

# 铁路接触网支柱的图像序列自适应识别方法

黄竹安<sup>1</sup>, 宋浩然<sup>2</sup>, 王浩然<sup>3</sup>, 刘俊博<sup>2</sup>, 顾子晨<sup>4</sup>, 戴鹏<sup>2</sup>

(1. 中车南京浦镇车辆有限公司, 南京 210031;

2. 中国铁道科学研究院集团有限公司基础设施检测研究所, 北京 100081;

3. 中国铁道科学研究院, 北京 100081; 4. 北京铁科英迈技术有限公司, 北京 100081)

**摘要:** 接触网支柱数字化管理是电气化铁路运维的关键环节, 基于移动视频建立接触网支柱数字台账是高效、经济、便捷的技术手段; 为实现对于移动视频图像序列中接触网支柱杆号的精准识别, 提出了一种基于区域相关和改进 SVTR 网络的接触网支柱识别算法; 针对视频图像中接触网支柱区域重叠、结构模式复杂的特点, 采用了 YOLO v4 网络对单帧图像中支柱区域和号牌标识区域分别进行检测, 并通过测算交叠区域来获得距观察点最近的杆位和对应的号牌区域; 此外, 针对接触网杆号牌尺度多样性和字符变长的问题, 在杆号文字识别问题中采用了 SVTR-tiny 网络, 并进一步引入迁移学习方法增强模型对于复杂杆号的识别精度和对于不同线路场景的泛化性能; 通过在实际高铁线路采集的移动视频数据集上进行测试, 结果表明算法在移动视频中视野最近杆位杆号区域的定位检出率可达 98.01%, 杆号文本的识别准确率达到 96.13%, 适用于我国高速铁路主要干线建设配套的接触网支柱结构。

**关键词:** 接触网支柱识别; 移动视频; YOLO v4 网络; 区域相关; SVTR-tiny 网络

## Image Sequence Adaptive Recognition Method for Railway Catenary Pillar

HUANG Zhu'an<sup>1</sup>, SONG Haoran<sup>2</sup>, WANG Haoran<sup>3</sup>, LIU Junbo<sup>2</sup>, GU Zichen<sup>4</sup>, DAI Peng<sup>3</sup>

(1. CRRC Nanjing Puzhen Co., Ltd., Nanjing 210031, China;

2. Infrastructure Inspection Research Institute, China Academy of Railway Sciences Co., Ltd., Beijing 100081, China;

3. China Academy of Railway Sciences, Beijing 100081, China;

4. INMAI Railway Technology Co., Ltd., Beijing 100081, China)

**Abstract:** Digital management of the contact network pillars is a critical component of operating and maintaining electrified railways. Creating a digital ledger of catenary pillars based on mobile video is an efficient, cost-effective, and convenient technological method. To achieve accurate recognition of the plate numbers on catenary pillars in mobile video image sequence, we propose a catenary pillar recognition algorithm based on region correlation and an improved SVTR network. Firstly, to address the challenges posed by overlapping contact network pillar areas and complex structural patterns in video images, we use the YOLO v4 network to separately detect catenary pillar areas and number plate areas in single-frame images. The nearest catenary pillar position and corresponding number plate area are then determined by calculating overlapping areas. Furthermore, we use an improved SVTR-tiny network for catenary pillar number text recognition to address the issues of diverse pillar number plate scales and variable character length, and transfer learning methods are introduced to enhance the model's recognition accuracy for complex plate numbers and its generalization performance to different line scenes. Through testing on a mobile video dataset collected from actual high-speed railway lines, the test results show that the algorithm achieves a positioning recall rate of 98.01% for the nearest pillar number in the field of view and a pillar number text recognition accuracy of 96.13%, and the method is suitable for the catenary pillar structure supporting the construction of major high-speed railway trunk lines in China.

**Keywords:** catenary pillar recognition; mobile video; YOLO v4 network; regional related; SVTR-tiny network

## 0 引言

接触网支柱是电气化铁路中的重要基础设施, 用于支撑接触网, 向列车提供电力。接触网支柱一般由钢制或混凝土制成, 高度通常为 8~10 m, 根据实际需要也有更高的支柱。它们以大约 50 m 的间距分布在轨道两侧, 在接触网

线路的弯道处密集排列, 并以自定义格式编号标识。构建接触网支柱的数字化数据库可以帮助铁路管理部门更加准确、快速地了解接触网支柱的分布、数量、状态等信息, 便于及时发现和解决问题。通过动态普查手段构建接触网支柱的数字数据库, 是实现铁路基础设施台账数字化管理,

收稿日期: 2023-03-31; 修回日期: 2023-04-10。

基金项目: 国家自然科学基金项目(52272427); 中国国家铁路集团有限公司科研计划项目(K2021T015); 中国铁道科学研究院集团有限公司院重大基金课题(2022YJ256)。

作者简介: 黄竹安(1976-), 男, 大学本科, 高级工程师。

通讯作者: 戴鹏(1982-), 男, 博士, 研究员。

引用格式: 黄竹安, 宋浩然, 王浩然, 等. 铁路接触网支柱的图像序列自适应识别方法[J]. 计算机测量与控制, 2023, 31(10): 222-227.

构建数字孪生铁路的关键环节。接触网支柱数据的采集主要有采用移动视频拍摄<sup>[1]</sup>和激光雷达<sup>[2]</sup>扫描等手段。其中移动视频可以获取颜色、纹理等细节信息,且成像设备价格低廉、部署方便、易于操作,使其成为首选考虑的成像手段。接触网支柱识别的传统方法是由专业技术人员进行外观检查和维护,这种方法操作过程繁琐,效率低下,而且在高速列车运行的情况下存在安全隐患。所以,基于高速移动视频的接触网支柱自动识别技术开始受到广泛的关注。

基于移动视频进行接触网支柱的自动识别主要存在以下难点:1)高速移动视频背景多变,且视频图像中可能随机出现多根支柱,而一般只需要关注距拍摄观测点最近的支柱就可以满足需求,同一场景中出现的多个同类目标会影响机器识别准确性;2)接触网支柱的号牌在单幅图像中占比极小,且不同结构接触网号牌的样式、字体、格式均存在差异,字符文本也表现为长度不定,数字和英文字母混合等特点,因此要求识别系统具备准确定位的能力和适应复杂文本的识别能力。

针对接触网支柱及杆号识别问题,国内外学者开展了广泛的研究。武翔宇提出结合支持向量机利用灰度直方图特征对接触网杆号进行定位的方法,通过优化算法实现精细分割,根据图像灰度统计值准确提取单个字符图像并分别进行识别<sup>[3]</sup>。陈忠革提出基于特征算子提取支柱和杆号区域,将杆号文本分割为单个文字,最终采用 HOG 特征和支持向量机实现字符识别的方法<sup>[4]</sup>。杨文静为消除采集图像质量差异带来的影响,提高杆号文字识别准确率,提出了基于重心形状的上下文特征识别支柱杆号文字的方法<sup>[5]</sup>。杨梅提出了一种基于形状模板匹配的自适应策略算法,通过 Shape Context 匹配算法针对已知的模板字符及杆号图像进行杆号的识别与定位,并利用自适应策略算法在实时分析场景下更准确识别不同类型的杆号<sup>[6]</sup>。闵锋等采用 HOG 特征提取和 SVM 分类器的算法定位杆号牌位置,随后利用投影矩阵的极值点得到候选分割点,分析其距离的变异系数得到合理的字符分割点<sup>[7]</sup>。郑巧等针对杆号牌可能出现的遮挡、倾斜问题,提出了一种基于几何特征检测及自适应投影分割的接触网号牌字符提取方法,采用拆分检测与线条检测的方式得到不同号牌矩形结构的精细区域,再采用自适应投影分割方法对单个字符进行提取<sup>[8]</sup>。近年来,随着深度学习技术<sup>[9]</sup>在图像识别领域取得的巨大成功,一些学者开始将深度学习的方法应用于接触网的杆号识别任务上。徐可佳采用 Faster R-CNN 算法定位号牌单个字符进行识别得到杆号信息,通过特定图像处理算法扩充样本数据,使端到端的在线识别检测算法具有更好的适应性能力<sup>[10]</sup>。侯明斌等提出一种包含深度残差 YOLO 模型(DR-YOLO)定位模块和 SRM 识别模块的支柱号图像序列识别方法<sup>[11]</sup>,该方法泛化能力好,可以字符序列为整体进行识别,解决了分割字符过程中产生的误差,可实现对于不定长文本杆号的识别。在铁路运行安全保障领域,杨栋等提出了基于智能识别技术的铁路安检辅助分析方法与系统<sup>[12]</sup>,

在多线程思路设计当中采用了改进的残差网络模块和 CenterNet 算法,实现了更有效的特征提取和检测。

尽管现有方法已经初步实现了对于接触网杆号区域进行定位并对于杆号信息进行自动识别的功能,但从应用效果来看,现有智能识别系统的精度难以适应不同类型线路接触网杆号模式的复杂多样性,存在误报率较高、泛化性不强等问题。为此,本文提出了一种针对高速移动视频中接触网支柱和杆号区域表现特性的识别算法,并在我国高速铁路实际线路数据上进行实测应用验证,证明了算法的有效性。该算法针对高速移动视频中接触网支柱的复杂多样性进行了设计,能够准确地识别接触网支柱及其杆号。本算法的应用可优化铁路管理部门对铁路基础设施台账数字化管理和数字孪生铁路的建设。

## 1 算法原理

算法的任务目标为从视频单帧图像中识别得到距拍摄观察点最近的接触网支柱,并识别出对应杆号。图 1 给出了 3 种不同接触网支柱结构的视频截图。从 3 张图中可以看出,前向拍摄视频中会出现多个支柱。不同接触网支柱结构,杆号牌的安装位置也存在差异。



(a) 杆号牌位于支柱本体 (b) 杆号牌位于支柱横腕臂 (c) 杆号牌位于支柱外侧

图 1 3 种不同结构接触网支柱及杆号标识牌图像

为了实现算法的目标,该算法设计为两步。首先,在第一步中,通过定位算法模型来定位整幅图像中的支柱区域,并使用区域相关分析技术计算这些区域的交叠区域,以获得距离拍摄观察点最近的支柱和对应的号牌区域。这个过程旨在实现支柱及其对应号牌的区域定位,以便后续处理。在第二步中,基于前面定位出的杆号,使用 OCR 文本识别模型实现对于接触网支柱杆号牌文字信息的识别。此步骤利用深度学习技术,将图像中的杆号牌转换为电子文本格式,以便进一步的处理和分析。整个算法的设计考虑到了视频单帧图像中存在多个支柱、不同结构导致的号牌安装位置差异等特点,并通过两个步骤的处理实现了对接触网支柱及其号牌的准确定位和识别。

算法流程如图 2 所示。

### 1.1 支柱及号牌区域定位

在移动视频图像具有待识别目标多、号牌目标较小的问题。传统的杆号牌定位算法通常采用首先定位接触网支

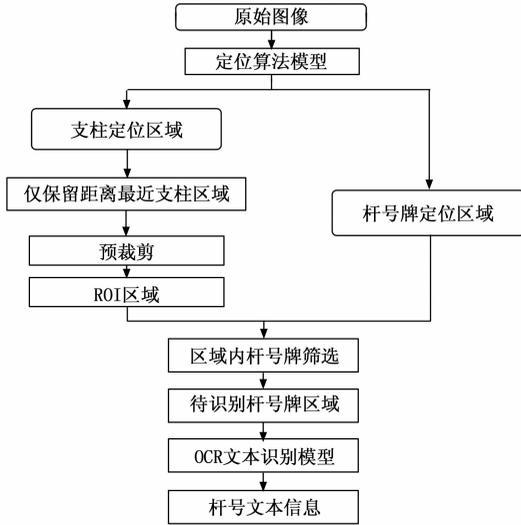


图 2 接触网杆号识别算法流程

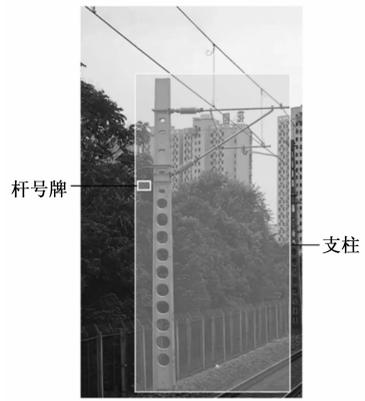


图 3 接触网区域定位标注示意图

柱区域, 然后根据预设相对位置关系确定号牌区域的方式进行定位。然而, 这种方法在应用时需要较多的人工干预, 并且存在对不同线路的泛化性和适应性不足的问题。而本文提出的方法采用多区域相关性分析的方法, 实现杆号及对应标识牌的自适应联合定位。具体来说, 该方法利用图像处理技术, 在整个图像中找到多个可能属于接触网支柱的区域, 并结合杆号牌的特征信息, 通过多区域相关性分析的方式自适应地确定每个支柱的杆号牌区域。具体原理如下:

1) 数据标注: 以前述提到的我国高速铁路主要干线接触网支柱视频作为基础数据, 根据不同结构和号牌标识的结构特点, 对接触网支柱区域、号牌区域作为两类检测目标对象进行标注, 如图 3 所示。

2) 区域定位模型: 对标注好的数据集采用 YOLO v4<sup>[13]</sup> 进行训练, 得到一个包含 2 个检测类别 (支柱、杆号牌) 的目标检测模型, 网络模型的结构如图 4 所示。

YOLOv4 模型的主干网络采用了 CSPDarknet53<sup>[14]</sup>, 沿用了 Darknet53 网络的滤波器尺寸和整体结构, 在每组跳跃连接加上一个跨阶段部分结构, 减少了模型整体参数且保留了图像在浅层网络的小目标的形态特征, 提升模型对杆号牌区域识别效果, 并快速实现对于杆号牌区域和接触网支柱的定位。

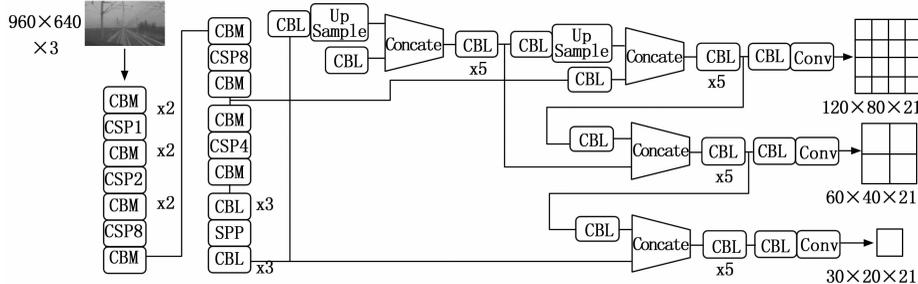


图 4 YOLOv4 模型网络结构

模型的特征融合模块采用了 PANet (path aggregation network)<sup>[15]</sup> 与 FPN (feature pyramid network)<sup>[16]</sup> 结合的结构, PANet 在 FPN 网络的结构基础上做出了改进, 网络加入了自底向上的融合结构, 原有的浅层特征经过较少的层数和高层次特征加以融合, 加深模型对接触网图像中目标形态特征和语义特征的理解, 从而同时提升对于尺寸较大的支柱和尺寸较小目标的杆号牌目标的识别效果。

此外, YOLOv4 还引入了 SAM (spatial attention module) 模块, 它可以自适应地调整特征图中不同区域的重要性, 以便更好地识别小目标。SAM 模块通过计算特征图中每个位置的平均值和最大值, 确定了不同区域的重要性, 并将其作为权重应用到特征图上。

为提升对图像的感知效果, YOLOv4 还采用了 SPP (spatial pyramid pooling) 结构, 该结构可以在固定的感受野内提取多个不同尺度的特征。这种方法可以增加特征图的维度, 并且可以在不改变感受野大小的情况下适应不同尺寸的物体。

在后处理方面, YOLOv4 使用 Distance-IoU NMS (DIOU-NMS) 进行目标框筛选。DIOU-NMS 考虑了目标框之间的距离和 IoU 值, 从而在 NMS 过程中更加准确地判断哪些框需要被保留下来。

YOLOv4 中采用了一系列创新的技术, 从而具有高效、准确、全面和实时等优点, 适用于本文支柱及号牌区域的定位问题。

3) 区域相关分析: 首先确定图像中与观察点最近的支柱。依据目标与观测点越近区域像素面积占比越大的原则, 取面积最大的支柱检测区域作为定位对象, 然后设计算法确定与之对应的杆号牌。为克服图像中存在多个支柱和号牌影响定位准确性的问题, 按照支柱主体结构将支柱图像裁剪为支柱外侧、横腕臂、支柱本体 3 个子区域, 如图 5 所示。

以下通过分析号牌区域与 3 个

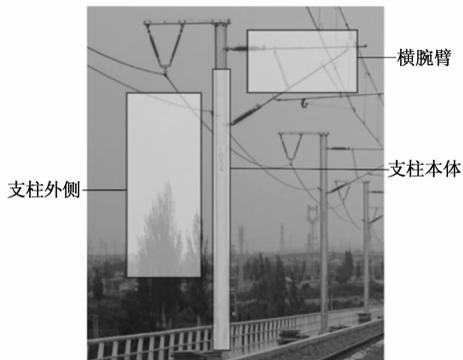


图 5 接触网支柱杆号牌 ROI 区域划分

裁剪区域的相关性确定支柱结构类型, 计算原理如下:

假定杆号牌区域为  $A$ , 裁剪区域定义为集合  $B$ , 则  $A$  和  $B$  可由式 (1) ~ (2) 表示:

$$A = \{(x, y) \mid x_{a_1} \leq x \leq x_{a_2}, y_{a_1} \leq y \leq y_{a_2}, x, y \in R\} \quad (1)$$

$$B = \{(x, y) \mid x_{b_1} \leq x \leq x_{b_2}, y_{b_1} \leq y \leq y_{b_2}, x, y \in R\} \quad (2)$$

其中: 杆号牌区域  $A$  的在图像中左上顶点和右下顶点的坐标分为  $(x_{a_1}, y_{a_1}), (x_{a_2}, y_{a_2})$ , 感兴趣区域  $B$  的左上顶点和右下顶点的坐标分为  $(x_{b_1}, y_{b_1}), (x_{b_2}, y_{b_2})$ 。

则杆号牌区域  $A$  和裁剪区域  $B$  的交并比计算公式如式 (3) 所示:

$$IOU = \frac{A \cap B}{A \cup B} =$$

$$\frac{(\min\{x_{a_2}, x_{b_2}\} - \max\{x_{a_1}, x_{b_1}\}) \times (\min\{y_{a_2}, y_{b_2}\} - \max\{y_{a_1}, y_{b_1}\})}{(x_{a_2} - x_{a_1})(y_{a_2} - y_{a_1}) + (x_{b_2} - x_{b_1})(y_{b_2} - y_{b_1}) - (\min\{x_{a_2}, x_{b_2}\} - \max\{x_{a_1}, x_{b_1}\}) \times (\min\{y_{a_2}, y_{b_2}\} - \max\{y_{a_1}, y_{b_1}\})} \quad (3)$$

$$IoU_{mc} = \max(IoU_{11}, IoU_{12}, IoU_{13}, \dots, IoU_{i1}, IoU_{i2}, IoU_{i3}) \quad (4)$$

对于图像中模型定位出的全部杆号牌边界框, 将其与 3 个分别预设的 ROI 区域分别按上述公式计算  $A$  类和  $B$  类区域交并比, 并得到各区域之间的相关性表示集合, 见式 (4)。式中,  $IoU_{i1}, IoU_{i2}, IoU_{i3}$  分别表示定位出的第  $i$  个边界框与 3 个裁剪区域的交并比。

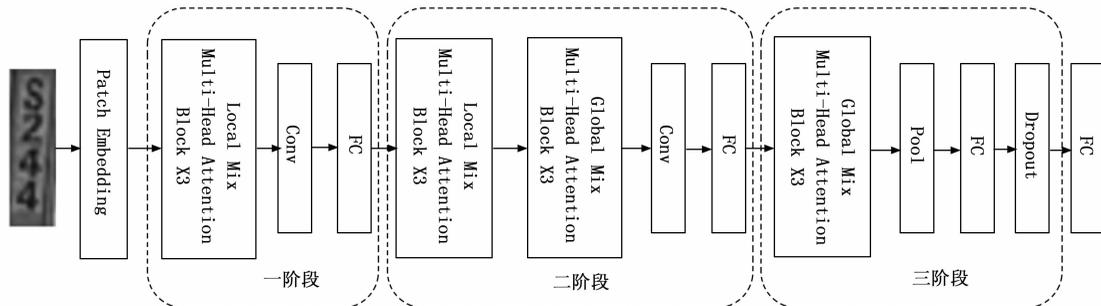


图 7 SVTR 杆号文本识别模型

取集合最大值元素对应的号牌区域并确定其对应裁剪区域, 即可确定支柱结构类型和待识别杆号文本。

### 1.2 支柱号牌识别

由于我国高铁线路类型多样, 不同类型的接触网支柱的号牌在文字样式、字体、文本长度、格式上均存在差异。图 6 给出了部分号牌示例。



图 6 部分线路杆号图像示例

本文设计一种能够自适应自动识别各类杆号文本的 OCR 模型, 该模型基于深度神经网络框架 SVTR-tiny<sup>[17]</sup> 构建。针对杆号牌文本的特点进行具体设计, 保证模型能够准确地识别不同类型的杆号文本内容。图 7 展示了该网络结构。

SVTR-tiny 网络采用了 Transformer 架构<sup>[18-19]</sup> 设计。首先将输入图像分成小的字符组件, 即文本块。之后采用组件级别的混合、合并和/或组合进行层次递归处理, 生成多粒度字符组件感知结果。感知模块中设计了全局和局部混合块以感知字符间和字符内部的模式。其中, 采用全局多头注意力机制混合模块用于提取号牌字符之间的相关性和文本与非文本的相关性, 采用局部的多头注意力机制混合模块用于提取字符的形态信息。网络的 3 个运算阶段功能如下: 第一阶段主要进行局部特征分析, 在较为浅层网络捕捉识别字符的形态信息; 第二阶段进行局部和全局区域的相关性预算, 提取字符形态信息和抽象的语义信息, 确定文本字符类型; 第三阶段进行全局区域的文字相关性语义信息提取。其后采用特征池化和全连接层进行特征组合后分类得到字符的识别结果。

SVTR-tiny 网络可通过单个视觉模型端到端地实现对于输入图像的文本识别, 通过自注意力机制直接实现对整幅图像中所有字符形态信息和图像字符相关性信息提取突破了传统的对单个字符进行分割提取后再进行特征提取和文本转录的限制, 可对不定长文本信息加以识别。

此外, 本方法针对杆号文字包含中英文的特点, 对研究中 SVTR-tiny 网络的基础数据集进行了扩展, 以提升模型对杆号文字的识别效果。首先将 SVTR-tiny 网络在大规

模中英文文本识别任务数据集上进行模型的预训练,之后再预训练后的模型在杆号文本图像数据集上进行迁移学习,从而使得该模型具备同时识别中英文文字的嫩体力,更好地适应杆号文本数字—英文混合,中文字符样式、字体多样性的特点。

## 2 实验与性能验证

通过在时速 300 km/h 高速综合检测列车动车组安装的高清摄像机采集的移动视频数据验证本文算法。数据包括 3 种不同类型的接触网支柱结构。其中第一种结构 2 012 张,第二种结构 2 314 张,第三种结构 2 145 张。选取 5 177 张混合后作为训练集,按照接触网支柱和号牌两类对象对训练集进行标注。选取 1 294 张作为测试集,用于对模型性能进行测试。

模型训练的实验计算平台硬件配置为 Intel Xeon@2.4GHz CPU、Nvidia Geforce RTX 3090 GPU×4,软件配置为 Ubuntu 18.04, Python 3.8, CUDA 11.3。

YOLOv4 定位模型的训练设置为每批次 64 张图片,采用 SGD 随机梯度下降训练策略,学习率设置为 0.001,总计进行 100 000 次迭代训练。SVTR 杆号文本识别模型每个训练批次设置为 128 张图片,采用 ADAM 优化算法<sup>[20]</sup>,一阶矩衰减系数设置为 0.9,二阶矩衰减系数设置为 0.99;学习率设置为 0.005,总计训练 100 000 次迭代训练。

### 2.1 检测指标

在对定位模型性能的评估中,选取接触网区域定位的检出率 (Recall) 和精确率 (Precision) 作为模型对于定位精确程度的评价指标,用于评价模型定位部件的能力。

检出率、精确率计算公式如下所示:

$$Recall = \frac{TP}{TP + FN} \times 100\% \quad (5)$$

$$Precision = \frac{TP}{TP + FP} \times 100\% \quad (6)$$

式中,  $TP$  为正确识别为对应类别区域的数目,  $FN$  为对应类别区域漏识的数目,  $FP$  为误报为对应类别区域的数目。

杆号文本模型的识别效果采用准确率作为评价指标计算公式如下所示:

$$Accuracy = \frac{1}{n} \sum_{i=1}^n I(f(x_i) = y_i) \quad (7)$$

式中,  $n$  表示测试样本数目;  $x_i$  表示输入第  $i$  张图片;  $f(x_i)$  表示模型预测杆号文本;  $y_i$  表示数据标签文本;  $I(\cdot)$  表示判断统计预测和实际标签是否一致的函数,当模型预测类别和数据标签类别一致时为 1,否则为 0。

### 2.2 测试集测试

将本文提出的接触网识别算法在移动视频上采集某三条实际线路的数据进行了测试,三条线路的杆号区域分别位于支柱本体、横腕臂、支柱外侧。实验测试结果评价了算法对于距离最近支柱的杆号区域的检出率、精确率以及检出杆号区域的文本识别的准确率,实验结果如表 1、图 8 所示。

表 1 实际线路接触网杆号算法实验结果 %

线路杆号区域类型	杆号区域检出率	杆号区域精确率	检出杆号区域的文本识别准确率
号牌位于支柱本体	98.77	95.48	96.86
号牌位于横腕臂	98.10	94.29	96.89
号牌位于支柱外侧	97.17	91.37	94.63
平均值	98.01	93.71	96.13

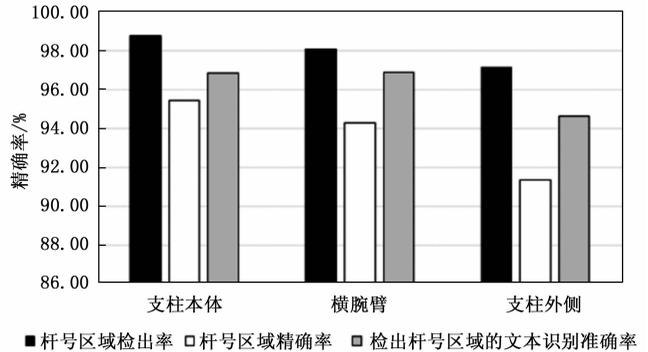


图 8 实际线路接触网杆号算法识别统计结果

实验结果表明,接触网杆号识别算法对于距离最近支柱杆号区域检出率均达到 97% 以上,精确率达到 90% 以上,检出杆号区域的文本识别准确率均达到 94% 以上,可以以较为精准的方式实现对于铁路接触网杆号的识别任务,表明该算法能够适应不同类型线路的杆号识别需求,在实际线路的识别中具有良好的可用性。对算法的识别结果进行分析可以得出,杆号区域定位出现漏报的原因多由于遮挡或杆号牌区域在图像中采集不完整引起,故相对于杆号牌区域位于横腕臂和支柱外侧,支柱本体区域的杆号牌检出率较高,而定位误报则为错误定位特征近似杆号牌区域的图像区域。文本识别错误则多由定位区域成像质量不佳或号牌遮挡引起模型识别错误。

算法在实际线路的部分杆号识别如图 9 所示,图片右下角为该支柱对应的接触网杆号牌识别结果。图 9 表明在支柱本体线路、支柱横腕臂线路和支柱外侧线路的不同杆号牌识别中,算法均获得了精确的识别结果。首先,针对支柱本体线路的杆号牌识别,算法的检出率和识别准确率均较高,这可能是由于支柱本体线路杆号牌的外观特征较为明显,杆号牌上的字符大小和排列方式相对固定,SVTR-tiny 网络可以较好地识别这种类型的杆号牌。同时,由于支柱本体线路的杆号牌在拍摄角度和距离上相对固定,因此 YOLOv4 网络也可以准确地检测出支柱区域和杆号牌标识区域,从而实现对该类型杆号牌的准确识别。

其次,对于支柱横腕臂线路的杆号牌识别,可能存在一些挑战,例如该类型杆号牌的尺寸和字符大小相对较小,排列方式也不太规则,这可能会影响 SVTR-tiny 网络的识别效果。此外,支柱横腕臂线路的拍摄角度和距离也可能与支柱本体线路不同,从而影响 YOLOv4 网络对支柱区域和杆号牌标识区域的检测效果。因此,在识别支柱横腕臂线路的杆号



(a) 杆号牌位于支柱本体线路



(b) 杆号牌位于支柱横腕臂线路



(c) 杆号牌位于支柱外侧线路

图 9 算法是实拍视频中的识别运用效果

牌时, 算法的检出率和识别准确率可能会稍微降低。

最后, 对于支柱外侧线路的杆号牌识别, 由于支柱外侧线路的杆号牌通常较为隐蔽, 位置比较靠外, 拍摄角度和距离也可能与支柱本体线路和支柱横腕臂线路不同, 因此, 算法的检出率和识别准确率可能会相对较低。针对这种情况, 可以考虑增加更多的训练样本以提高模型的泛化性能, 或者尝试优化算法的网络架构和参数以提高检测和识别效果。

综上所述, 针对不同类型的杆号牌, 算法的检出率和识别准确率可能会有所不同, 这可能与杆号牌的外观特征、拍摄角度和距离等因素有关。为了进一步提高算法的识别精度和鲁棒性, 未来可以针对不同类型的杆号牌进行深入的分析和优化。

### 3 结束语

本文提出一种可适应我国高铁主要干线接触网设备的接触网支柱定位和号牌识别智能算法。针对接触网支柱区域和杆号牌分别进行区域检测, 然后通过裁剪支柱图像并进行区域相关性分析定位的策略, 可自适应判断接触网支柱结构类型并定位号牌区域。设计的基于 SVTR-tiny 网络的识别模型可适应接触网杆号文本中英文字符混合、字体多样、文本长度变化等特点。经实际应用测试验证, 本文提出识别方法具有较高的识别精度和泛化性, 可满足我国主要铁路干线接触网支柱识别需求。该方法可提高接触网支柱数字化管理效率和精度, 从而为铁路供电基础设施的养护维修和保障列车行车安全提供了有力支持。

#### 参考文献:

[1] 周 雯. 面向高速铁路运行安全的智能图像识别方法研究

[D]. 北京: 中国铁道科学研究院, 2020.

- [2] 刘尚昆, 朱广平, 王 波, 等. 基于激光雷达的铁路接触网支柱识别技术 [J]. 铁道建筑, 2021, 61 (10): 133-135.
- [3] 武翔宇. 基于图像的接触网支柱检测与编号识别 [D]. 成都: 西南交通大学, 2015.
- [4] 陈忠革. 接触网安全巡检系统中的支柱号牌识别算法 [J]. 电气化铁道, 2016 (6): 34-38.
- [5] 杨文静, 狄 岚, 梁久祯. 基于多帧杆号字符识别的接触网检测系统 [J]. 数据采集与处理, 2016, 31 (6): 1250-1258.
- [6] 杨 梅, 谭泽富, 蔡 黎. 基于自适应形状上下文特征的实时杆号识别 [J]. 计算机工程与应用, 2018, 54 (12): 187-191, 197.
- [7] 闵 锋, 吴 涛, 陈智羽. 接触网支柱号牌定位与字符分割方法 [J]. 计算机工程与设计, 2020, 41 (3): 789-794.
- [8] 郑 巧, 唐 鹏, 韦 璞. 基于自适应投影分割的接触网号牌字符提取 [J]. 信息技术, 2016 (5): 58-62.
- [9] LECUN Y, BENGIO Y, HINTON G. Deep learning [J]. Nature, 2015, 521 (7553): 436-444.
- [10] 徐可佳, 张军磊, 汪正鑫. 基于深度学习的接触网支柱号在线识别算法研究 [J]. 电气化铁道, 2017, 28 (Z1): 140-145, 150.
- [11] 侯明斌, 于 龙, 高仕斌. 高铁接触网支柱号的图像序列识别方法 [J]. 电气化铁道, 2020, 31 (6): 29-35.
- [12] 杨 栋, 李 超, 吴兴华, 等. 基于智能识别技术的铁路安检辅助分析装置研究 [J]. 计算机测量与控制, 2022, 30 (8): 25-30.
- [13] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. Scaled-YOLOv4: scaling cross stage partial network [C] // Computer Vision and Pattern Recognition, 2021: 13029-13038.
- [14] WANG C Y, LIAO H Y M, WU Y H, et al. CSPNet: A new backbone that can enhance learning capability of CNN [C] // Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops, 2020: 390-391.
- [15] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8759-8768.
- [16] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2117-2125.
- [17] DU Y, CHEN Z, JIA C, et al. SVTR: scene text recognition with a single visual model [C] // European Conference on Artificial Intelligence, 2022: 884-890.
- [18] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C] // Proceedings of the Advances in Neural Information Processing Systems, 2017: 6000-6010.
- [19] KAI H, YUNHE W, HANTING C, et al. A survey on vision transformer [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45 (1): 87-110.
- [20] KINGMAD, BA J. Adam: a method for stochastic optimization [C] // International Conference on Learning Representations, 2014, abs/1412.6980.