

基于 YOLOv4 的输电线路外破隐患识别算法^①



田二胜, 李春蕾, 朱国栋, 粟忠来, 张小明, 徐晓光

(许继集团有限公司, 许昌 461000)

通讯作者: 李春蕾, E-mail: xj_lichunlei@163.com

摘要: 针对人工巡检及传统视频监控方式不能及时识别输电线路外破隐患的问题, 本文提出基于 YOLOv4 的输电线路外破隐患识别算法. 该算法采用改进 K-means 算法对图片样本集目标的大小进行聚类分析, 筛选出符合检测目标特征的锚框, 之后利用 CSPDarknet-53 残差网络提取图片深层次网络特征数据, 并采用 SPP 算法对特征图进行处理增加感受野, 提取更高层次的语义特征. 最后结合实际输电线路现场监控图片, 测试结果表明该算法能够及时准确检测到外破隐患.

关键词: YOLOv4; 残差网络; K-means; SPP 算法

引用格式: 田二胜, 李春蕾, 朱国栋, 粟忠来, 张小明, 徐晓光. 基于 YOLOv4 的输电线路外破隐患识别算法. 计算机系统应用, 2021, 30(7): 190-196. <http://www.c-s-a.org.cn/1003-3254/8082.html>

Identification Algorithm of Transmission Line External Hidden Danger Based on YOLOv4

TIAN Er-Sheng, LI Chun-Lei, ZHU Guo-Dong, SU Zhong-Lai, ZHANG Xiao-Ming, XU Xiao-Guang

(XJ Group Co. Ltd., Xuchang 461000, China)

Abstract: In this paper, we propose an algorithm based on YOLOv4 to solve the problem that manual inspection and traditional video monitoring methods cannot identify the external hidden dangers of transmission lines in time. In this algorithm, cluster analysis is performed with the improved K-means algorithm on the size of the targets in the image sample set to select the anchor frames that conform to the characteristics of detection targets. After that, the CSPDarknet-53 residual network is used to extract the deep-seated network feature data of the images, and the feature map is processed by the SPP algorithm to increase the receptive field and extract higher-level semantic features. Finally, in combination with the monitoring pictures of transmission lines, the test results show that the proposed algorithm can detect external hidden dangers timely and accurately.

Key words: YOLOv4; Residual Network (ResNet); K-means; SPP algorithm

随着电网建设覆盖范围的不断增加, 输电线路总里程也逐渐增多, 具有规模大、距离长、分布地域广等特点, 通常分布在偏远地带, 运行环境比较恶劣. 山火、滑坡等自然灾害损害架空输电线路的正常运行, 同时还存在施工机械、超高车辆、船舶导致外力破坏的危险. 传统采用人工巡线的方式存在劳动强度大、工作效率低的问题, 对突发外破事件应对时效性差.

随着智能监测技术的发展, 以视频、图像为主的

输电线路巡检的方式受到越来越多的关注, 然而传统的输电线路在线监测系统通常是将视频、图像的数据传回监控平台, 通过人工筛选的方式识别施工机械等外破隐患发生, 存在工作效率低、无法实时识别并预警的问题^[1]. 随着卷积神经网络算法在图像识别中的应用越来越广泛^[2-6], 为基于深度学习的图像识别技术在输电线路巡检中的研究和应用提供新的解决方法. 张骥等^[7]提出 Faster RCNN 算法提取特征, 用于输电线

① 收稿时间: 2020-11-05; 修改时间: 2020-12-12, 2020-12-29, 2021-01-18; 采用时间: 2021-01-26; csa 在线出版时间: 2021-06-30

路目标的图像识别,代替人工筛选,提高隐患预警实时性.许志君^[8]采用单步识别的SSD算法,与Faster RCNN相比,识别速度更快,满足边缘计算识别要求.赖秋频等^[9]提出YOLOv2,引入锚框机制,提高模型收敛速度和识别准确率.徐守坤等^[10]提出YOLOv3算法,采用Darknet53残差网络提取深层次网络的图像数据特征,适用于输电线路复杂背景下的目标识别,与YOLOv2采用Darknet19网络结构相比,目标识别效果更好.然而,以上大多数的图像识别算法模型计算时间长,不能满足输电线路图像识别模型边缘部署对算法识别实时性的要求.

针对以上问题,本文提出基于YOLOv4的输电线路外破隐患识别算法.该算法采用改进K-means算法对图片样本集目标的大小进行聚类分析,筛选出符合检测目标特征的锚框,之后利用CSP与Darknet-53残差网络相结合,能够提取图片深层次网络特征数据,同时减小计算量,并采用SPP算法对特征图进行处理增加感受野,提取更高层次的语义特征.最后将实际的输电线路现场监控图片作为测试样本集,验证该算法的性能.

1 YOLOv4 算法

YOLOv4 图像识别算法^[11]与YOLOv3算法相比,改进了网络结构和模型训练技巧等多个方面.YOLOv4利用CSP与Darknet-53残差网络相结合,能够提取图片深层次网络特征数据,同时减小计算量;增加SPP网络模块,增大感受野,提取更高层次的语义特征,能够适应不同尺寸和背景特征的目标识别.此外,采用Mosaic样本增强方法,将随机翻转、旋转等预处理后的4张图片拼接起来用于模型训练,减少模型对样本背景的依赖.

YOLOv4网络结构主要包括主干网络、Neck模块,网络结构如图1所示.其中Darknet-53残差网络结构包括一系列的 1×1 和 3×3 卷积以及不同大小的残差模块组成^[12,13].主干网络采用CSPDarknet53结构形式,将CSPNet网络结构应用于Darknet-53中残差模块,使上层的特征图一个分支进行残差运算,同时另外一个分支和经过残差运算后的特征图相结合.此外,CSPDarknet53在卷积层之后采用Mish激励函数,代替YOLOv3中的Leaky ReLU函数,提高非线性映射能力.Neck模块包括SPP网络结构和PANet网络结构.其中,SPP网络

结构采用 1×1 、 5×5 、 9×9 、 13×13 最大池化层分别对上层特征图进行处理;PANet采用高低层特征融合方法实现3个不同大小的特征图Y1、Y2、Y3输出,用于不同大小的目标识别. C 为识别目标类别数.

将一张图片划分为 $S\times S$ 网格,假如被识别的目标中心位置落在某个网格内,则该网格用于该目标的识别^[14].

每个目标识别结果需要对5个参数进行预测:目标中心位置偏移量 (x,y) ,YOLOv4以相对网格左上角的偏移代替中心位置的坐标.目标边框尺寸相对整张图片的宽高尺寸为 (w,h) ,以及该目标识别结果的置信度 $P_{\text{confidence}}$.

$$P_{\text{confidence}} = P_{\text{object}} \times IOU_{\text{prediction}}^{\text{truth}} \quad (1)$$

式中, P_{object} 为预测目标的网格存在目标的概率,存在则概率为1,不存在为0. $IOU_{\text{prediction}}^{\text{truth}}$ 为预测结果的边框与原图片中标注的目标边框位置重合度^[15,16].

$$IOU_{\text{prediction}}^{\text{truth}} = \frac{box_{\text{prediction}} \cap box_{\text{truth}}}{box_{\text{prediction}} \cup box_{\text{truth}}} \quad (2)$$

式中, $box_{\text{prediction}}$ 为预测结果的边框, box_{truth} 为原图片中的目标边框.

$$P_{C_{\text{prediction}}} = P(\text{Class}_i | \text{object}) \times P_{\text{confidence}} \quad (3)$$

式中, $P_{C_{\text{prediction}}}$ 表示该网格预测的属于第 i 类目标及边框位置置信度的概率. $P(\text{Class}_i | \text{object})$ 为当存在目标的情况下该目标属于第 i 类目标的概率.

因此,YOLOv4算法输出3个特征图Y1、Y2、Y3的大小表示将图片分别划分为 19×19 、 38×38 、 76×76 个网格,每个特征图对应3个锚框,每个网格预测4个位置 (x,y,w,h) 和 $P_{\text{confidence}}$,输出尺寸为 $S\times S\times [3\times (4+1+C)]$.

2 基于YOLOv4的输电线路外破隐患目标识别

2.1 改进K-means筛选锚框

采用改进K-means算法对输电线路外破隐患图片样本集目标的坐标进行聚类分析,筛选出符合检测目标特征的锚框,用于预测目标边框位置.K-means算法采用基于欧氏距离的聚类分析^[17],这对于不同大小的目标筛选是不合适的.本文采用改进K-means算法,将预测边框和实际目标边框重合度 $IOU_{\text{prediction}}^{\text{truth}}$ 作为距离的度量^[18,19].聚类距离公式如下:

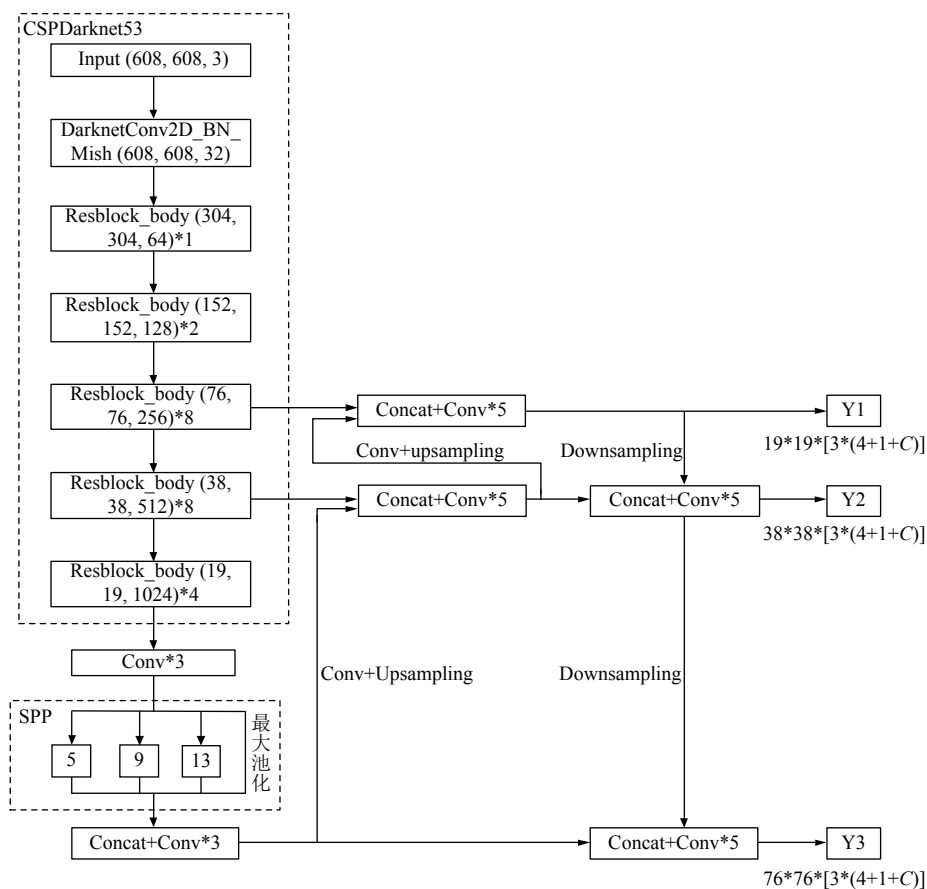


图1 YOLOv4 网络结构

$$d(box_{truth}, center) = 1 - IOU_{prediction}^{truth}(box_{truth}, center) \quad (4)$$

式中, $d(box_{truth}, center)$ 为实际目标边框与聚类中心的距离, $center$ 为聚类中心.

采用改进 K-means 算法对输电线路外破隐患图片样本集进行聚类分析, 得到用于预测目标的 9 个锚框, 3 个特征图 Y1, Y2, Y3 分别对应于 3 个锚框. 将锚框按照尺寸从大到小进行排序, 将尺寸大的特征图对应于小的锚框, 能够得到更丰富目标图片的边缘信息, 因此排名前 3 个锚框对应于大小为 19×19 的 Y1, 较大的 3 个锚框对应于大小为 38×38 的 Y2, 排名后 3 个锚框对应于大小为 76×76 的 Y3.

2.2 基于 YOLOv4 的输电线路外破隐患目标识别过程

YOLOv4 对输入的外破隐患目标图片样本集进行预处理: 包括统一调整分辨率大小, 对样本图片进行色彩调整、翻转、旋转, 用于减少样本背景对目标识别的影响. 之后采用 YOLOv4 算法进行模型训练. 其中, 损失函数计算公式如下:

$$\begin{aligned}
 f_{LOSS} = & \lambda_{coord} \sum_{i=1}^{S^2} \sum_{j=1}^b I_{ij}^{obj} \left[(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 \right] \\
 & + \lambda_{coord} \sum_{i=1}^{S^2} \sum_{j=1}^b I_{ij}^{obj} \left[\left(\sqrt{w_i} - \sqrt{\hat{w}_i} \right)^2 + \left(\sqrt{h_i} - \sqrt{\hat{h}_i} \right)^2 \right] \\
 & + \sum_{i=1}^{S^2} \sum_{j=1}^b I_{ij}^{obj} (c_i - \hat{c}_i)^2 + \lambda_{noobj} \sum_{i=1}^{S^2} \sum_{j=1}^b I_{ij}^{noobj} (c_i - \hat{c}_i)^2 \\
 & + \sum_{i=1}^{S^2} I_i^{obj} \sum_{c \in classes} \left(p_i(c) - \hat{p}_i(c) \right)^2 \quad (5)
 \end{aligned}$$

式中, f_{LOSS} 表示损失函数, 由 5 部分组成: 第 1、2 部分表示对目标坐标位置的预测, 第 3 部分表示对目标边框置信度的预测, 第 4 部分表示对该网格内未包含目标置信度的预测, 第 5 部分表示对目标所属类别的预测. λ_{coord} 表示目标坐标位置的损失权重, λ_{noobj} 表示未包含目标边框置信度的损失权重. I_i^{obj} 为网格 i 包含目标, I_{ij}^{obj} 表示网格 i 中第 j 个锚框包含目标, I_{ij}^{noobj} 表示网格 i 中第 j 个锚框不包含目标. (x_i, y_i, w_i, h_i) 表示网格 i 中预

测的目标边框, $(\hat{x}_i, \hat{y}_i, \hat{w}_i, \hat{h}_i)$ 表示网格 i 中实际的目标边框. c_i 为网格 i 预测的类别, \hat{c}_i 表示网格 i 目标所属实际的类别. $p_i(c)$ 为网格 i 内预测的目标概率, $\hat{p}_i(c)$ 表示网格 i 内实际的目标概率. S^2 表示一张图片被划分的网格数. b 为每个网格需要预测的边框数^[20].

基于 YOLOv4 的输电线路外破隐患目标检测流程图如图 2 所示, 包括 3 个阶段: 对图片样本集进行预处理、模型训练及待检测目标图片的测试.

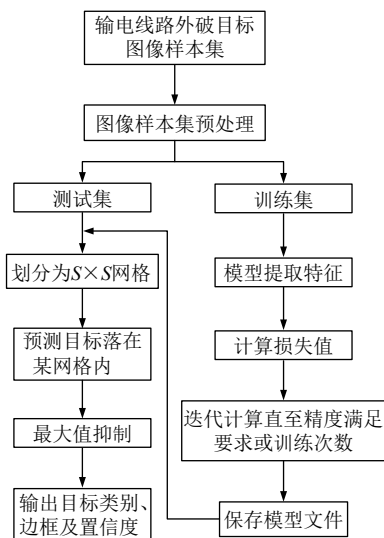


图 2 输电线路外破隐患目标检测流程图

检测步骤如下:

(1) 对样本集进行预处理

1) 收集输电线路外破隐患图片样本集, 对图片样本名称进行统一编号, 之后将图片分辨率统一调整为 608×608 ;

2) 采用样本增强的方法对样本图片进行色彩调整、翻转、旋转等处理;

3) 运用标注工具对样本集进行目标类别和边框的标注;

4) 按照 VOC2007 数据集的格式整理样本集及标注文件, 并生成训练样本集和测试样本集.

(2) 模型训练

1) 根据本项目的样本集数据, 修改配置文件中的学习率、批次数等相关参数;

2) 运用 CSPDarknet53 网络提取输入样机图片的外破隐患目标特征;

3) 运用 Neck 模块将特征向量进行高低层次特征融合, 生成 3 个不同尺寸的特征图, 用于目标多尺度

预测;

4) 将预测结果和标注文件中实际目标边框、类别代入损失函数计算公式, 得到损失值;

5) 利用神经网络反向传播方法自动调整网络结构参数, 重复步骤 2)~4), 直到模型满足精度要求或达到训练次数, 生成模型文件.

(3) 模型测试

1) 输入外破隐患图片测试样本集;

2) 利用网络前向传播得到目标的边框和类别;

3) 根据目标预测边框、实际边框计算 IOU 重合度;

4) 根据步骤 3) 得到的 IOU 值和类别概率得到置信度;

5) 设置 IOU 阈值筛选到 IOU 值较小的预测框, 对剩余的预测框采用最大值抑制算法得到最终的预测结果.

3 实验分析

本文采用 Ubuntu18.04 操作系统, GPU 硬件选择 NVIDIA GTX 1080Ti, 软件为 CUDA9.0, CUDNN7.0. 搭建基于 Keras 框架的模型训练及测试环境, 采用 YOLOv4 算法模型对输电线路外破隐患目标进行识别, 输出被检测图片中外破隐患目标的类别和边框.

3.1 数据集制作

根据实际的输电线路现场监控视频及图像信息, 制作模型训练及识别的数据集, 包含施工机械、吊车、塔吊 3 种外破隐患目标类型, 共计 3825 张. 其中施工机械类 1362 张, 吊车类 1184 张, 塔吊类 1279 张. 该数据集涵盖不同背景内容、光照强度、天气情况下的目标图片样本, 能够减小图像背景对目标识别的影响, 提高模型精度. 将数据集按照 7:3 的比例分为训练集和测试集, 划分结果如表 1 所示.

表 1 数据集划分结果

类别	训练集	测试集	总数
施工机械	954	408	1362
吊车	829	355	1184
塔吊	896	383	1279
总数	2679	1146	3825

3.2 输电线路外破隐患目标识别

将外破隐患目标图像数据集按照 VOC2007 数据格式进行整理, 并将数据集中目标标注的坐标位置进行归一化处理, 作为模型训练的样本输入, 归一化公式

如下:

$$\begin{cases} x = \frac{x_{\max} + x_{\min}}{2width}, y = \frac{y_{\max} + y_{\min}}{2hight} \\ w = \frac{x_{\max} - x_{\min}}{width}, h = \frac{y_{\max} - y_{\min}}{hight} \end{cases} \quad (6)$$

式中, (x, y) 为归一化处理后的中心点坐标, (w, h) 表示归一化处理后的图片宽度和高度. (x_{\min}, y_{\min}) 为原始图片中目标的左上角坐标, (x_{\max}, y_{\max}) 表示原始图片中目标的右下角坐标.

将训练集图片中外破隐患目标作为聚类算法输入, 采用改进 K-means 算法筛选出适用于本项目的锚框, 共计 9 个锚框, 分别是 (39, 27), (72, 51), (85, 213), (112, 83), (147, 289), (179, 161), (221, 321), (323, 379), (539, 431), 该组锚框与图片中施工机械等外破隐患目标大小快速适配, 提高模型训练速度.

3.3 模型的训练与测试

首先, 修改预训练模型的配置文件相关参数, 以适用于本项目特点, 部分参数值如表 2 所示. 此外, 将本项目的 9 个锚框替换配置文件中相应位置的锚框. 然后, 在 Keras 框架下加载预训练模型, 输入图片样本集进行模型训练.

表 2 配置文件部分参数值

参数名称	设置值
批次大小	32
组数	8
衰减权重	0.0003
学习率	0.001

保存训练后生成的模型文件, 结合数据集制作时得到的测试样本集, 进行模型性能的测试.

挑选不同项目目标图像特点的测试样本, 用于检测模型在复杂背景、不同姿态、不同光照、不同天气、多个目标情况下的性能. 图 3(a)、图 3(b) 为同类目标在不同背景内容下的识别结果, 实验结果表明该模型能够识别不同背景下的目标, 减少对图片背景内容的依赖. 图 4(a)、图 4(b) 为同一目标不同姿态下的识别结果, 实验结果表明该模型能够识别不同姿态的目标, 增强对目标姿态的识别. 图 5(a)、图 5(b) 为同一目标不同光照强度下的识别结果, 实验结果表明该模型能够识别不同光照强度下拍摄的目标图片, 提升对图片色彩、亮度变化的适应能力. 图 6(a)、图 6(b) 为同一目标不同天气情况下的识别结果, 实验结果表明该模型在有雾天气下对目标的准确识别. 图 7(a)、

图 7(b) 为一张图片中存在多个目标的识别结果, 实验结果表明该模型能够实现对多个目标的完整识别, 提高识别准确率.



(a) 农田背景下施工机械识别



(b) 高速公路背景下施工机械识别

图 3 不同背景内容下目标识别



(a) 侧面姿态下施工机械识别



(b) 正面姿态下施工机械识别

图 4 不同姿态下目标识别



(a) 光照强度较弱下吊车识别



(b) 光照强度较强下吊车识别

图5 不同光照强度下目标识别



(a) 天气晴朗下施工机械识别



(b) 有雾天气下施工机械识别

图6 不同天气情况下目标识别

将本文算法与常用的图像识别算法进行对比, 对比结果如表3所示. 从表3可以看出, 与常用的 Faster RCNN、SSD 算法相比, 本文算法的各个类别 AP 及 mAP

值均高于其它算法, 平均单张图片识别速度更快, 表明本文算法的优越性.



(a) 多个施工机械目标识别



(b) 多个塔吊目标识别

图7 同一张图片多个目标识别

表3 不同算法对比结果

不同算法	AP(%)			mAP(%)	识别速度
	施工机械	吊车	塔吊		
Faster RCNN	91.02	89.72	90.43	90.39	0.19
SSD	89.92	87.03	88.45	88.46	0.08
本文算法	92.61	91.53	92.18	92.11	0.07

4 结论与展望

文中提出基于 YOLOv4 的输电线路外破隐患识别算法, 采用 CSPDarknet-53 残差网络提取图片深层次网络特征数据, 并结合 SPP 算法对特征图进行处理增加感受野, 提取更高层次的语义特征. 同时, 采用改进 K-means 算法对图片样本集目标的大小进行聚类分析, 筛选出针对检测目标特征分析的锚框, 用于模型训练及检测. 最后采用本文算法和其它两种常用的图像识别算法对实际的输电线路现场图片样本集进行测试, 实验结果表明本文算法具有更高的识别准确率, 且平均单张图片识别速度更快, 满足模型前端部署的要求, 验证该算法的优越性.

参考文献

1 赵建青, 姚瑶, 邱玩辉, 等. 基于输电线路在线巡视系统的

- 智能预警系统研究. 电力系统保护与控制, 2013, 41(23): 49–54. [doi: [10.7667/j.issn.1674-3415.2013.23.008](https://doi.org/10.7667/j.issn.1674-3415.2013.23.008)]
- 2 Girshick R. Faster R-CNN. Proceedings of 2015 IEEE International Conference on Computer Vision. Santiago, Chile. 2015. 1440–1448.
- 3 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](https://doi.org/10.1109/TPAMI.2016.2577031)]
- 4 Liu W, Anguelov D, Erhan D, *et al.* SSD: Single shot multibox detector. Proceedings of the 14th European Conference on Computer Vision. Amsterdam, The Netherlands. 2016. 21–37.
- 5 Redmon J, Divvala S, Girshick R, *et al.* You only look once: Unified, real-time object detection. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA. 2016. 779–788.
- 6 Redmon J, Farhadi A. YOLO9000: Better, faster, stronger. Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA. 2017. 6517–6525.
- 7 张骥, 余娟, 汪金礼, 等. 基于深度学习的输电线路外破图像识别技术. 计算机系统应用, 2018, 27(8): 176–179.
- 8 许志君. 基于深度卷积神经网络的高压输电线路图像识别 [硕士学位论文]. 广州: 华南理工大学, 2019.
- 9 赖秋频, 杨军, 谭本东, 等. 基于 YOLOv2 网络的绝缘子自动识别与缺陷诊断模型. 中国电力, 2019, 52(7): 31–39.
- 10 徐守坤, 倪楚涵, 吉晨晨, 等. 基于 YOLOv3 的施工场景安全帽佩戴的图像描述. 计算机科学, 2020, 47(8): 233–240.
- 11 Bochkovskiy A, Wang CY, Liao HYM. YOLOv4: Optimal speed and accuracy of object detection. arXiv: 2004.109342020, 2020.
- 12 Redmon J, Farhadi A. YOLOv3: An incremental improvement. arXiv: 1804.02767, 2018.
- 13 王思元, 王俊杰. 基于改进 YOLOv3 算法的高密度人群目标实时检测方法研究. 安全与环境工程, 2019, 26(5): 194–200.
- 14 Huang G, Liu Z, Van Der Maaten L, *et al.* Densely connected convolutional networks. Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA. 2017. 2261–2269.
- 15 Zhang GX, Zeng Z, Zhang SW, *et al.* SIFT matching with CNN evidences for particular object retrieval. Neurocomputing, 2017, 238: 399–409. [doi: [10.1016/j.neucom.2017.01.081](https://doi.org/10.1016/j.neucom.2017.01.081)]
- 16 赖奎, 姚军艳, 马承志, 等. 输电线路智能巡检系统的设计研究. 广东电力, 2016, 29(7): 105–110. [doi: [10.3969/j.issn.1007-290X.2016.07.021](https://doi.org/10.3969/j.issn.1007-290X.2016.07.021)]
- 17 王琳, 卫晨, 李伟山, 等. 结合金字塔池化模块的 YOLOv2 的井下行人检测. 计算机工程与应用, 2019, 55(3): 133–139. [doi: [10.3778/j.issn.1002-8331.1710-0236](https://doi.org/10.3778/j.issn.1002-8331.1710-0236)]
- 18 徐诚极, 王晓峰, 杨亚东. Attention-YOLO: 引入注意力机制的 YOLO 检测算法. 计算机工程与应用, 2019, 55(6): 13–23. [doi: [10.3778/j.issn.1002-8331.1812-0010](https://doi.org/10.3778/j.issn.1002-8331.1812-0010)]
- 19 石磊, 王亚敏, 曹仰杰, 等. 基于深度卷积神经网络的车型识别. 计算机科学, 2018, 45(5): 280–284.
- 20 Tang PJ, Wang HL, Kwong S. G-MS2F: GoogLeNet based multi-stage feature fusion of deep CNN for scene recognition. Neurocomputing, 2017, 225: 188–197. [doi: [10.1016/j.neucom.2016.11.023](https://doi.org/10.1016/j.neucom.2016.11.023)]