%)	87.37	82.33	88.16	91.85	92.64	94.67
检出率 (%)	89.63	70.34	83.69	91.07	92.46	94.72
参数量	$1.37 \times 10^8$	$6.15 \times 10^7$	$6.39 \times 10^7$	$7.02 \times 10^6$	$3.72 \times 10^7$	$7.49 \times 10^6$
计算量 (GFLOPs)	369.72	65.60	59.95	15.95	106.47	16.21
推理速度 (f/s)	24.34	60.40	41.82	66.31	56.35	65.47

### 3 总结

在复合片自动修复应用中,对复合片检测的准确性和实时性都有较高的要求.本文在传统 YOLOv7 网

络的基础上,加入了深度可分离卷积、SimAM 注意力机制、SPPFCSPC,并且使用 K-means++对原始锚框进行了重新聚类,通过启发式算法定位出漏检的复合片. 相较传统的 YOLOv7 网络,改进后的网络在降低模型大小的同时,提高了网络的准确率和特征提取能力,提高了背景杂乱、光照不均等问题下的检测精度,可满足复合片检测需求. 但该方法仍然存在不足,例如准确性有待提高,对缺损部分过多的复合片容易漏检等问题. 在后续工作中,将研究如何汇总深度学习更多的优秀算法和网络结构,进一步提高网络的检测准确性和实时性,以及针对漏检情况提出更加高效准确的解决方案.

### 参考文献

- 1 陈楠. 浅析钢体 PDC 钻头修复与应用. 西部探矿工程, 2022, 34(9): 112–114. [doi: 10.3969/j.issn.1004-5716.2022.09.039]
- 2 Girshick R, Donahue J, Darrell T, *et al.* Rich feature hierarchies for accurate object detection and semantic segmentation. Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Columbus: IEEE, 2014. 580–587. [doi: 10.1109/cvpr.2014.81]
- 3 Girshick R. Fast R-CNN. Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV). Santiago: IEEE, 2015. 1440–1448. [doi: 10.1109/iccv.2015.169]
- 4 Ren SQ, He KM, Girshick R, *et al.* Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137–1149. [doi: 10.1109/tpami.2016.2577031]
- 5 Redmon J, Divvala S, Girshick R, *et al.* You only look once: Unified, real-time object detection. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016. 779–788. [doi: 10.1109/cvpr.2016.91]
- 6 Redmon J, Farhadi A. YOLOv3: An incremental improvement. arXiv:1804.02767, 2018.
- 7 Bochkovskiy A, Wang CY, Liao HYM. YOLOv4: Optimal speed and accuracy of object detection. arXiv:2004.10934, 2020.
- 8 Jocher G. YOLOv5. <https://github.com/ultralytics/yolov5>. (2020-08-09)[2022-03-23].
- 9 Wang CY, Bochkovskiy A, Liao HYM. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver: IEEE, 2023. 7464–7475. [doi: 10.1109/cvpr52729.2023.00721]
- 10 Ashok P, Vashisht P, Kong H, *et al.* Drill bit damage assessment using image analysis and deep learning as an alternative to traditional IADC dull grading. SPE Annual Technical Conference and Exhibition. Denver: OnePetro, 2020. [doi: 10.2118/201664-MS]
- 11 Ashok P, Chu J, Witt-Doerring Y, *et al.* Drill bit failure forensics using 2D bit images captured at the rig site. Proceedings of the 2021 SPE/IADC International Drilling Conference and Exhibition. Stavanger: OnePetro, 2021. [doi: 10.2118/204124-MS]
- 12 Lu BP, Xu T, Huang YB, *et al.* Applications of computer vision and deep learning in visual features extraction of drill bits. Proceedings of the 2022 International Petroleum Technology Conference. Riyadh: OnePetro, 2022. [doi: 10.2523/IPTC-22624-MS]
- 13 Chollet F. Xception: Deep learning with depthwise separable convolutions. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017. 1800–1807. [doi: 10.1109/cvpr.2017.195]
- 14 Yang LX, Zhang RY, Li LD, *et al.* SimAM: A simple, parameter-free attention module for convolutional neural networks. Proceedings of the 38th International Conference on Machine Learning. Vienna: PMLR, 2021. 11863–11874.
- 15 He KM, Zhang XY, Ren SQ, *et al.* Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904–1916. [doi: 10.1109/TPAMI.2015.2389824]

(校对责编: 牛欣悦)