

基于深度学习的高铁电线杆及杆号检测与识别方法

田 宇¹, 徐志洁^{1*}, 张健钦², 张冬梅¹, 范博文¹

¹北京建筑大学理学院, 北京

²北京建筑大学测绘与城市空间信息学院, 北京

收稿日期: 2022年9月14日; 录用日期: 2022年10月14日; 发布日期: 2022年10月24日

摘 要

面向高铁电线杆智能监测任务, 提出一种基于YOLOv3的端到端的高铁电线杆自动检测及杆号识别算法。该算法首先对电线杆及杆号区域进行检测, 并根据杆号区域检测坐标自动裁剪, 然后识别杆号区域中的数字, 最后将电线杆检测结果与数字识别结果自动结合。通过构建高铁电线杆图像数据集以及杆号区域数据集, 进行大量实验。实验结果表明, 我们提出的方法对电线杆及杆上编号的检测与识别准确率分别达到了97.50%、95.30%, 能有效地完成高铁最优电线杆及杆号的自动检测任务。

关键词

目标检测, 高铁电线杆, 杆号, YOLOv3, 两阶段检测

Detection and Recognition Method of Pole and Number in High-Speed Railway Based on Deep Learning

Yu Tian¹, Zhijie Xu^{1*}, Jianqin Zhang², Dongmei Zhang¹, Bowen Fan¹

¹School of Science, Beijing University of Civil Engineering and Architecture, Beijing

²School of Geomatics and Urban Spatial Informatics, Beijing University of Civil Engineering and Architecture, Beijing

Received: Sep. 14th, 2022; accepted: Oct. 14th, 2022; published: Oct. 24th, 2022

Abstract

This work offers an end-to-end automatic detection of poles and number identification technique

*通讯作者。

文章引用: 田宇, 徐志洁, 张健钦, 张冬梅, 范博文. 基于深度学习的高铁电线杆及杆号检测与识别方法[J]. 计算机科学与应用, 2022, 12(10): 2318-2330. DOI: 10.12677/csa.2022.1210237

in high-speed railway based on YOLOv3 for the intelligent monitoring task of poles in high-speed railway. To begin, the algorithm recognizes the pole and the number area, then crops automatically based on the detection findings of the number area's coordinates. Then identify the numbers in the number area. Finally, the results of the pole detection and number identification are integrated automatically. By building the picture data set of poles and number region in high-speed railway, a great number of experiments were carried out. The experimental results show that the proposed method's detection and recognition accuracy for the pole and the number on the pole is 97.50% and 95.30%, respectively, indicating that it can effectively complete the automatic detection tasks of the most optimal pole and the number on the pole in high-speed railways.

Keywords

Object Detection, High-Speed Railway Poles, Pole Numbers, YOLOv3, Two-Stage Detection

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在我国庞大的高铁电力系统中,大量的电线杆设备需要人员进行定期的巡检维护,但高铁线路交错复杂且地处偏远地带,因而可能长期处于无人值守状态。高铁上的电线杆是用来给电力机车供电的输电线路,属于高速铁路接触网。高速铁路接触网,是沿铁路线上空架设,向电力机车供电的输电线路。高铁列车运行利用的电流就是通过机车上端的接触网来输送的。接触网一旦停电,或列车电弓与接触网接触不良,便会对列车的供电产生影响。因此接触网的质量和在工作状态将直接影响着高铁列车的运输效率。

由于接触网是露天设置,没有备用,线路上的负荷同时随着电力机车的运行而沿接触线移动和变化。首先,如果能够借助于计算机视觉技术实现智能实时目标检测,那么高铁电力系统就能够极大地减少人力成本的投入,提升管理效率,同时这对于保证输电线路的正常工作有着非常积极的影响,并且对于电力机车的正常运转有着关键性的意义。其次,大量的数据使得数据存储存在大量的空间占用问题以及数据利用率低等问题,因此通过计算机视觉技术对最优电线杆图像进行自动检测并保存,可以大大减少非必要图像存储空间的占用,从而提升存储空间以及有效数据的利用率。

传统的目标检测算法,主要依赖于人工获取有关目标的特征信息,这需要极大的人工成本。随着深度学习的发展为计算机视觉带来了很大的突破。传统的目标检测算法已经不能满足精度与速度上的要求。基于深度学习的目标检测算法主要分为两大类:以 YOLO 系列网络[1] [2] [3]为代表的 One-stage 网络和以 Faster RCNN [4]为代表的 Two-stage 网络。Two-stage 目标检测算法:R-CNN (Regions with Convolutional Neural Network Features)系列[4] [5] [6],优点是检测精度高,但耗时长,不能满足实时检测的需求;One-stage 目标检测算法:SSD (Single Shot Multibox Detector)算法[7]、YOLO (You Only Look Once)系列[1] [2] [3],一阶段目标检测模型都是端到端模型,端到端即将一幅图像输入到模型中,模型会直接输出在图像内识别到的目标的位置及类别。因此,一阶段模型在满足检测精度的同时,极大地减少时间损耗,满足实时检测的需求,更适合应用于工业检测。

2017年谢兴阳等人[8]利用无损检测技术对电线杆的健康状态进行评估,提出了一种基于窄带频率激励和先进信号处理技术的新型电线杆无损检测技术系统。为了提升损伤检测的精度,该系统采用了频率响应函数方法以及主成分分析从单频应力波中提取信号特征,从而对电线杆的健康状态进行评估。2017

年李涛等人[9]针对监控视频中的电线杆局部遮挡区域问题,利用双边滤波、HSV 梯度改进 LSD 直线检测,提出了一种基于改进的线段分割检测算法。该方法聚焦于电线杆的局部区域检测,是为了解决监控视频中被电线杆遮挡的信息缺失问题,并非是对完整的电线杆进行检测。2019 年江南等人[10]对高空电线杆损坏检测方法进行了研究,提出了一个基于绝缘子检测高空电线杆损坏的方法。作者认为高空电线杆是否损坏主要由绝缘子的情况所决定,因此绝缘子是否存在缺陷成为高空电线杆损坏检测的关键指标。这一方法是基于传统的边缘检测算法以及椭圆拟合方法,根据拟合误差判断绝缘子是否损坏,因此并不能很好地应用于现有的大量数据检测工作。

以上的研究方法都是基于传统的信号处理方式或机器学习的方式,深度学习算法的出现,解决了传统机器视觉方法存在的问题。2016 年刘志浩等人[11]从农网改造的实际角度出发,运用 YOLO 模型对电力线路的各型电线杆进行检测,实现电线杆自动识别与定位。但所用数据集仅 737 张图像,数量过少且标注信息仅包含电线杆部分结构,并不能对完整电线杆进行检测定位。2020 年庞宁[12]提出了基于深度学习的输电线杆塔鸟巢检测与识别算法,该方法以 Faster RCNN 算法网络结构为基础,利用 ZF-NET 网络提取鸟巢图像特征图,将此特征图作为输入,生成矩形候选区域,计算检测窗口有无鸟巢。然而在数据集中,有些鸟巢的结构完整,特征明显,有些鸟巢还处于未完成阶段,仅有几根树枝组成,特征不明显,这使得模型对鸟巢特征的学习受到极大影响,导致模型的泛化能力较差。2021 年刘凯歌等人[13]基于 Faster RCNN 对铁轨电线杆进行检测,数据集包含 1895 张无重复场景图像,由于铁轨沿途电线杆图像采集均采用高速连拍相机,重复场景的存在是无可避免的,这在实际应用场景中是不实际的。

此外,还有大量的研究工作基于 YOLOv3 模型应用于交通领域的相关检测任务。2020 年 JinLiu 和 Dongquan Zhang [14]对车辆目标检测任务进行了研究,他们根据实际交通状况和效率要求,通过添加尺度预测层并结合 K-means++ 算法提高聚类效果,提出了基于 YOLOv3 的车辆目标检测算法 F-YOLOv3,并取得了较好的性能,极具推广前景。2021 年 Huazhou Dou 等人[15]对交通标志检测进行了相关研究,他们通过修改 YOLOv3 的主干特征提取网络,提出一种基于 YOLOv3 的轻量级算法,实现了实时、准确的交通标志检测。2021 年 Dongying Gao 等人[16]同样是对交通标志进行检测并将其应用于盲人出行,他们通过结合 EfficientNet-lite 网络与 YOLOv3 的特点,在检测精度适当降低的前提下,减小了模型尺寸和对设备性能的依赖,但是检测精度仍需提高。

本文为将目标检测算法应用于高铁电线杆及杆号检测,选择了运行速度快,端到端的 YOLOv3 模型 [3],引用 YOLOv3-tiny 版本对高铁电线杆及杆号进行多目标检测的研究,在检测出重复杆号的基础上筛选过滤重复电线杆,最终选出最优电线杆(清晰、完整的电线杆图像)。针对高铁电线杆图像中,电线杆作为大目标,数字杆号作为小目标的特殊情形,直接使用单阶段模型进行训练检测,会造成检测精度低,小目标无法检出的情况。因此,我们设计采用两阶段进行目标检测,使用不同的数据集对 YOLOv3-tiny 模型进行训练更新。在训练及检测过程前,利用 K-means 聚类算法[17]在自制的电线杆数据集上对先验框进行重聚类,得出适用于我们构建的电线杆数据集的先验框,训练过程在设定 Epoch 范围时以最低验证集损失为标准保存最优模型。最终将第一阶段模型应用于检测定位电线杆及杆号区域,并在检测过程中自动裁剪杆号区域。此时第二阶段模型用于识别杆号区域内的数字,最终将杆号识别结果,与第一阶段电线杆检测结果自动合并。

我们基于 YOLOv3 模型设计电线杆检测及杆号识别系统,通过反复实验设计了对高铁电线杆以及杆上杆号进行两阶段检测与识别的方法,这是一个端到端且实时高效的电线杆检测系统。本文的其余部分分别为:在第 2 部分详述了 YOLOv3 的检测原理,第 3 部分描述了两阶段电线杆检测及杆号识别过程,第 4 部分主要包含数据集构建、两阶段实验结果及分析以及消融实验,最后,第 5 部分给出了我们的研究结论、算法存在的不足以及未来的研究方向。

2. 相关方法介绍

YOLO 系列是一种使用深度卷积神经网络学得特征来检测对象的目标检测器。YOLOv3 是一个全卷积神经网络(Fully Convolutional Network, FCN) [18], 由 75 个卷积层以及残差连接块和上采样层构成, 它使用步幅为 2 的卷积层对特征图进行下采样, 卷积过程中宽高不断压缩, 通道数不断扩张。网络中无池化层, 防止了由于池化使低级特征丢失, 导致检测精度低的问题。在此过程中, 上采样与之前层的特征图拼接, 实际上为特征金字塔的过程, 利用特征金字塔可以进行多尺度的特征融合, 提取出更有效的特征, 从而有助于网络学习细粒度特征, 帮助检测较小目标。网络最后一个卷积层的卷积核的尺寸是: $1 \times 1 \times (k \times (5 + C))$, 这里 1×1 是卷积核的大小, k 表示每个网格单元可以预测的边界框数量(一般默认为 3), 每个边界框有 $5 + C$ 个属性[坐标 (t_x, t_y, t_w, t_h) 、置信度、 C 个类别的类别概率], 即预测过程是利用 $(5 + C) \times k$ 个 1×1 的卷积核进行卷积预测, 达到提高识别准确率和效率的目的。由于使用了 1×1 卷积, 因此预测特征图的大小不发生变化。

YOLOv3 的检测思想是多尺度特征检测, 包含三个尺度: $13 \times 13, 26 \times 26, 52 \times 52$ 大尺度的特征图, 有助于检测相对较小的目标。检测时, 首先忽略掉分数低于阈值的边界框, 然后根据非极大值抑制(Non-Maximum Suppression, NMS)过滤余下的边界框, 即按照得分高低对边界框进行排序, 然后计算得分最高的边界框和其他边界框的 IOU 值, 以 IOU 值衡量边界框的重复度, 过滤掉重复度较高的边界框。YOLOv3 的损失函数如公式(1)所示, 主要由三部分组成: 第一部分(lbox)为检测目标框位置所产生的坐标损失和尺寸损失, 第二部分(lobj)为检测目标框置信度所产生的损失, 其中第一项为正样本检测框的置信度损失, 第二项为负样本检测框的置信度损失, 第三部分(lcls)为检测类别概率所产生的损失, 这三部分共同组成了目标检测器 YOLOv3 的损失函数。

$$\begin{aligned}
 lbox &= \sum_{i=0}^{K \times K} \sum_{j=0}^M I_{ij}^{obj} (2 - w_i * h_i) \left[-x_i * \log(x_i^*) - (1 - x_i) * \log(1 - x_i^*) \right] \\
 &\quad + \sum_{i=0}^{K \times K} \sum_{j=0}^M I_{ij}^{obj} (2 - w_i * h_i) \left[-y_i * \log(y_i^*) - (1 - y_i) * \log(1 - y_i^*) \right] \\
 &\quad + \sum_{i=0}^{K \times K} \sum_{j=0}^M I_{ij}^{obj} (2 - w_i * h_i) \left[(w_i - w_i^*)^2 + (h_i - h_i^*)^2 \right] \\
 lobj &= \sum_{i=0}^{K \times K} \sum_{j=0}^M I_{ij}^{obj} \left[C_i * \log(C_i^*) + (1 - C_i) * \log(1 - C_i^*) \right] \\
 &\quad + \sum_{i=0}^{K \times K} \sum_{j=0}^M I_{ij}^{noobj} \left[C_i * \log(C_i^*) + (1 - C_i) * \log(1 - C_i^*) \right] \\
 lcls &= \sum_{i=0}^{K \times K} \sum_{j=0}^M I_{ij}^{obj} \sum_{c \in \text{classes}} \left[p_i(c) * \log(p_i^*(c)) - (1 - p_i(c)) * \log(1 - p_i^*(c)) \right] \\
 \text{loss}(\text{object}) &= lbox - lobj - lcls \tag{1}
 \end{aligned}$$

其中, $K \times K$ 为输入图像被模型划分后形成的所有网格单元, 在 YOLOv3 中为 $13 \times 13, 26 \times 26, 52 \times 52$, M 为每个网格包含的检测框(anchor box)个数, I_{ij}^{obj} 用于判断第 i 个网格单元中第 j 个检测框(anchor box)是否有需要检测的目标, 如果存在, 则为 1, 反之则为 0; I_{ij}^{noobj} 的含义与之相反; $x_i^*, y_i^*, w_i^*, h_i^*, C_i^*, p_i^*(c)$ 分别表示真实标注框的位置、尺寸、置信度及类别概率 $x_i, y_i, w_i, h_i, C_i, p_i(c)$ 为模型预测目标框的位置、尺寸、置信度和类别概率。需要注意的是: x_i, y_i, w_i, h_i 并非目标框的绝对位置和尺寸, 仅为相对于网格中心位置和先验框尺寸的相对值。

YOLOv3-tiny 与 YOLOv3 的最大不同在于特征提取网络上, YOLOv3-tiny 的主干特征提取网络是

7 个卷积与 6 个最大池化层的按序堆叠, 特征检测尺度也只有两个: $13 \times 13, 26 \times 26$, 网络结构如图 1 所示, 因此 YOLOv3-tiny 的检测精度会随着特征层的无法深入而降低。同时也正因为 YOLOv3-tiny 是轻量级模型, 因此它的优势在于网络结构简单, 计算量较小, 速度快, 可以在移动端或设备端运行, 常应用于工业项目检测。本文为了达到速度与耗时要求, 选用 YOLOv3-tiny 模型进行最优电线杆检测及杆上编号识别。

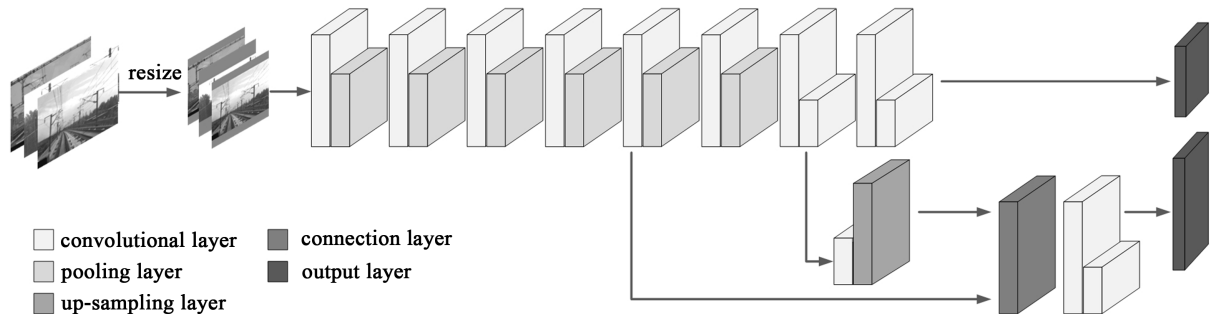


Figure 1. The network architecture of YOLOv3-tiny

图 1. YOLOv3-tiny 网络结构

3. 两阶段电线杆检测及杆号识别

为了同时识别定位最优电线杆的位置(最优即电线杆相对完整、清晰度较好)以及最优杆的杆号, 考虑到我们所使用的数据集的特殊性, 如图 3, dataset-1 所示, 电线杆与数字杆号相比, 两者大小差异较大, 这对于小目标数字的检测不具有任何优势, 因此在单阶段 YOLOv3-tiny 模型中, 想要同时检测出电线杆并识别出杆号是具有一定难度的。这一点在实验中也得到验证, 电线杆目标较大, 检测效果理想, 数字相对图像而言, 目标极小且像素低, 检测效果很差。

基于以上任务存在的难度, 我们提出一种两阶段的检测与识别算法, 首先分别使用第一阶段高铁电线杆数据集、第二阶段杆号区域内数字数据集训练更新两个 YOLOv3-tiny 模型, 得到针对电线杆、杆号区域两类目标以及十类数字目标, 两阶段目标检测与识别的权重参数。然后两阶段模型应用时, 第一阶段模型用于检测电线杆及杆号区域, 并自动裁剪检测定位出的杆号区域, 然后送入第二阶段模型用于识别杆号, 最终将检测结果自动结合, 得到可视化并带有杆号识别结果的电线杆检测结果。

然后通过逻辑编码设计了电线杆检测过滤操作以获得最优杆检测结果, 若同一杆号电线杆检测出多张结果, 则进行这一过滤操作, 即自动筛选出同一杆号电线杆检测结果中得到最大置信度的一张图像作为最优杆检测结果。若未识别出重复杆号的电线杆图像, 则不进行过滤操作, 自动保存这一检测结果。通过实验验证, 我们设计的电线杆检测及杆号识别两阶段算法, 可以完成最优杆检测任务, 同时可以最大程度地提升空间存储利用率。任务流程如图 2 所示。

3.1. 针对检测目标重聚类

YOLO 系列从 YOLOv2 开始引入“锚框(anchor box)”机制[2], 这一机制有助于避免模型在训练的时候盲目地寻找目标, 从而加快模型收敛速度。YOLOv3 使用 K-means 算法[17]在数据集中所有样本的真实框上聚类, 得到锚框。对于锚框的个数, 作者采用实验的方式, 分别用不同数量的 anchor 应用到模型, 然后在模型复杂度和高召回率之间找到最优的一组 anchor box, 最终得出 anchor box 的数量为 9 时效果最佳。YOLOv3 存在三种尺度的输出, 因此使用 9 个锚框。在 YOLOv3-tiny 上, 只有两个尺度的输出, 因此 anchor box 的个数为 6。

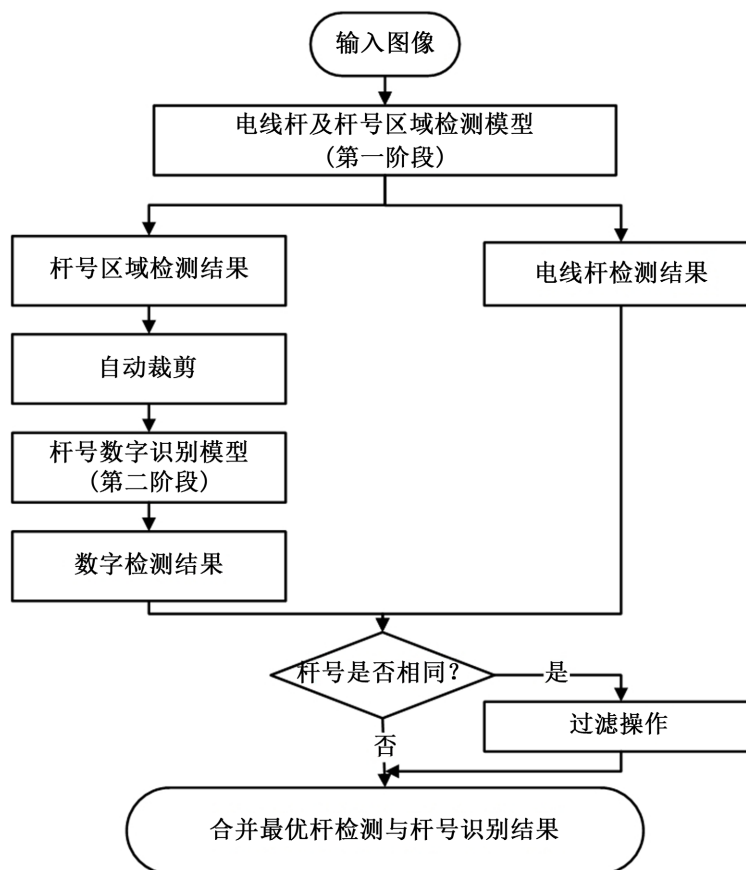


Figure 2. The task process of poles detection and numbers recognition in high-speed railway
图 2. 高铁电线杆检测及杆号识别任务流程

由于默认的 anchor 方案主要是对 VOC 数据集中的检测目标而设定的,对本次检测的目标不具有针对性,在实验中有所验证。考虑到所采用数据集的特点,即在第一阶段目标检测任务中,目标电线杆与杆号区域的大小差异较大,且目标种类数目为 2,如图 3 所示,为了提升检测精度,我们使用 K-means 聚类算法,在第一阶段真实标注边界框(*i.e.* Bounding Box of Ground Truth)的基础上重聚类,提取带有代表性的宽高,即“锚框”(anchor box)。锚框的作用实际上是对预测的目标范围进行约束,并加入了尺寸先验经验,从而实现多尺度学习的目的。

3.2. 聚类后进行线性缩放

在进一步的实验中,我们发现仅对检测数据集进行重聚类,也无法达到检测要求。经分析由于电线杆与杆号区域的大小差异较大,且目标种类少,类型单一,同类目标标注框相差不大,尺寸比较集中,如果仅通过 K-means 聚类算法重聚类获得新的锚框(anchor box),尺寸自然就很集中,几乎相差不大,这就无法体现 YOLO 模型多尺度输出的优势。同时我们发现较多标注框比重聚类得到的锚框尺寸大,这样导致在训练时对模型的要求更高。为了增加锚框的多样性,使之适应不同尺度的目标,我们使用了一种将锚框进行线性尺度的缩放的方式,即将聚类所得锚框的尺寸进行伸缩,提升模型的检测效果。这一操作机制的效果在实验中得以验证。其计算公式如(2)所示,引入线性缩放机制,即将锚框的尺寸进行拉伸。

线性缩放的公式如下:

$$\begin{cases} x'_1 = \alpha x_1, \\ x'_n = \beta x_n, \\ x'_i = \frac{x_i - x_1}{x_n - x_1} (x'_n - x'_1) + x'_1, \\ y'_i = x'_i \frac{y_i}{x_i}. \end{cases} \quad (2)$$

其中, x_i, y_i 表示通过 K-means 聚类获得的初始锚框宽高信息, x'_i, y'_i 表示通过线性缩放后锚框的宽高信息, i 的取值范围为 $[1, n]$, n 为所需锚框个数, 我们设计的算法中所需锚框个数为 6。 α, β 为缩放比例系数, 取值区间为 $[0, 1]$, 经实验, 我们将其设置为 0.5。

4. 实验与分析

4.1. 数据集构建

为了保证数据的多样性, 我们通过收集两条高铁线路上, 共 14 个地点的电线杆图像来构建数据集。训练数据集从 67346 张电线杆图像中进行筛选, 最终得到一阶段数据集为 11040 张图像, 其中 9439 张用作训练集, 497 张用作验证集, 1104 张用作测试集, 并使用 LabelImg 标注 YOLO 所需的标签格式: {类别 1, $(x_{\min}, y_{\min}), (x_{\max}, y_{\max})$, 类别 2, ...}。在本文中, 称一阶段数据集为 dataset-1, 该数据集的部分图片展示如图 3 所示。在第一阶段电线杆及杆号区域检测数据集的基础上, 利用目标杆号区域的坐标裁剪出第二阶段的数据集, 经筛选后共有 7833 张图像, 并对其进行如上描述的标注工作, 其中 6697 张用作训练集, 353 张用作验证集, 783 张用作测试集。本文中, 称第二阶段数据集为 dataset-2, 该数据集的部分图片展示如图 4 所示。

由于采集了两条高铁线路上的电线杆图像, 因此数据集中存在两种差异较大的电线杆形状, 如图 3 所示。同时杆号区域所在图像的位置以及杆号位数也不同, 有竖向排列的四位数或三位数杆号, 也有横向排列的三位数杆号, 如图 4 所示。第一阶段模型是对电线杆及杆号区域进行二分类, 第二阶段模型是对数字区域所包含的数字进行十分类, 均使用 LabelImg 将其真实标注保存为 xml 文件, 然后对目标的真实标注类别及边界框坐标进行提取, 保存为 txt 文件用于训练、验证及测试环节。



Figure 3. The partial images in dataset-1

图 3. dataset-1 部分图像



Figure 4. The partial images in dataset-2

图 4. dataset-2 部分图像

4.2. 实验环境

高铁最优电线杆检测及杆号识别系统训练与测试的 GPU 环境为 NVIDIA GeForce RTX 2060, 处理器为 intel i7-10875H, 操作系统为 Windows10, 编译环境为 Python3.6。我们的模型架构在 Tensorflow [19]上部署, 深度学习参数配置平台为 Keras, 并同时安装了 CUDA 9.0, CUDNN 7.0.4 以支持 GPU 的使用。

4.3. anchor 方案对比实验

在小型数据集(320 张图像)上进行的实验证明, 重聚类加线性放缩所获得的新锚框, 对于检测精度有很大的提升, 对比情况如表 1 所示。我们以平均准确率(mAP)和召回率(Recall)作为衡量标准。从表 1 第一行中可以看出, YOLOv3-tiny 在 VOC 数据集上聚类所得的锚框, 并不适用于我们所使用的电线杆数据集, 主要原因在于 VOC 数据集中样本的目标种类多, 尺寸不一, 大小差距较大, 聚类出来的先验框形状不一, 充分发挥了 YOLOv3-tiny 多尺度特征检测的优势。

由于我们第一阶段的检测目标是电线杆及杆号区域, 类别少, 且这两类目标大小差异大, 因此直接使用默认的锚框用于训练我们的电线杆数据集时显然不合理。如表 1 第一行结果所示, 两类目标的平均检测精度仅有 66.93%, 无法满足工业检测需求。所以, 我们尝试利用 K-means 聚类算法对第一阶段电线杆数据集进行重新聚类, 获得适合第一阶段电线杆数据集的先验框, 然而经实验发现仅聚类对于模型的提升并无效果, 我们可以注意到仅通过聚类而无进行线性缩放时的电线杆检测平均精度为 0.00, 经分析, 我们认为主要是由两点原因造成的, 首先在小型数据集上训练数据过少导致模型学习特征不充分, 其次在数据集信息描述中(4.3 节), 可以发现数据集中电线杆形态差异较大, 导致重聚类所得锚框尺寸单一, 从而使得多尺度目标检测器 YOLOv3-tiny 模型无法体现多尺度检测的优势, 致使模型检测效果较差, 因此我们通过结合 3.2 节描述的线性缩放机制, 对锚框进行再处理, 从而离散尺寸过于集中的锚框, 进一步提升电线杆(大目标)、杆号区域(小目标)的检出率。如表 1 第三行结果所示, 通过聚类算法结合线性缩放机制, 这两类目标的平均检测精度相较默认 anchors, 重聚类 anchors 分别提升了 16.74%、59.59%, 由此对比实验结果可以看出对于高铁最优电线杆自动检测任务而言, 重聚类结合线性缩放极大地提升了模型的学习效果及检测精度。我们将在后续的大型数据集上继续沿用此方案。

Table 1. Influence from different anchor box set for the first stage of the model

表 1. 不同的锚框设置对于第一阶段模型性能的影响

锚框方案	杆号区域/AP	电线杆/AP	mAP
YOLOv3-tiny默认锚框	75.52%	58.33%	66.93%
重聚类锚框	48.16%	0.00%	24.08%
重聚类加线性缩放	80.59%	86.74%	83.67%

4.4. 两阶段实验结果及分析

由于我们的实验将电线杆检测及杆号识别分为两阶段进行, 因此分别从两个阶段分析实验结果。首先在第一阶段检测过程中, 主要针对电线杆以及杆号区域进行检测, 并自动裁剪存在两个目标的数字区域, 保证电线杆与杆号区域同时检测到, 这样才可以达到最优杆检测的目的, 同时数字杆号完整清晰也是电线杆完整清晰的标准之一。

从表 2 中可以看出, 在保持高召回率的基础上, 两阶段模型也保持着相对较高的平均准确率, 两阶段模型的测试定性结果如图 5 所示。从图 5(a1)~(a4)中可以看出通过大量的训练数据, 模型对于不同形状

电线杆的特征学习情况良好，能区分不同特征的电线杆，并对杆号区域也实现了较精准的定位。第一阶段模型对于电线杆及杆号区域的识别定位准确率分别达到了 97.50%、94.91%。在第二阶段模型中，十类数字目标的识别及定位准确率达 95.30%。

Table 2. Detection result of two-stage model

表 2. 两阶段模型的测试结果

定量结果	检测目标	检测与识别准确率	平均准确率	召回率
第一阶段	电线杆	97.50%	92.27%	93.00%
	杆号区域	94.91%		
第二阶段	数字杆号	95.30%	96.42%	97.10%

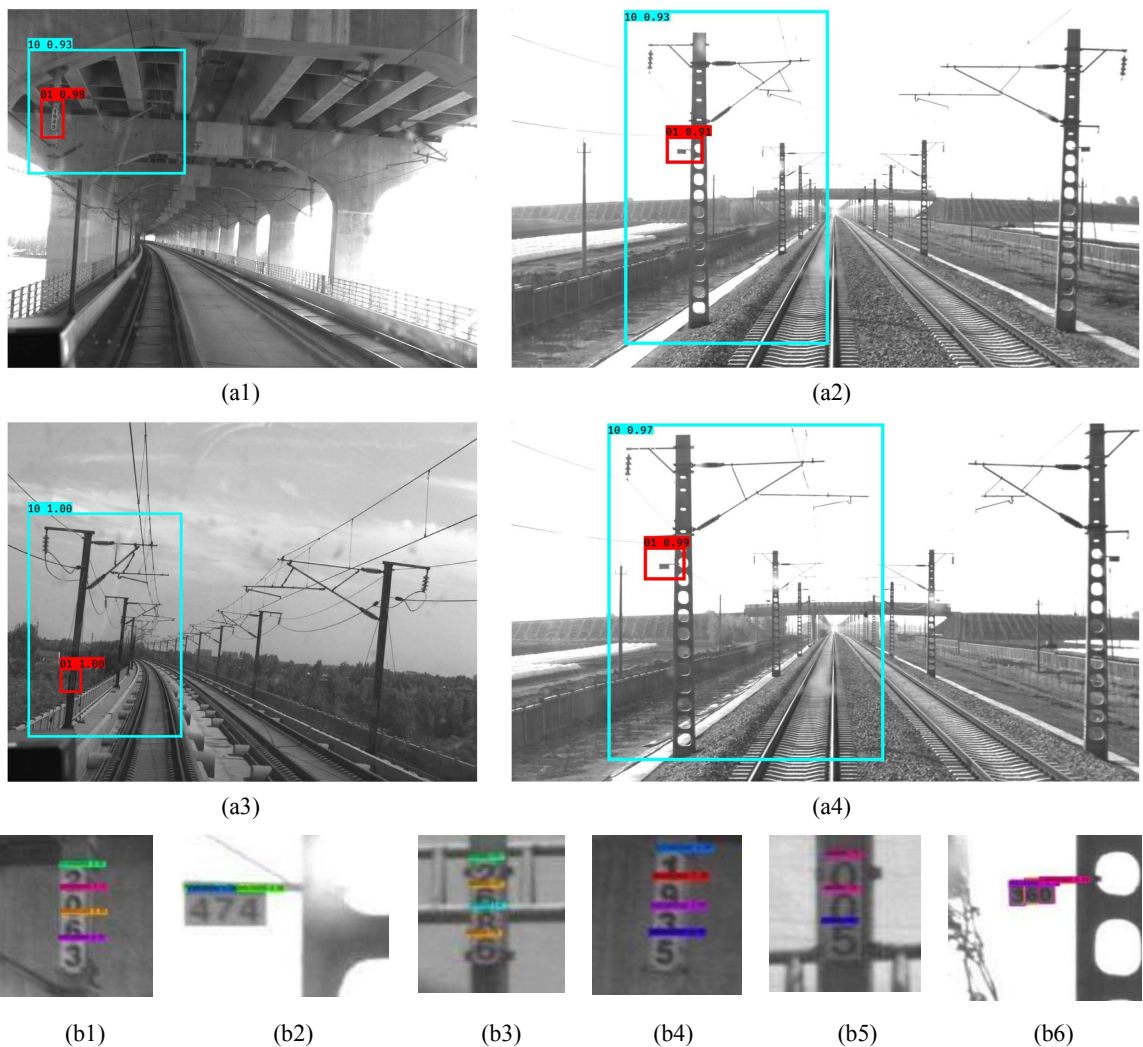


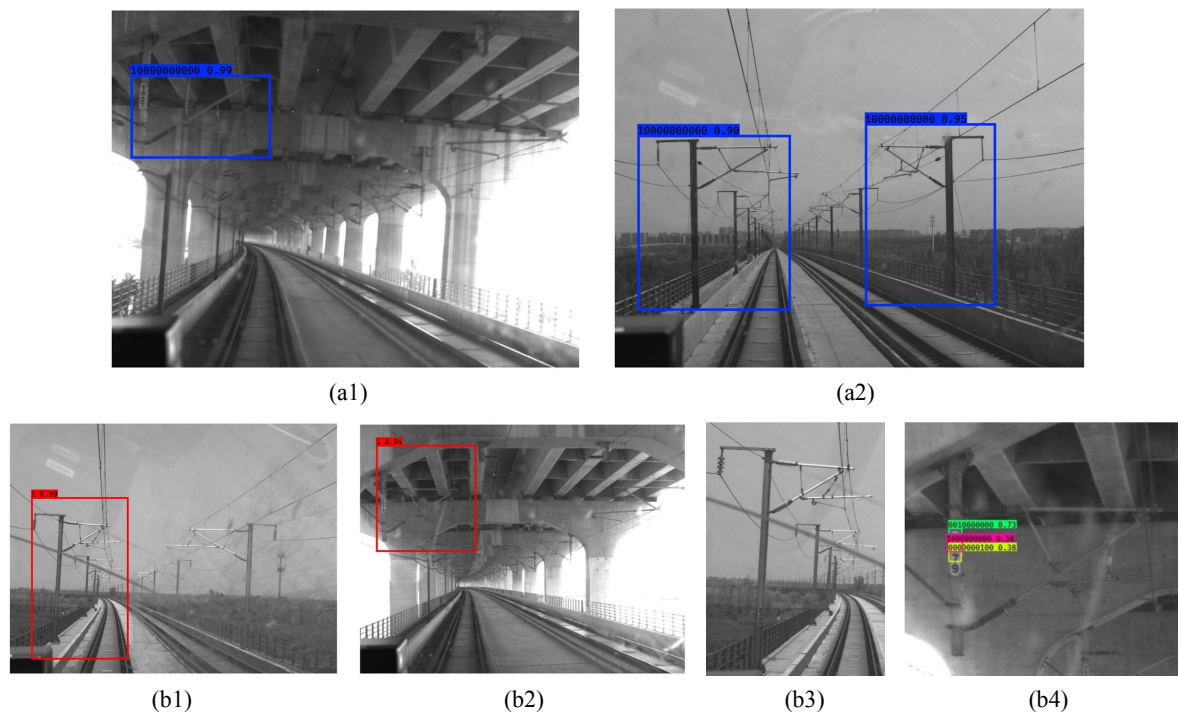
Figure 5. Detection results of two-stage on real world. (a) Results of first stage; (b) Results of second stage

图 5. 两阶段实际场景检测结果。(a) 第一阶段检测结果；(b) 第二阶段检测结果

下面我们给出以往研究方案的定性结果及分析，如图 6 所示，其中图 6(a1)~(a2)给出了我们最初的检测算法，在一个阶段内进行十一类检测(电线杆、十类数字)，从所给检测结果例图可以看出，模

型能够很好地捕捉学习大目标电线杆的特征，然而对于数字这类小目标无法检出；图 6(b1)~(b4)给出了两阶段初步检测算法结果，与我们设计算法的不同之处在于第二阶段是对第一阶段完整的电线杆目标检测结果进行自动裁剪，并送入第二阶段识别杆上数字，经分析，只有裁剪后局部电线杆的数字才能识别，如图 6(b4)所示，相较而言图 6(b3)上的数字没有检出，这是由于局部电线杆的数字目标相对较大，易识别，全局电线杆上数字无法检出，是由于数字与电线杆目标差异较大，数字小目标相对不易识别；经分析以上方案检测结果存在的不足之处，第三行(图 6(c1)~(c6))是在以上设计算法的基础上，通过仅关注局部杆号牌区域从而缩小杆号区域的标注范围，手动裁剪出杆号牌的大致区域，并在这一区域上标注十类数字目标得到数字数据集，送入模型训练并检测，如此通过缩小检测范围获得的结果如图 6(c1)~(c9)所示，可以看出数字识别效果得到显著提升，因此我们可以得到一个猜想：如果将电线杆检测与杆号识别任务分为两个阶段进行，并且在第一阶段可以对局部杆号牌区域进行自动检测并裁剪，那么当送入第二阶段时数字识别效果将非常理想，最后将两阶段结果进行串联，便可以实现高铁电线杆的自动检测和杆号识别任务。

由以上实验方案及相关猜想铺垫，图 6 第四行(图 6(d1)~(d2))中的检测算法将第一阶段设计为二分类检测任务，分别为电线杆及局部杆号牌区域两类，然而完整电线杆上杆号区域的检测效果较差，原因在于此类图像亮度低，杆号区域特征不明显；第五行(图 6(e1)~(e2))是在以上算法的基础上继续改进，放大杆号区域的真实边界框，具体规则为杆号区域边界框高低不变，宽度适当加大，从而添加背景辅助特征以提升检测精度，即我们的最终算法设计方案，如第 3 节描述，定性结果如图 5 所示，定量结果如表 2 所示，可以看出与以往实验方案相比，第一阶段电线杆、杆号区域检测结果及第二阶段杆号识别结果均得到了较大提升，然后通过逻辑编码的方式，将两阶段模型结合便能有效地完成高铁最优电线杆及杆号的自动检测与识别任务。需要注意的是，在第一、第二、第三行的初步检测算法中我们没有使用重聚类及线性缩放，我们设想如果加入以上调整方案，实验效果会得到明显提升，但仍不足以满足工业需求，因此在调整为两阶段任务时我们方才引用了重聚类及线性缩放。



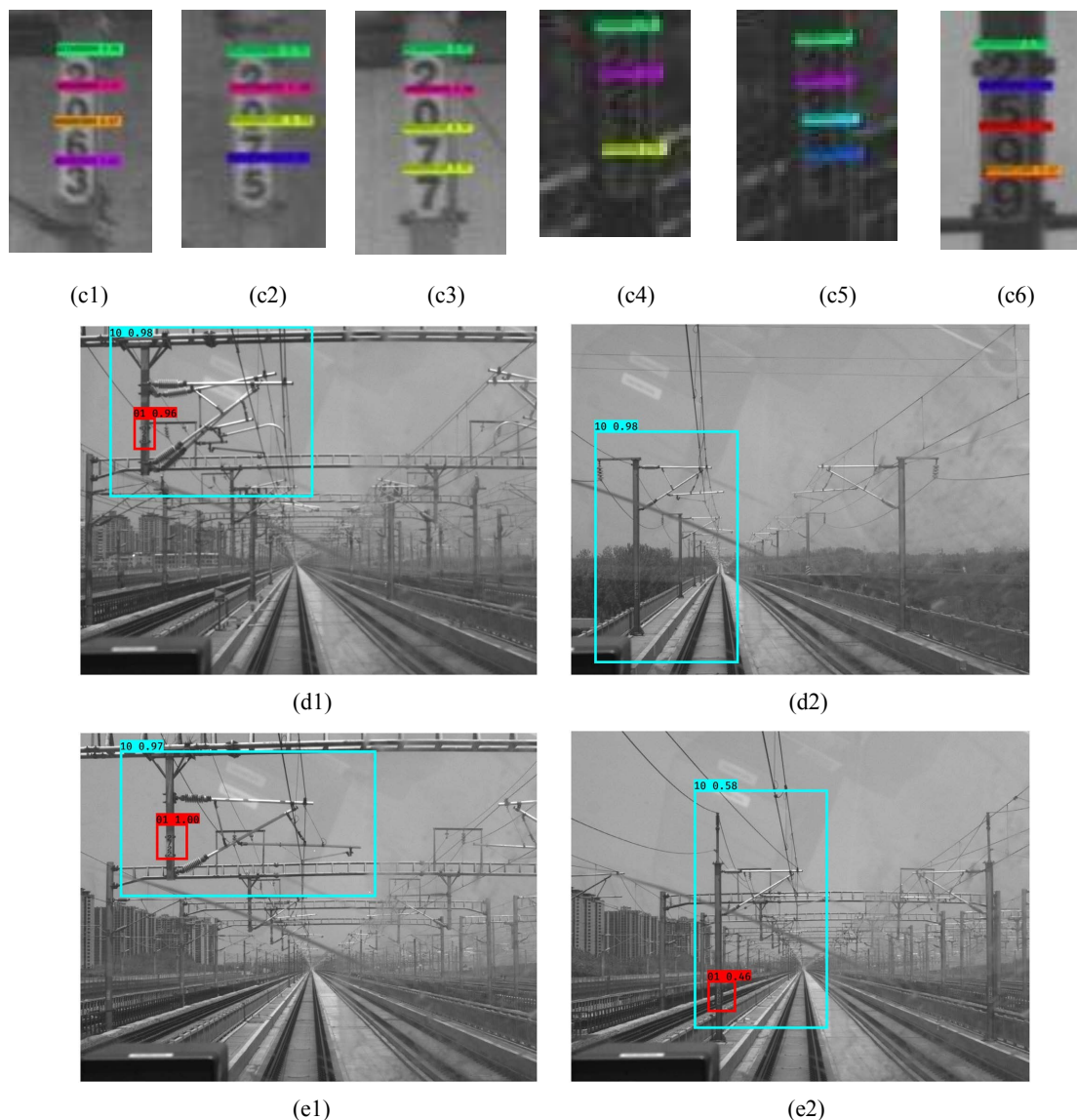


Figure 6. The qualitative results of different experiment design. (a) Detection results for eleven categories (poles and ten numbers); (b) Preliminary detection results of two-stage model (b1, b2): Detection results of one category for first stage; (b3, b4): Recognition results of numbers for second stage); (c) Manual narrowing of the pole number area and recognition results of numbers; (d) Detection results of two categories (poles and number areas) for first stage; (e) Detection results of two categories (poles and number areas) for first stage after enlarge the number areas

图 6. 不同实验设计的定性结果。(a) 十一类(电线杆、十类数字)检测结果; (b) 两阶段模型初步检测结果(b1, b2): 第一阶段一类检测结果; (b3, b4): 第二阶段数字识别结果); (c) 手动缩小杆号区域及数字识别结果; (d) 第一阶段二分类(电线杆、杆号区域)检测结果; (e) 放大杆号区域后第一阶段二分类(电线杆、杆号区域)检测结果

4.5. 模型消融实验

前面提到第一阶段电线杆数据集的构建采用了两条高铁线路上的电线杆图像，并且电线杆的形态差异较大，因此为了定量的衡量模型对两种不同形态的电线杆特征的学习情况及检测效果，我们分别针对单条线路上的电线杆图像进行测试，测试结果如表 3 所示。因数据保密原则，我们将两条线路分别命名为线路一、线路二。

从表 3 的结果中我们可以看出，模型在第一条高铁线路上的检测精度相比混合数据平均检测精度高

出 2.6%，表明该模型对于线路一上的电线杆特征学习效果理想，符合工业检测要求。同时为了提升模型的泛化能力，我们对第二条高铁线路的电线杆图像同时进行学习并测试，然而相较混合数据平均检测精度降低了 3.58%，也基本符合工业检测要求，说明我们设计所采用的模型泛化能力基本满足需求，但仍有待提升。值得注意的是，高铁线路二的电线杆目标检测精度较高，而杆号区域检测精度较低。经分析，如图 3，datase-1 第二张图像所示，杆号区域为电线杆左侧的一个小型矩形牌(可以放大后进行观察)，相较于整个图像而言，特征较小且模糊不易识别，未来我们会对目标检测领域中的小目标检测进行研究并对算法进行改进，以提升线路二上小目标的检测精度，进一步提升模型的泛化能力。

Table 3. Detection results of two-stage model on two high-speed routes

表 3. 两阶段模型在两条线路的测试结果

高铁线路	杆号区域/AP	电线杆/AP	mAP
线路一	98.13%	91.62%	94.87%
线路二	81.51%	95.86%	88.69%

5. 结论及未来研究方向

在这项工作中，我们分析了最优电线杆检测及杆号识别这一任务的关键问题，即由于电线杆与数字目标差异较大，因此无法同时检出。为此，在保证检测速率的基础上，以 YOLOv3-tiny 目标检测算法为主干，将最优杆检测与杆号识别分为两阶段进行，首先使用 K-means 聚类算法重新确定适合第一阶段电线杆数据集的先验框尺寸，使模型更符合电线杆目标的检测，提高模型预测目标坐标的能力；其次，对先验框做了线性缩放，使其更具备多尺度检测的优势，获得更高的检测精度；最后通过多次标签调整、模型调参及逻辑编码，实现了两阶段串联的设计，自动获得最优电线杆检测及杆号识别结果。电线杆的检测和建档极大地方便了对电线杆的管理和监测，也解决了大量数据的存储占用空间过大问题。

另外，我们的方法也存在一些不足之处：首先，对于亮度低，对比度低，图像模糊，特征不明显的图像，我们的方法检测精度较低，对于电线杆与杆号区域的特征学习造成一定难度，因此在未来的研究中，我们计划进一步考虑如何学习对比度低、亮度低的模糊图像特征，从而进一步提高检测精度。其次，由于数字过小和采集图像模糊，我们的方法对于一些形状类似数字(如 3 和 8，5 和 3)会产生一些误分类的情况。在实际应用中，不同高铁线路上的电线杆形状各异，因此还需要提升模型的泛化能力，使其能广泛应用于不同的电线杆检测项目，提升推广前景。此外，电线杆上的鸟巢极大地影响着电力系统的正常运转，对电线杆上的鸟巢进行检测也是我们后续的一个研究方向。

基金项目

这项研究由以下三个基金支持研究：国家自然科学基金(No. 41771413)，北京市自然科学基金(No. 8202013)。2022 年北京建筑大学研究生创新项目(No. PG2022145)。

参考文献

- [1] Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. (2016) You Only Look Once: Unified, Real-Time Object Detection. *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, 27-30 June 2016, 779-788. <https://doi.org/10.1109/CVPR.2016.91>
- [2] Redmon, J. and Farhadi, A. (2017) YOLO9000: Better, Faster, Stronger. *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, 21-26 July 2017, 6517-6525. <https://doi.org/10.1109/CVPR.2017.690>
- [3] Redmon, J. and Farhadi, A. (2018) YOLOv3: An Incremental Improvement. *Computer Science*. arXiv: 1804.02767.

- <http://arxiv.org/abs/1804.02767>
- [4] Ren, S.Q., He, K.M., Girshick, R. and Sun, J. (2017) Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans on Pattern Analysis and Machine Intelligence*, **39**, 1137-1149. <https://doi.org/10.1109/TPAMI.2016.2577031>
- [5] Girshick, R. (2015) Fast R-CNN. *Proceedings of the 2015 IEEE International Conference on Computer Vision and Pattern Recognition*, Santiago, 7-13 December 2015, 1440-1448. <https://doi.org/10.1109/ICCV.2015.169>
- [6] Girshick, R., Donahue, J., Darrell, T. and Malik, J. (2016) Region-Based Convolutional Networks for Accurate Object Detection and Segmentation. *IEEE Trans on Pattern Analysis and Machine Intelligence*, **38**, 142-158. <https://doi.org/10.1109/TPAMI.2015.2437384>
- [7] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., et al. (2016) SSD: Single Shot MultiBox Detector, *European Conference on Computer Vision 2016*, Amsterdam, 11-14 October 2016, 21-37. https://doi.org/10.1007/978-3-319-46448-0_2
- [8] 谢兴阳, 余阳, 任峰. 基于窄带频率激励和先进信号处理技术的新型电线杆无损检测系统研究[J]. 自动化与仪器仪表, 2017(1): 105-106+109. <https://doi.org/10.14016/j.cnki.1001-9227.2017.01.105>
- [9] 李涛, 陈黎, 聂晖. 基于改进线段分割检测的电线杆遮挡检测算法[J]. 计算机工程, 2017, 43(9): 250-255. <https://doi.org/10.3969/j.issn.1000-3428.2017.09.044>
- [10] 江南, 李怡然, 黄毅标, 张海滨, 孔令一, 黄超, 等. 基于绝缘子的高空电线杆损坏检测技术[J]. 机械设计与制造工程, 2019, 48(3): 107-111. <https://doi.org/10.3969/j.issn.2095-509X.2019.03.025>
- [11] 刘志浩, 冯柳平, 曹晓鹤. 基于深度学习的电线杆检测方法[J]. 北京印刷学院学报, 2016, 24(6): 44-47+52. <https://doi.org/10.19461/j.cnki.1004-8626.2016.06.011>
- [12] 庞宁. 基于深度学习的输电线杆塔鸟巢检测与识别[J]. 自动化与仪器仪表, 2020(4): 195-198+204. <https://doi.org/10.14016/j.cnki.1001-9227.2020.04.195>
- [13] 刘凯歌, 王琪, 孟祥越, 张祥德. 基于改进 Faster R-CNN 的被遮挡电线杆检测算法[J]. 无线电工程, 2021, 51(7): 540-545.
- [14] Liu, J. and Zhang, D.Q. (2020) Research on Vehicle Object Detection Algorithm Based on Improved YOLOv3 Algorithm. *Journal of Physics: Conference Series*, **1575**, Article ID: 012150. <https://doi.org/10.1088/1742-6596/1575/1/012150>
- [15] Dou, H.Z., Zhang, H.Y. and Li, B. (2021) A Fast Traffic Sign Detection Algorithm Based on Modified YOLOv3. *Journal of Physics: Conference Series*, **1880**, Article ID: 012025. <https://doi.org/10.1088/1742-6596/1880/1/012025>
- [16] Gao, D.Y., Gao, T.Y., Shi, S.S., Zhang, Z.C. and Ding, Z. (2020) Research on Traffic Information Detection of the Visually Impaired Based on Improved YOLOv3. *Journal of Physics: Conference Series*, **1802**, Article ID: 032026. <https://doi.org/10.1088/1742-6596/1802/3/032026>
- [17] Hartigan, J.A. and Wong, M.A. (1979) Algorithm AS 136: A K-Means Clustering Algorithm. *Journal of the Royal Statistical Society*, **28**, 100-108. <https://doi.org/10.2307/2346830>
- [18] Shelhamer, E., Long, J. and Darrell, T. (2017) Fully Convolutional Networks for Semantic Segmentation. *IEEE Trans on Pattern Analysis and Machine Intelligence*, **39**, 640-651. <https://doi.org/10.1109/TPAMI.2016.2572683>
- [19] Abadi, M., Agarwal, A., Barham, P., et al. (2016) Tensorflow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. arXiv Preprint arXiv:1603.04467.