

## 基于改进 EfficientDet 的线束端子线芯检测算法

朱世松<sup>1\*</sup>, 孙秀帅<sup>1</sup>, 赵理山<sup>1</sup>, 芦碧波<sup>1</sup>, 姚东林<sup>2</sup><sup>1</sup>河南理工大学计算机科学与技术学院, 河南 焦作 454003;<sup>2</sup>空军装备部驻武汉地区第一军事代表室, 湖北 武汉 430000

**摘要** 线束端子压接高度不当会导致线芯被剪断或线芯之间空隙较大,人工检测线芯数量和判断端子压接后线芯是否断裂,存在劳动强度大、易产生视觉疲劳等问题,因此提出一种基于深度学习的线束端子显微图像线芯检测算法,用于线芯计数。针对端子线芯显微成像呈密集且不规则排列的特性,使用 K-means 多维度聚类算法对线芯边界框进行聚类,生成与线芯边界框相匹配的锚框;为处理端子图像中不同属性锚框类别极不均匀的问题,利用梯度均衡机制重构损失函数。与多个目标检测算法的对比实验结果表明,所提改进算法在兼顾实时性的基础上平均精度均值达 96.2%,能保持与人工同样的计数精度,可用于线束端子线芯计数和后续的压接质量评估等工作。

**关键词** 深度学习; 目标检测; 线束端子; 梯度均衡机制; 显微成像

中图分类号 TP391.4

文献标志码 A

DOI: 10.3788/LOP202259.1815008

## Detection Algorithm of Wire Harness Terminal Core Based on Improved EfficientDet

Zhu Shisong<sup>1\*</sup>, Sun Xiushuai<sup>1</sup>, Zhao Lishan<sup>1</sup>, Lu Bibo<sup>1</sup>, Yao Donglin<sup>2</sup><sup>1</sup>School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo 454003, Henan, China;<sup>2</sup>The First Military Representative Office of the Air Force Equipment Department in Wuhan, Wuhan 430000, Hubei, China

**Abstract** The improper crimping height of the wire harness terminal leads to the wire core being cut off or having a large gap. In manual detecting number of wire cores and judging whether the wire core is broken after terminal crimping, there are many challenges, such as high labor intensity and visual fatigue. In this paper, a wire core detection method based on deep learning in the microscopic image of the wire harness terminal for counting is proposed. K-means multidimensional clustering algorithm was used to cluster the wire core bounding boxes to generate the anchor boxes that conform to the distribution of the core bounding boxes, aiming at the properties of dense and irregular arrangement of terminal wire core microimaging. The gradient equalization mechanism was used to reconstruct the loss function to deal with the extremely uneven category of anchor boxes with different attributes in the terminal image. The results show that the algorithm proposed in this paper achieves a mean average precision of 96.2% while maintaining the real-time performance and the same counting accuracy as the manual, compared with the other object detection algorithms. The proposed algorithm can be used for wire core counting and crimping quality evaluation of wire harness terminals.

**Key words** deep learning; object detection; harness terminal; gradient harmonizing mechanism; microscopic imaging

## 1 引言

线束是电气设备中电流和信号传输的载体,广泛应用于汽车、轮船、航天航空及大型工业设备中。端子作为线束生产过程中的重要连接组件,压接工艺的优

劣显著影响线束的机械强度和电气性能,端子压接模具设定的压接高度应当既能避免单根线股因变形不均而被剪断的可能,又应保证线束截面中无明显空隙<sup>[1]</sup>。端子压接过程中为防止压接高度过小造成大量端子内部线芯断裂,统计端子样件线芯数目是端子截面分析

收稿日期: 2021-07-09; 修回日期: 2021-07-19; 录用日期: 2021-08-02

基金项目: 国家自然科学基金(41773024,U1404103)、河南省高等学校重点科研项目资助计划(21A520016)、2021年度河南科技智库调研课题(HNKJZK-2021-56C)

通信作者: \*zss@hpu.edu.cn

中的重要工序。为能够清晰地观察线束端子剖面,专业人员在专用剖面分析设备上对待测端子截面依次进行切片、抛光、蚀刻及酸洗等操作,最后放在显微镜下拍摄和观察,分析端子显微成像中的面积、内外径轮廓及线芯数量,以评估端子的导线致密度。

人工测量的方式存在较大的局限性,难以满足快速的端子质检需求。为实现线束端子智能检测,文献[2]结合机器视觉研发线束端子在线检测系统,用于测量端子截面压接宽度和两侧翼片角度等部分几何量。文献[3]使用水平集方法分割线束端子剖面内外径轮廓,便于后续端子截面数据提取和质量检测。文献[4]针对线束端子压接后外观缺陷,提出一种以线芯轮廓最小外接矩形数目为特征的视觉检测算法。文献[5]和文献[6]分别以偏微分理论和变指数色度滤波结合 Canny 算子边缘检测方法分割端子剖面轮廓,解决了传统检测方法中端子截面边界轮廓提取不准确的问题。据当前资料研究分析,大部分端子截面分析工作的研究着重于端子线芯内外径轮廓的数据提取,主要方法包括阈值分割、边缘检测及数学形态学处理等,关于端子截面线芯智能计数的研究却鲜有报道。研究一种快而准的线芯计数方法对实现线束智能化质检及生产具有重要价值。

由于存在线芯数量多、形状各异、边界模糊等情况,郭芳等<sup>[7]</sup>采用人机交互的方式进行线芯计数。线芯之间由于相互粘连,若采用实例分割计数需精细绘制每条线芯的形状轮廓,任务量大,且无法保证模型分割效果。基于密度图估计的计数模型适用于群体数目

估计<sup>[8]</sup>,并不适用于端子截面线芯的精确计数工作,且模型的实时性有待考究。近年来,随着 R-CNN<sup>[9]</sup>、Fast R-CNN<sup>[10]</sup>、Faster R-CNN<sup>[11]</sup>、Mask R-CNN<sup>[12]</sup>、SSD<sup>[13]</sup>、YOLO<sup>[14-16]</sup>等基于深度学习的目标检测算法的发展,研究人员将计数问题转换为目标检测问题,这些算法在羊群数量估计<sup>[17]</sup>、细胞有丝分裂分析<sup>[18]</sup>、小麦产量统计<sup>[19]</sup>、密集型钢筋计数<sup>[20]</sup>等工作中取得了较好效果。

受上述方法的启发,为解决当前线束端子线芯计数及质量评估等技术难题,本文将线芯计数问题转换为线芯目标检测问题,并基于 EfficientDet<sup>[21]</sup> 构造了一种线束端子显微图像线芯检测算法,用于线束端子截面分析工作。根据线芯检测结果可实现线芯的自动计数,以评估端子是否满足压接规范和技术标准。实验结果验证了所提算法的可行性与高效性。

## 2 面向线束端子显微图像的线芯检测算法

### 2.1 算法流程及原理

考虑实际应用中端子检测精度要求较高而计算资源有限,选取 EfficientDet-D0 网络模型进行线芯检测。针对端子剖面中线芯呈密集分布且不规则排列而难以检测的问题,在 EfficientDet 基础上,根据端子显微图像中线芯成像特性,使用 K-means 边界框聚类算法设定与线芯相匹配的锚框长宽比,引入梯度均衡机制构造损失函数,解决端子显微成像中大量的易分类样本和离群样本在训练时主导梯度更新方向的问题。所提算法流程如图 1 所示。

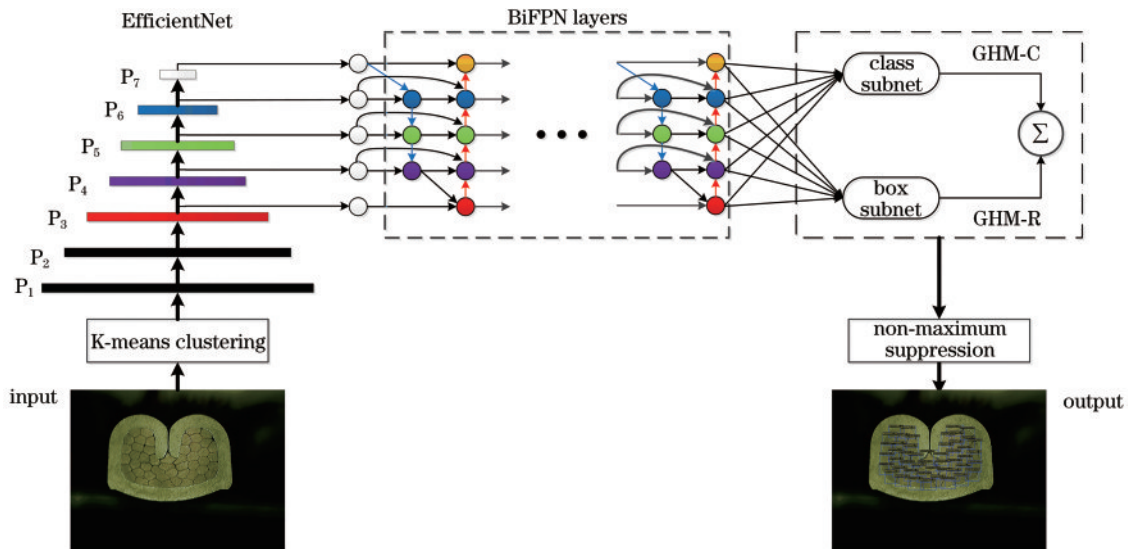


图 1 基于改进 EfficientDet 的线束端子显微图像线芯检测流程

Fig. 1 Core detection process of harness terminal microscopic image based on improved EfficientDet model

模型主要由主干特征提取网络、特征融合网络、类别和边界框预测子网络构成。BiFPN 特征融合模块用于处理物体检测时多尺度变化的问题,通过在 PANET<sup>[22]</sup> 的基础上去除对特征融合贡献较小的节

点,在同一层输入节点和输出节点之间添加一条额外的跳跃连接边,跨尺度的特征信息能够得到充分利用。

完成锚框参数的优化后,首先将端子显微图像送入高效的 EfficientNet 主干网络提取数据特征,经过不

断重复下采样和通道维数扩张后,获得  $P_1$ 、 $P_2$ 、 $P_3$ 、 $P_4$  和  $P_5$  五个特征层。由于  $P_1$  和  $P_2$  层生成于网络的初始阶段,携带语义信息较弱,后续并未使用  $P_1$  和  $P_2$  层作为有效特征层。 $P_5$  层经过两次下采样后分别得到  $P_6$  和  $P_7$  层,共获得  $P_3$ 、 $P_4$ 、 $P_5$ 、 $P_6$  和  $P_7$  五个有效特征层,然后对五个有效特征层反复应用 BiFPN 进行加权特征融合。为了准确预测目标的边界框,在获得五个不同具有高级语义信息和空间信息的有效特征层后, EfficientDet 将每个特征层划分为  $N \times N$  个网格,以每个网格为锚点,生成 9 个不同大小的锚框,即预设的边界框。图 2(a) 为分辨率为  $512 \times 512$  的原始图像经特征提取后生成的  $8 \times 8$  大小特征图,该特征层中的 anchor box 数量为  $8 \times 8 \times 9 = 576$ ,图 2(b) 为  $4 \times 4$  大小的特征图,该层的锚框数量为  $4 \times 4 \times 9 = 144$ 。

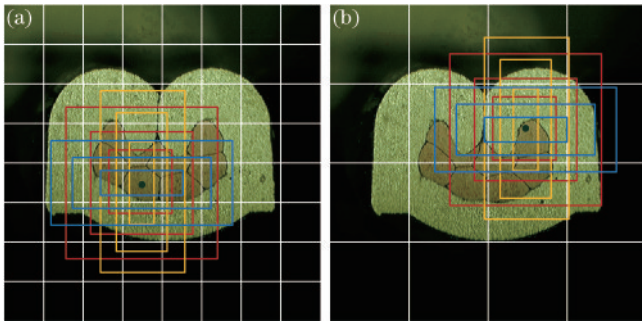


图 2 锚框生成示例。(a)  $8 \times 8$  大小特征图; (b)  $4 \times 4$  大小特征图  
Fig. 2 Examples of anchor box generation. (a)  $8 \times 8$  feature map; (b)  $4 \times 4$  feature map

本文将 GHM-C 损失函数用于锚框分类, GHM-R 损失函数用于锚框回归,计算损失之和并进行反向传播,完成网络对线束端子线芯的拟合。将待检线束端子图像送入拟合后的网络中,若锚框包含待检目标,且置信度较高,则通过边框的中心和宽高偏移量回归预测结果,对锚框的中心以及宽高坐标进行调整,使其逼近目标的真实边界框。经非极大值抑制,消除冗余的边界框后获得最终的预测边界框,完成线芯的识别与定位。

## 2.2 线芯边界框聚类

线束端子的特性使线芯集中于端子图像的中心区域,线芯形状各异,呈不规则排列。而 EfficientDet 模型预测目标边界框时依赖于预先设定的锚框,若锚框的尺寸与线束端子中的线芯真实边界框差异较大,则会影响模型的收敛速度和检测效果。故需重新对线束端子显微图像的预设锚框长宽比进行设计。

基于 K-means 算法对线束端子显微图像数据集中的边界框进行聚类。K-means 多维度聚类方法既能有效地保持锚框与 ground-truth 的精确匹配程度,又能避免额外的计算开销。K-means 边界框多维度聚类过程如下。

1) 随机选取样本中的  $K$  个真实边界框作为初始

化簇心,即 anchor box。

2) 分别计算样本中每个真实边界框与  $K$  个 anchor box 之间的距离,将距离簇心最近的样本划分为该类。距离度量公式为

$$D(b, c) = 1 - \text{IOU}(b, c), \quad (1)$$

其中交并比 (IOU) 公式为

$$\text{IOU}(b, c) = S_{b \cap c} / S_{b \cup c}, \quad (2)$$

式中:  $S_{b \cap c}$  表示聚类中心与真实边界框相交区域的面积;  $S_{b \cup c}$  表示聚类中心与真实边界框并集区域的面积;  $b$  为样本中的真实边界框;  $c$  为聚类过程产生的簇心。交并比用于度量产生的聚类中心与真实边界框之间的重叠程度。

3) 计算所在簇里所有真实边界框参数的均值,更新簇心。

4) 重复步骤 2) 和 3), 直至簇心不再变化, 聚类完成。

实验中以簇心和所在簇中所有边界框的平均交并比衡量本次聚类结果的准确率, 准确率越高, 则代表所得锚框与真实边界框匹配程度越好。

## 2.3 梯度均衡机制

端子线芯显微图像中锚框可划分为正负样本和难易样本, 线束端子图像背景占据了大量的空间, 只有少量的锚框中存在线芯目标。候选区域中的难易样本数量极度不平衡会导致大量的易分负样本参与训练, 主导梯度更新的方向, 这些负样本对模型的准确率贡献极小, 甚至还可能引起模型退化。Focal Loss<sup>[23]</sup> 在交叉熵损失 (CE Loss) 的基础上降低置信度较高样本的权重, 使模型关注难分样本。CE Loss 和 Focal Loss 的表达式分别为

$$L_{\text{CE}}(p, y) = \begin{cases} -\log p, & y = 1 \\ -\log(1 - p), & y = 0 \end{cases}, \quad (3)$$

$$L_{\text{F}}(p, y) = \begin{cases} -\alpha(1 - p)^\gamma \log p, & y = 1 \\ -(1 - \alpha)p^\gamma \log(1 - p), & y = 0 \end{cases}, \quad (4)$$

式中:  $p \in [0, 1]$  表示模型预测样本属于某一类的概率值;  $y \in \{0, 1\}$  表示样本的期望输出值;  $\alpha$  和  $\gamma$  为一对可调节因子, 其中  $\alpha \in [0, 1]$ , 用于平衡正负样本, 焦参数  $\gamma$  平滑调整易分样本的权重。

Focal Loss 引入调节因子来平衡正负样本以及难易样本。然而在线束端子数据集中, 静态的损失并不能随着训练过程的进行而不断适应不同数据分布变化的情况, 调节因子需要花费大量的时间来确定其最优解。且模型若过多地关注离群点, 反而会造成过拟合现象, 降低模型的准确性。为解决此问题, 本文引入 gradient harmonizing mechanism (GHM)<sup>[24]</sup> 重构模型损失函数, 根据样本的梯度范数分布减小易分样本和离群点对模型的影响, 其中 GHM-C Loss 用于分类,

GHM-R Loss 用于边界框回归。定义的梯度范数  $g$  为

$$g = |p - y| = \begin{cases} 1 - p, & y = 1 \\ p, & y = 0 \end{cases}, \quad (5)$$

$g$  值的大小用于定义样本分类的难易属性。定义梯度密度  $D_G(g)$  来衡量一定范围内梯度范数的样本数量, 表达式为

$$D_G(g) = \frac{1}{l_\epsilon(g)} \sum_{n=1}^N \delta_\epsilon(g_n, g), \quad (6)$$

$$\delta_\epsilon(x, y) = \begin{cases} 1, & y - \frac{\epsilon}{2} \leq x < y + \frac{\epsilon}{2} \\ 0, & \text{otherwise} \end{cases}, \quad (7)$$

$$l_\epsilon(g) = \min\left(g + \frac{\epsilon}{2}, 1\right) - \max\left(g - \frac{\epsilon}{2}, 0\right), \quad (8)$$

式中:  $g_n$  表示第  $n$  个样本的梯度范数;  $\delta_\epsilon(g_n, g)$  表示  $N$  个样本中梯度范数分布在区间  $\left[g - \frac{\epsilon}{2}, g + \frac{\epsilon}{2}\right)$  内的样本个数;  $l_\epsilon(g)$  表示以  $g$  为中心, 偏移量为  $\epsilon$  的区间长度; 第  $n$  个样本的梯度密度调节参数  $\beta_n = \frac{N}{D_G(g_n)}$ 。所使用的 GHM-C Loss 和 GHM-R Loss 的表达式分别为

$$L_{\text{GHM-C}} = \frac{1}{N} \sum_{n=1}^N \beta_n L_{\text{CE}}(p_n, y_n) = \sum_{n=1}^N \frac{L_{\text{CE}}(p_n, y_n)}{D_G(g_n)}, \quad (9)$$

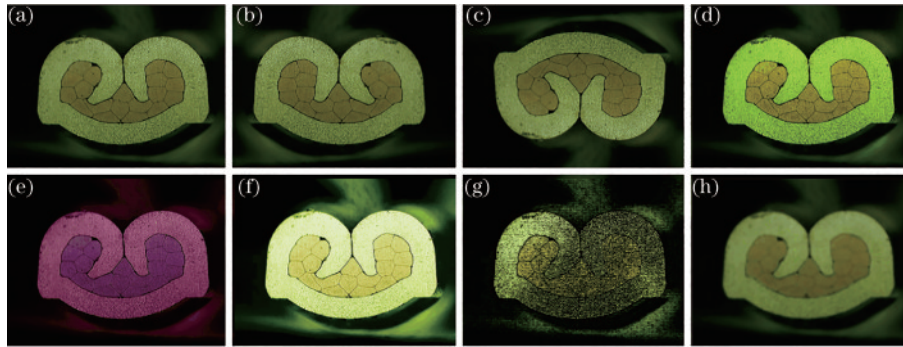


图 3 线束端子显微图像数据增强示例。(a)原图;(b)水平翻转;(c)垂直翻转;(d)饱和度增强;(e)色调增强;(f)亮度增强;(g)椒盐噪声;(h)高斯噪声

Fig. 3 Examples of data augmentation for microscopic images of wire harness terminals. (a) Original photo; (b) flip horizontal; (c) flip vertical; (d) saturation enhancement; (e) hue enhancement; (f) intensity enhancement; (g) salt and pepper noise; (h) Gaussian noise

增强后端子显微图像数据集信息如表 1 所示。

为保持与 EfficientDet 算法特征层相同的锚框覆

表 1 线束端子显微图像数据集的图片数量

Table 1 Number of images in the harness terminal microscopic image dataset

Dataset	Original	Data augmentation	Core of cable
Train	385	2582	77696
Test	96	96	3074
Total	481	2678	80770

$$L_{\text{GHM-R}} = \frac{1}{N} \sum_{n=1}^N \beta_n L_{\text{ASL}}(d_n) = \sum_{n=1}^N \frac{L_{\text{ASL}}(d_n)}{D_G(N_{\text{gr}})}, \quad (10)$$

$$L_{\text{ASL}}(d) = \sqrt{d^2 + \mu^2} - \mu, \quad (11)$$

式中:  $d = t_i^{(*)} - t_i$ ;  $t = (t_x, t_y, t_w, t_h)$  表示期望的边框中心坐标及宽高偏移量;  $t^{(*)} = (t_x^{(*)}, t_y^{(*)}, t_w^{(*)}, t_h^{(*)})$  表示模型

预测的边框中心坐标及宽高偏移量;  $N_{\text{gr}} = \left\lfloor \frac{d}{\sqrt{d^2 + \mu^2}} \right\rfloor$

定义为  $L_{\text{ASL}}$  的梯度范数。ASL<sub>1</sub>(authentic smooth L<sub>1</sub>) 为改进后的 SL<sub>1</sub>(smooth L<sub>1</sub>) 损失;  $\mu = 0.02$ 。

### 3 实验结果与分析

实验环境为: Ubuntu 20.04 操作系统, Intel(R) Core(TM) i9-10900K CPU@3.7 GHz 中央处理器, 64G RAM, NVIDIA GeForce RTX 3090 图形处理器, CUDA Version 11.2, Python 3.8, Pytorch 1.8.0 深度学习框架。

#### 3.1 数据集及实验设置

实验所用线束端子显微图像均是在专用线束端子截面分析仪上采集拍摄的, 共获得原始数据集共 481 张。随机选取 80% 作为实验所用训练集, 余下 20% 作为测试集。由于线束端子显微图像获取流程繁琐且原始数据有限, 对训练集随机采取几何变换和颜色变换等数据增强方式进行扩充, 以提高模型的泛化能力。线束端子显微图像数据增强示例如图 3 所示。

盖密集程度, 进行端子线芯边界框聚类实验时, 在不改变每个特征层预设锚框尺度个数与大小基础上, 取  $K=3$ , 聚类出 3 种符合线芯锚框形状分布的长宽比。对 3 种不同尺度的锚框设定 3 种长宽比, 特征层上的每个锚点便可生成 9 种符合线芯边界框形状分布的锚框。对线束端子数据集样本反复进行聚类实验 10 次, 所得聚类准确率最高的三组锚框长宽分别为 (36, 38)、(58, 64)、(46, 49), aspect ratio 为 0.9、0.93、0.95, 准确率为 81.91%。优化线束锚框初始长宽比后, 引入梯

度均衡机制对线束端子显微图像数据集进行拟合。批量大小设为 8, 迭代次数为 20000, 采用分段常数学习率衰减策略, 动量参数 0.9, 权重衰减参数 0.0005, 采用 stochastic gradient descent (SGD) 进行损失优化, 损失不再下降, 则训练结束。训练过程中学习率和损失变化曲线如图 4 所示。模型训练时总损失包含锚框分类损失和边框回归损失, 损失函数分别为 GHM-C 和

GHM-R。由图 4 可看出, 模型由于采用梯度均衡策略统计数据集中每类样本梯度范数分布的比例, 平衡了数据集中正负样本以及难易样本占损失函数的权重, 使得模型训练更加高效和稳健。在训练初期, 模型的分损失和回归损失均快速下降, 模型在收敛时, 总的平均损失值为 0.159681, 其中分类损失值为 0.027298, 边框回归损失值为 0.132383。

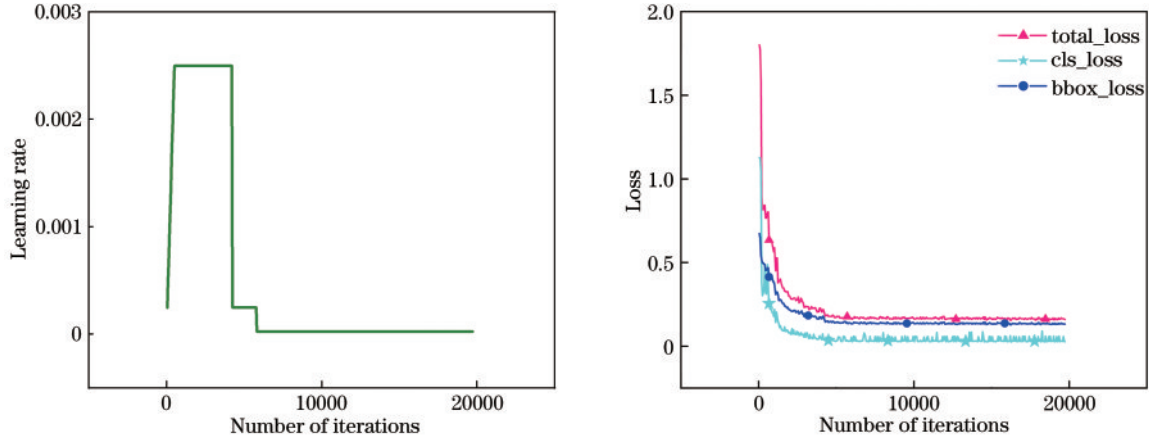


图 4 学习率及损失变化曲线

Fig. 4 Learning rate and loss curves

### 3.2 模型评估指标

使用平均精度均值 (mAP)、frames per second (FPS) 和召回率 (recall) 等常用指标来评估模型性能的优良。其中 mAP 是模型在多个检测类别上平均精度 (AP) 的均值; FPS 为模型单位时间内处理图像的总数量, 用于衡量模型的推理速度; recall 定义为所有目标都被模型检测到的比例, 用于衡量模型的查全率。

$$P_{\text{precision}} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FP}}} = \frac{N_{\text{TP}}}{N_{\text{detection}}}, \quad (12)$$

$$R_{\text{recall}} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}}} = \frac{N_{\text{TP}}}{N_{\text{ground-truth}}}, \quad (13)$$

式中:  $N_{\text{TP}}$  表示模型检测正确的线芯数量, 即与真实框  $\text{IOU} > 0.5$  的锚框;  $N_{\text{FP}}$  表示检测错误的线芯数量, 即与真实框  $\text{IOU} < 0.5$  的锚框;  $N_{\text{FN}}$  表示模型漏检的线芯数量。以 recall 为横轴, precision 为纵轴, 则 AP 即为 P-R 曲线下面积, 表示召回率在 0 至 1 之间对应精度的均值, 公式为

$$P_A = \int_0^1 P_{\text{precision}}(R_{\text{recall}}) dR_{\text{recall}} \quad (14)$$

$\sum_{c=1}^C P_A(c)$  表示所有类别的平均精度总值,  $C$  为所检测目标的总类别数, 则 mAP 表示为

$$P_{\text{mAP}} = \frac{\sum_{c=1}^C P_A(c)}{C} \quad (15)$$

### 3.3 定性分析

为验证所提基于 EfficientDet 改进的线束端子显微图像线芯检测算法的检测效果, 通过线束端子显微

图像测试集来评估 EfficientDet 和改进后 EfficientDet 模型的性能, 并同主流的单阶段目标检测算法 YOLOv3、RetinaNet 和 NAS-FPN<sup>[25]</sup> 进行对比。图 5 为不同方法在部分测试集上的放大后检测效果。

图 5(a)~(c) 为 7 芯线束端子, 图 5(d) 为 20 芯线束端子。YOLOv3 在图 5(a) 和图 5(c) 中的定位效果好, 但在图 5(b) 和图 5(d) 中发生了漏检, 在图 5(b) 中只检测出了其中 5 根线芯, 漏检 2 根, 在图 5(d) 中实检 16 根, 漏检 4 根。YOLOv3 虽然也引入了多尺度特征融合的策略, 并使用 DarkNet-53 主干网络提取图像特征, 但其只在  $13 \times 13$ 、 $26 \times 26$ 、 $52 \times 52$  三个有效特征层上进行边框分类及回归, 导致对密集小目标的检测效果不理想。EfficientDet 则分别在  $P_3$ 、 $P_4$ 、 $P_5$ 、 $P_6$  和  $P_7$  五个加强提取后的有效特征层进行最终预测。较小特征层上的元素感受野较为广阔, 用于捕捉输入图像中大尺寸目标的特征, 相反, 较大特征层上元素的感受野较为窄小, 则用于捕捉小尺寸的目标。EfficientDet 感受野较为细腻, 有效减小了对密集型目标的检测难度。然而原 EfficientDet 虽然没有出现漏检的情形, 但在对四幅图像的检测中都存在少量重复检测的错误, 且误检的边界框置信度都较高, 对于图 5(a) 中发生复检的两个边界框, 线芯的置信度达 93% 和 97%。RetinaNet 除了在图 5(a) 中存在一处误检外, 在另外三幅图像中都准确检测并定位出了线芯。改进后的 EfficientDet 在图 5 四幅图像中都取得了较好的检测效果, 无错检、漏检和复检的情况, 且边框的置信度都接近于 1。实验结果验证了改进后的 EfficientDet 网络能够更好地学

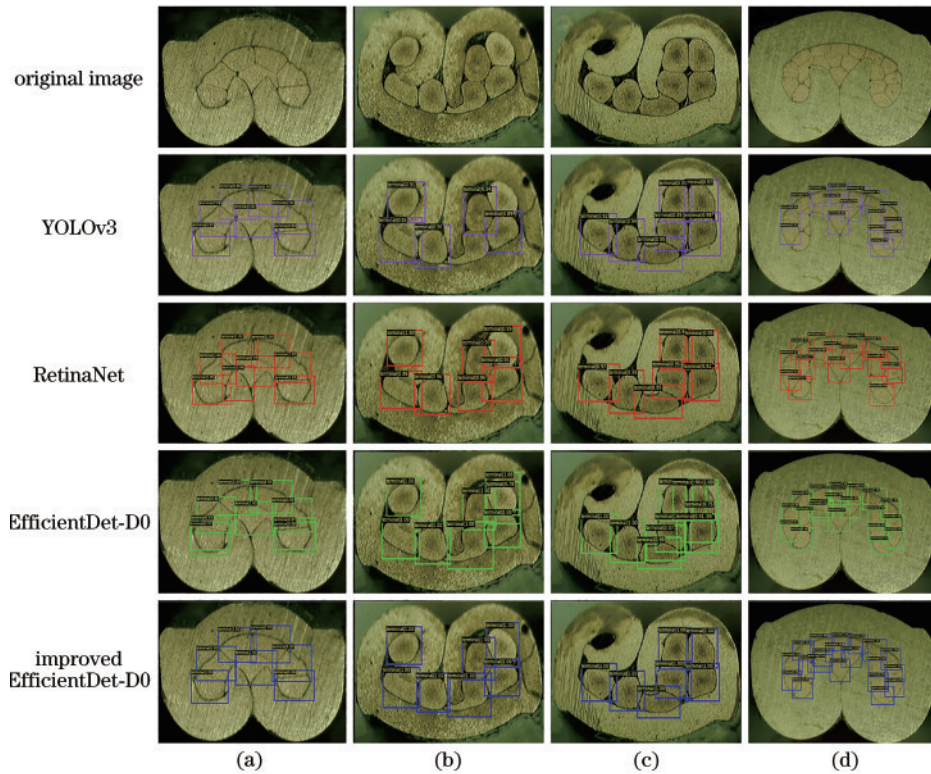


图 5 不同模型检测效果对比

Fig. 5 Comparison of detection results of different models

习显微成像下线束端子中线芯的特征,适用于显微成像下对线束端子中线芯的精确检测。

### 3.4 定量分析

为定量评估所做工作在 EfficientDet 模型上的性能提升效果,分析了改进算法、仅优化线芯锚框长宽比的 EfficientDet 模型、仅引入梯度均衡机制的 EfficientDet 模型的性能。图 6 为在线束端子显微图像

测试集上 IOU 阈值为 0.5 时的 P-R 曲线。图 6(a)为仅使用 K-means 聚类优化锚框长宽比后的模型的测试结果。图 6(b)中引入梯度均衡机制后, EfficientDet 模型的 AP 值比仅优化线芯锚框长宽比的 EfficientDet 模型高近 2.8 个百分点。所提改进算法的 AP 值达 96.23%,在线束端子显微图像测试集上的性能达到了最优。

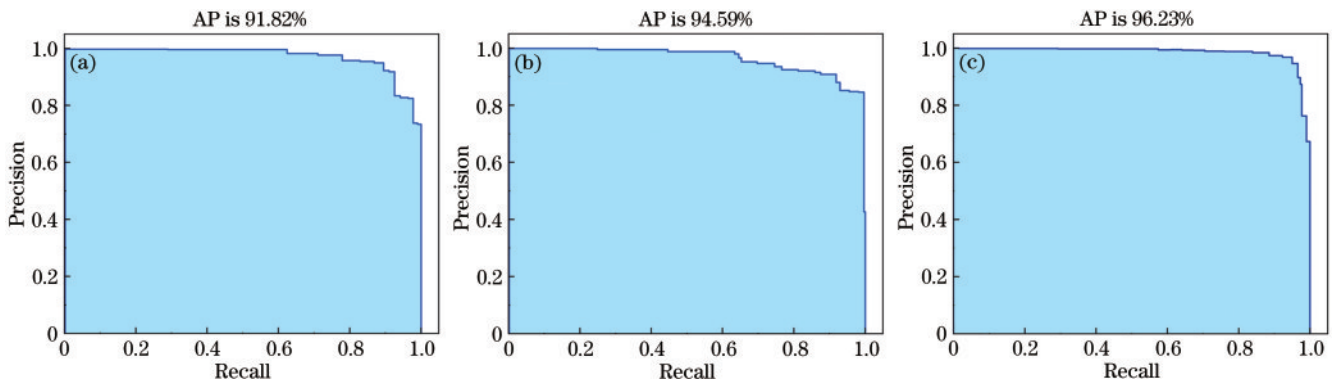


图 6 P-R 曲线及曲线下面积。(a)多维度聚类;(b)梯度均衡机制;(c)多维度聚类和梯度均衡机制

Fig. 6 P-R curves and area under the curve. (a) K-means; (b) GHM; (c) K-means and GHM

表 2 为 IOU 阈值为 0.5 时,不同目标检测模型在线束端子显微成像测试集上的检测性能对比。表 2 第 2 列为相应模型所使用的主干网络及特征融合网络,第 3 列为模型使用的损失函数,第 5 列为模型在处理 96 张测试集时测试 10 次所消耗的平均时长。

YOLOv3、RetinaNet、NAS-FPN 及 EfficientDet 同

遵循单阶段目标检测器范式。YOLOv3 在预测边界框的中心坐标和宽高的偏移量时使用均方差损失 (MSE Loss) 来拟合真实框,预测边框分类及评估置信度时使用交叉熵损失。RetinaNet、NAS-FPN 和 EfficientDet 同使用 Focal Loss 作为损失函数,但使用不同的特征融合网络。NAS-FPN 在 RetinaNet 的基

表 2 单阶段目标检测模型在线束端子显微图像数据集上的性能对比

Table 2 Performance comparison of single-stage target detection model on beam terminal micro image dataset

Detection framework	Backbone	Loss type	mAP / %	Elapsed / s	Speed / (frame · s <sup>-1</sup> )
YOLOv3	DarkNet-53	CrossEntropy+MSE	90.2	1.73	55.61
RetinaNet	ResNet-50-FPN	FocalLoss+L1Loss	89.8	2.75	34.90
NAS-FPN	ResNet-50-NAS-FPN	FocalLoss+L1Loss	89.5	4.14	23.20
EfficientDet-D0	EfficientNet-B0-BiFPN	FocalLoss+L1Loss	90.7	1.57	61.14
Improved EfficientDet-D0	EfficientNet-B0-BiFPN	GHM-C+GHM-R	96.2	1.64	58.23

础上利用神经架构搜索生成的特征网络结构进行特征融合,因而结构更加复杂,导致其对图像的推理速度变慢,仅为 23.2 frame/s,处理时间相比于 RetinaNet 延长了 1.39 s。所提改进算法在训练过程中优化了锚框的初始长宽比,使模型更容易拟合线束端子中线芯的边框特征分布,并使用 GHM-C 和 GHM-R 损失函数,解决正负样本、难易样本类别不均衡的问题。梯度均衡机制可以增强模型的鲁棒性,提高模型对线芯的检测能力。实验结果表明改进后的模型具有更高的准确率。相比于 EfficientDet,所提改进 EfficientDet 的推理速度虽然有所下降,但 mAP 提高了 5.5 个百分点,明显高于其他目标检测算法,能够保持较高检测精度的同时兼顾模型推理的实时性。

选取不同线芯数目的端子图像,用于比较实际工作环境中所提算法与人工计数方式的差异。图 7 给出了所提算法在 37 芯端子和 84 芯端子显微图像中放大后的线芯检测结果。受线束端子截面制作工艺和光照等不良因素影响,个别图像中线芯边界轮廓模糊,在视觉上难以识别,所提算法未将其检出。在绝大部分图像中,所提算法都能保持与人工同样的计数结果,计数对比结果如图 8 所示。对比实验结果表明,所提算法适用于线芯智能计数及后续的端子截面分析工作。

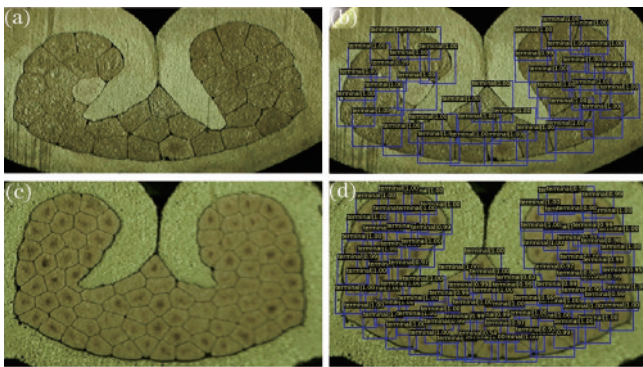


图 7 线束端子显微图像线芯检测结果。(a) 37 芯端子; (b) 37 芯端子检测结果; (c) 84 芯端子; (d) 84 芯端子检测结果

Fig. 7 Detection results of wire core of micrograph of wire harness terminals. (a) Harness terminal with 37 wire cores; (b) detection result of 37 cores terminal; (c) harness terminal with 84 wire cores; (d) detection result of 84 cores terminal

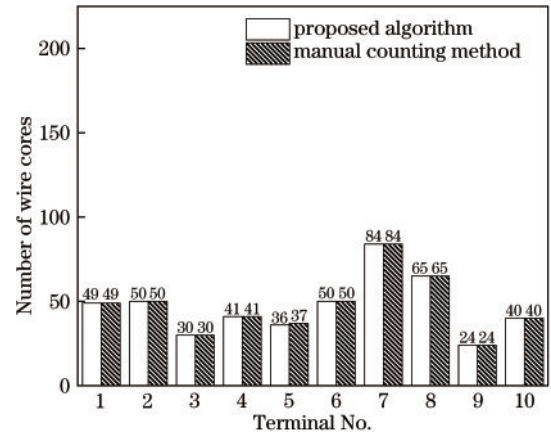


图 8 所提算法与人工计数方式的对比

Fig. 8 Comparison of counting results between the proposed algorithm and manual method

## 4 结 论

人工计数方式在对密集型的线芯计数时易受自身因素及外界因素的影响,导致线芯漏检或误检。因此,基于深度学习提出一种面向线束端子显微成像的线芯检测算法,改进 EfficientDet 算法并将其应用于线芯自动计数。使用 K-means 聚类算法优化先验框,提升先验框与真实边界框的匹配程度,降低模型对先验框的敏感性。针对单阶段目标检测模型中前景背景类别严重失衡的问题,定义梯度范数来区分样本分类的难易程度,以梯度密度来度量不同梯度范数内的样本数量。通过重构损失函数,抑制样本中属性密度较大的类别所占损失的贡献度。实验结果表明了所提改进算法的可行性与高效性,为线束端子截面智能分析的相关研究提供了参考。

## 参 考 文 献

- [1] 黄健峰. 线束端子压接工艺的研究与可靠性分析[D]. 北京: 北京邮电大学, 2016: 20-22.  
Huang J F. The study of crimping terminal process and reliability analysis[D]. Beijing: Beijing University of Posts and Telecommunications, 2016: 20-22.
- [2] 李秀娟, 李贝贝, 于心俊. 线束端子在线检测系统[J]. 自动化仪表, 2015, 36(6): 65-68.  
Li X J, Li B B, Yu X J. Online detection system of the wiring harness terminal[J]. Process Automation

- Instrumentation, 2015, 36(6): 65-68.
- [3] 芦碧波, 刘利群, 郑艳梅, 等. 一种线束端子显微图像全自动分割方法[J]. 光电工程, 2016, 43(10): 49-55.  
Lu B B, Liu L Q, Zheng Y M, et al. A method for segmenting the microscopic cable harness image automatically[J]. Opto-Electronic Engineering, 2016, 43(10): 49-55.
- [4] 袁彬淦, 钟铭恩, 倪晶鑫. 线束端子压接后外观缺陷的视觉检测算法研究[J/OL]. 系统仿真学报: 1-8 [2021-05-28]. <http://kns.cnki.net/kcms/detail/11.3092.v.20210224.0854.002.html>.  
Yuan B G, Zhong M E, Ni J X, Research on visual inspection algorithm of crimping appearance defects for wiring harness terminals[J/OL]. Journal of System Simulation: 1-8[2021-05-28]. <http://kns.cnki.net/kcms/detail/11.3092.v.20210224.0854.002.html>.
- [5] 侯守明, 王阳, 唐琪博, 等. 一种基于偏微分方程和 Canny 算子的图像分割方法[J]. 山西大学学报(自然科学版), 2017, 40(4): 676-682.  
Hou S M, Wang Y, Tang Q B, et al. Crimp terminals sectional contour extraction based on partial differential equation and canny operator[J]. Journal of Shanxi University (Natural Science Edition), 2017, 40(4): 676-682.
- [6] 芦碧波, 扈天卿, 刘甜甜. 面向线束端子显微图像分割的变指数色度滤波[J]. 光学精密工程, 2019, 27(8): 1894-1900.  
Lu B B, Hu T Q, Liu T T. Variable exponential chromaticity filtering for microscopic image segmentation of wire harness terminals[J]. Optics and Precision Engineering, 2019, 27(8): 1894-1900.
- [7] 郭芳, 王媛媛. 基于人机交互和目标检测的端子图像测量系统设计与研究[J]. 齐齐哈尔大学学报(自然科学版), 2021, 37(1): 47-50, 61.  
Guo F, Wang Y Y. Application of terminal image measurement system based on human computer interaction and target detection[J]. Journal of Qiqihar University (Natural Science Edition), 2021, 37(1): 47-50, 61.
- [8] 蒋妮, 周海洋, 余飞鸿. 基于计算机视觉的目标计数方法综述[J]. 激光与光电子学进展, 2021, 58(14): 1400002.  
Jiang N, Zhou H Y, Yu F H. Review of computer vision based object counting methods[J]. Laser & Optoelectronics Progress, 2021, 58(14): 1400002.
- [9] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, OH, USA. New York: IEEE Press, 2014: 580-587.
- [10] Girshick R. Fast R-CNN[C]//2015 IEEE International Conference on Computer Vision, December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 1440-1448.
- [11] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [12] He K M, Gkioxari G, Dollár P, et al. Mask R-CNN [C]//2017 IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2980-2988.
- [13] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector[M]//Leibe B, Matas J, Sebe N, et al. Computer Vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9905: 21-37.
- [14] 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, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 779-788.
- [15] Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6517-6525.
- [16] Redmon J, Farhadi A. YOLOv3: an incremental improvement[EB/OL]. (2018-04-08)[2020-11-20]. <https://arxiv.org/abs/1804.02767>.
- [17] 赵建敏, 李雪冬, 李宝山. 基于无人机图像的羊群密集计数算法研究[J]. 激光与光电子学进展, 2021, 58(22): 2210013.  
Zhao J M, Li X D, Li B S. Algorithm of sheep dense counting based on unmanned aerial vehicle images[J]. Laser & Optoelectronics Progress, 2021, 58(22): 2210013.
- [18] 柯宝生, 李颖, 任振波, 等. 基于深度学习的活体细胞有丝分裂检测方法[J]. 光学学报, 2021, 41(15): 1511001.  
Ke B S, Li Y, Ren Z B, et al. Deep learning-based detection method for mitosis in living cells[J]. Acta Optica Sinica, 2021, 41(15): 1511001.
- [19] Cao L B, Zhang X X, Pu J Y, et al. The field wheat count based on the efficientdet algorithm[C]//2020 IEEE 3rd International Conference on Information Systems and Computer Aided Education (ICISCAE), September 27-29, 2020, Dalian, China. New York: IEEE Press, 2020.
- [20] 明洪宇, 陈春梅, 刘桂华, 等. 基于 RetinaNet 的密集型钢筋计数改进算法[J]. 传感器与微系统, 2020, 39(12): 115-118.  
Ming H Y, Chen C M, Liu G H, et al. Improved counting algorithm for dense rebars based on RetinaNet [J]. Transducer and Microsystem Technologies, 2020, 39(12): 115-118.
- [21] Tan M X, Pang R M, Le Q V. EfficientDet: scalable and efficient object detection[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 10778-10787.
- [22] Liu S, Qi L, Qin H F, et al. Path aggregation network for instance segmentation[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE



- Press, 2018: 8759-8768.
- [23] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]//2017 IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2999-3007.
- [24] Li B Y, Liu Y, Wang X G. Gradient harmonized single-stage detector[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33: 8577-8584.
- [25] Ghiasi G, Lin T Y, Le Q V. NAS-FPN: learning scalable feature pyramid architecture for object detection [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 7029-7038.