

DOI: 10.11992/tis.201905026

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.TP.20190612.1619.002.html>

# 基于改进的 Faster R-CNN 高压线缆目标检测方法

刘召<sup>1</sup>, 张黎明<sup>2</sup>, 耿美晓<sup>1</sup>, 么军<sup>2</sup>, 张金禄<sup>2</sup>, 胡益菲<sup>2</sup>

(1. 清研同创机器人(天津)有限公司, 天津 300300; 2. 国网天津市电力公司, 天津 300010)

**摘要:** 利用带电作业机器人取代人类的手工作业, 可以有效地减少高电压、强电场对人体的危害, 大大提高作业的效率。为解决带电作业机器人在复杂背景环境中对线缆目标的智能检测问题, 提出基于改进的 Faster R-CNN 高压线缆目标检测方法。为了提高网络提取图像高级特征的能力, 引入跳转连接并调整激活层、卷积层的顺序; 然后对候选框生成机制进行改进, 提升网络对小目标检测的性能; 最后利用 ROI 池化层提取每个区域的特征, 同时完成分类和框回归任务。通过构建高压线缆图像数据集, 基于改进的 Faster R-CNN 模型进行大量实验, 最后取得了较好的精度和较快的速度。

**关键词:** 目标检测; 深度学习; 高压线缆; 复杂背景; 小目标; 带电作业; Faster R-CNN; 区域候选

**中图分类号:** TP18; TP391 **文献标志码:** A **文章编号:** 1673-4785(2019)04-0627-08

中文引用格式: 刘召, 张黎明, 耿美晓, 等. 基于改进的 Faster R-CNN 高压线缆目标检测方法 [J]. 智能系统学报, 2019, 14(4): 627-634.

英文引用格式: LIU Zhao, ZHANG Liming, GENG Meixiao, et al. Object detection of high-voltage cable based on improved Faster R-CNN[J]. CAAI transactions on intelligent systems, 2019, 14(4): 627-634.

## Object detection of high-voltage cable based on improved Faster R-CNN

LIU Zhao<sup>1</sup>, ZHANG Liming<sup>2</sup>, GENG Meixiao<sup>1</sup>, YAO Jun<sup>2</sup>, ZHANG Jinlu<sup>2</sup>, HU Yifei<sup>2</sup>

(1. Tsinghua Tongchuang Robot Co., Ltd, Tianjin 300300, China; 2. State Grid Tianjin Electric Power Company, Tianjin 300010, China)

**Abstract:** The use of live working robots to replace human manual operation can effectively reduce the harm of a high-voltage and strong electric field to the human body and considerably improve the working efficiency. To solve the intelligent high-voltage cable object detection problem for live working robots under a complicated background environment, a high-voltage cable object detection method based on the improved Faster R-CNN is proposed. To improve the capability of extracting the high-level features of images in the network, skip connections are introduced and the order of the activation and convolution layers is adjusted. Then, the proposal bounding box generation mechanism is improved to enhance the performance of the proposed method for small object detection. Finally, the features of each region are extracted using the ROI pooling layers, and the classification and bounding box regression tasks are accomplished at the same time. Through the construction of high-voltage cable image datasets and the performance of numerous experiments based on the improved Faster R-CNN model, good accuracy and fast speed have been achieved.

**Keywords:** object detection; deep learning; high-voltage cable; complicated background; small object; live working; Faster R-CNN; region proposal

随着我国电力行业的升级改造, 更加智能的带电作业机器人成为当前的研究热点。2009 年我国进行了高压带电作业机器人产品化样机的验

证<sup>[1]</sup>, 但现有高压带电作业机器人系统一般采用主从控制方式, 依然是人工手动控制<sup>[2]</sup>, 需要借助人眼判断目标的位置来进行操作。虽然在一定程度上提高了工作效率保护了作业人员, 但在高空高压环境下依然存在巨大的安全隐患, 造成无法挽回的安全和财产损失。因此, 利用当前性能优

收稿日期: 2019-05-14. 网络出版日期: 2019-06-13.

基金项目: 天津市智能制造科技重大专项 (17ZXZNGX00120).

通信作者: 耿美晓. E-mail: [lin@thtrobot.com](mailto:lin@thtrobot.com).

良的深度学习方法代替人眼对线缆目标进行自动识别检测,研制更智能的带电作业机器人成为亟待解决的问题。

传统的目标检测方法以图像识别为基础,主要包括4个步骤:提取候选框、对每个框提取特征、图像特征分类、非极大值抑制完成框回归<sup>[3-4]</sup>。一般对图像使用穷举法选出物体可能出现的区域框,对这些区域框提取特征并使用图像识别方法分类,得到所有分类成功的区域后,通过非极大值抑制输出结果。传统方法中最成功的是基于多尺度形变部件模型(deformable part model, DPM)<sup>[5]</sup>,它利用部件间的关系来描述物体,但该方法相对复杂,检测速度也较慢。所以,人工提取的特征不能适应当今的海量数据,难以应用于背景复杂的高压线缆图像数据中,且不同场景的图像往往要重新设计合适的特征,实际应用时也要依据具体情况而定,这些限制直接影响着最终的检测效果及效率。

2006年人工智能专家Hinton提出深度学习(deep learning, DL)概念<sup>[6]</sup>,开启了人工智能领域研究深度学习的热潮。深度学习具有很强的特征学习能力,它采用逐层训练的方法缓解了传统神经网络算法在训练多层神经网络时出现的局部最优问题<sup>[7-8]</sup>。基于这些特征,深度学习在图像识别、语音识别、自然语言处理、工业过程控制等方面已显示出独特的优势。

近两年深度学习技术在目标检测领域也取得了长足的进展。与传统依赖手工设计的特征提取算法不同,深度卷积神经网络对几何变换、形变、光照具有一定程度的不变性,可以有效克服复杂场景下目标检测困难,并且可在训练数据驱动下自适应地构建特征描述,具有更高的灵活性和泛化能力。基于深度学习的目标检测主要分为基于区域建议方法和无区域建议的方法<sup>[3,9]</sup>。基于区域建议的方法是以R-CNN(Region-based convolutional neural networks)<sup>[10]</sup>为基础,在PASCAL VOC2012数据集上的检测平均精度mAP达到53.3%,较传统方法有很大的提升。此后研究者在此基础上提出了很多改进模型,如SPP-net(Spatial pyramid pooling networks)<sup>[11]</sup>、Fast R-CNN<sup>[12]</sup>等。而无区域建议的目标检测方法主要是YOLO(You only look once)<sup>[13]</sup>及其改进模型。基于区域建议的方法在精度上更具优势,而无区域建议的方法检测速度更快。基于深度学习的目标检测方法在低空目标、人体行为检测、无人机图像等领域已取得较好应用<sup>[14-16]</sup>。

本文针对带电作业机器人在复杂背景环境中线缆目标的智能检测问题,首次将Faster R-CNN

模型引入到高压线缆及相关物体的目标检测任务中,根据高压线缆的特点对原始Faster R-CNN进行了改进,引入跳转连接使提取的特征更具表达能力,增加锚框数量使模型对不同尺寸目标都有较好的检测性能,最终实现了高空高压环境下各种线缆的目标检测,在精度上取得了目前最好的检测效果,并且检测速度更快,为带电作业机器人后续的操作提供了可靠的视觉支持。

## 1 基于深度学习的高压线缆目标检测

根据高压线缆目标检测的实际需求,本文以Faster R-CNN<sup>[17]</sup>模型为基础并提出了有效的改进策略,最终设计的模型原理框架如图1所示。它主要分为4个主要的步骤来实现:1)将整幅高压线缆图像输入深度卷积神经网络生成特征图谱;2)将特征图谱输入给区域候选网络(region proposal network, RPN)生成目标候选区域,避免了使用选择性搜索耗时长的问題,同时可以获得更加准确的定位结果;3)采用感兴趣区域池化层(region of interest, ROI)提取候选区域的特征,它可以将输入的任意尺寸的特征图谱转化成统一维度的输出向量,从而解决了卷积神经网络输入必须是固定尺寸的问题;4)利用一个多任务分类器做候选区域的位置回归,目标类别的判定采用深度学习算法精确检测出定位器的边界框位置。其中本文的改进首先在特征提取部分,引入跳转连接并调整激活层和卷积层的顺序,提高了网络提取图像高级特征的能力;其次在目标候选区域生成阶段增多锚框的种类和数量,提升对小目标检测的性能。

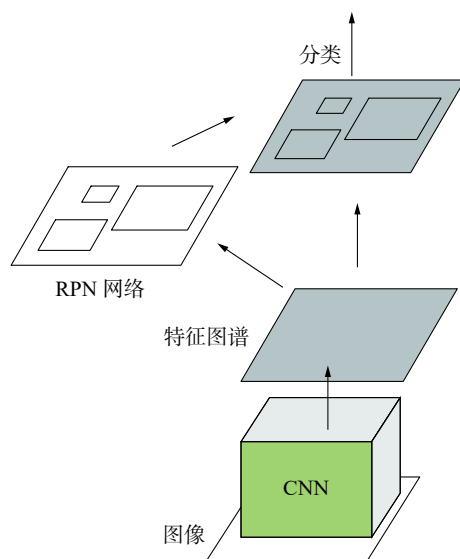


图1 基于Faster R-CNN高压线缆目标检测框架  
Fig. 1 Object detection framework of high voltage cable based on Faster R-CNN

### 1.1 卷积神经网络

卷积神经网络 (convolutional neural networks, CNN) 是一类包含卷积计算且具有深度结构的前馈神经网络, 是深度学习中最具代表性的网络结构之一。传统的 CNN 网络一般包括卷积层、池化层和全连接层, 随着网络层数的加深逐层提取图像的特征。一个典型是 CNN 结构如图 2 所示。

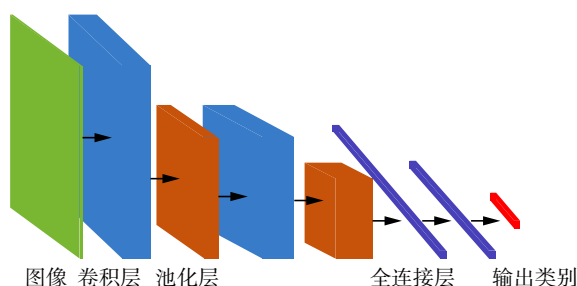


图2 CNN的一般结构框架

Fig. 2 General structure of CNN

### 1.2 区域候选网络

一般的目标检测方法都是使用选择性搜索来获取候选区域, 但这种方法不仅耗时且准确率不高, 因此考虑用一个卷积神经网络来进行候选区域的提取, 这也就是RPN网络的核心思想。

RPN的网络结构如图3所示。在RPN网络之前会用一个卷积神经网络对需要检测的图像进行特征提取, 这个前置的CNN提取的特征维度为 $51 \times 39 \times 256$ 。对该特征图谱再进行一次卷积操作, 保持维度不变, 同样得到一个 $51 \times 39 \times 256$ 维度的特征图谱。为了便于下文叙述, 先定义一个“锚”的概念: 对于这个维度为 $51 \times 39 \times 256$ 的卷积特征图谱, 认为其一共有 $51 \times 39$ 个“锚”。让该卷积特征图谱的每个“锚”都对原图中对应位置上9种尺寸的候选框进行检测, 检测的目标是判断候选框中是否包含一种物体, 因此共有 $51 \times 39 \times 9$ 个候选检测框。这每个“锚”对应的9个候选框的面积分别为 $128^2$ 、 $256^2$ 、 $512^2$ 。每种面积又分为3种长宽比, 分别为 $2:1$ 、 $1:2$ 、 $1:1$ 。

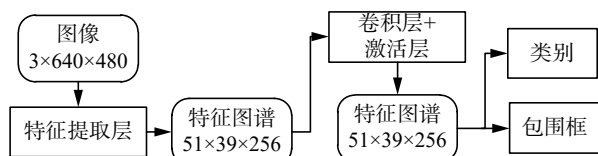


图3 RPN网络结构

Fig. 3 RPN networks structure

对于这 $51 \times 39$ 个“锚”和 $51 \times 39 \times 9$ 个检测框, 每个“锚”的计算步骤如图4所示。其中,  $k$ 为单个“锚”对应的候选框的个数, 此时 $k=9$ 。使用一个 $3 \times 3$ 的窗口在特征图谱上滑动, 将每个“锚”转

换为一个统一的256维的特征向量, 这个特征对应着两部分的输出。一部分表示该“锚”的候选框中包含物体的概率, 每个框对应着两个输出: 是物体的概率和不是物体的概率, 所以其总的输出长度为 $2k$ 。另一部分表示框回归, 每个框对应4个框回归参数, 因此该部分总输出的长度为 $4k$ 。

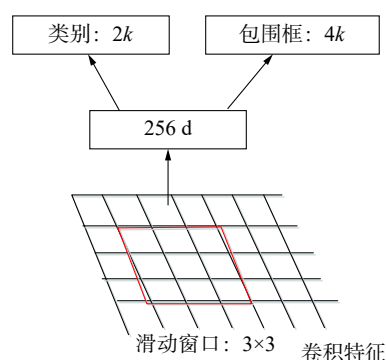


图4 “锚”与网络输出的对应关系

Fig. 4 Corresponding relation between frame and network output

### 1.3 Fast R-CNN 目标检测结构

Fast R-CNN网络将特征提取和区域分类两个步骤融合在一个卷积神经网络中, 不同于之前的方法是提取完每个区域的特征之后, 再以传统的SVM作为分类器。这样Fast R-CNN就可以同时进行特征提取网络和分类网络的训练, 从而获得更高的准确度, 其网络结构如图5。

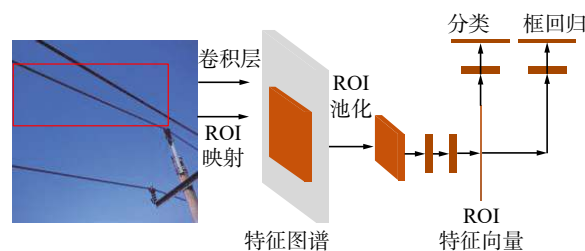


图5 Fast R-CNN网络结构

Fig. 5 Fast R-CNN networks structure

对于待检测图像中的候选区域, 将它映射到前置CNN所提取的卷积特征图谱中对应的区域, 即图5中的感兴趣区域映射, 这样操作是因为卷积特征图谱实际上和原始图像在位置上存在一定的对应关系, 即待检测图像中的候选区域是对应在卷积特征图谱中相同的位置。然后使用ROI池化层对该区域再进行特征提取, 如图6所示。ROI池化层对于不同尺寸大小的输入图像都能得到一个固定维度的输出向量。一般的卷积神经网络的输入需要固定尺寸图像, 而候选区域的尺寸是大小不一的。如果先将不同尺寸的区域



缩放到统一大小再输入卷积神经网络提取特征,会降低后续的分类准确率,因此使用 ROI 池化层避免这个问题。

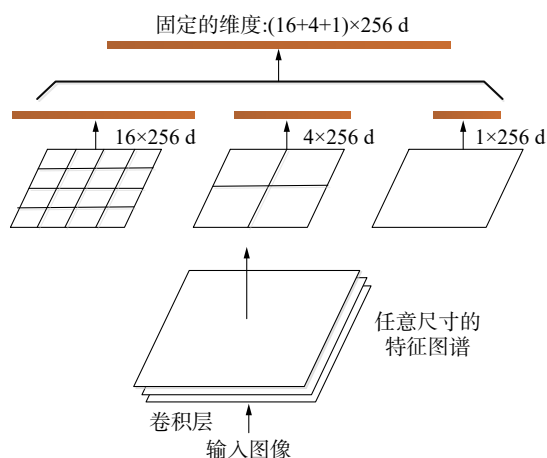


图6 ROI 池化层

Fig. 6 ROI pooling layer

为了具体阐述 ROI 池化层的作用,假设前置 CNN 输出特征图谱的宽度为  $w$ , 高度为  $h$ , 通道为  $c$ , 其中  $c$  是一个常数。首先, ROI 池化层把特征图谱划分为  $4 \times 4$  的区域网格, 这样每个网格的维度变成: 宽为  $w/4$ 、高为  $h/4$ 、通道为  $c$ 。对每个网格中的每个通道都取最大值作为该网格的输出, 这样  $4 \times 4$  的网格最终形成了一个  $16c$  维的特征向量。然后, 把特征图谱划分成  $2 \times 2$  和  $1 \times 1$  的网格, 用同样的方法提取每个网格的输出特征, 提取的特征的维度分别为  $4c$  和  $c$ 。最后, 将得到的所有特征向量拼接起来, 就得到了  $16c + 4c + c = 21c$  维的特征向量。显而易见, 这个输出特征的维度与输入特征图谱的  $w$ 、 $h$  均无关。因此 ROI 池化层可以把输入的任何宽度、高度的卷积特征图谱转换为固定维度的向量。

得到 ROI 池化层的特征向量后, 分成两个分支分别送到全连接层进行分类和框回归。对于分类, 如果待检测图像中有  $K$  类物体需要检测, 那么最终的输出应该是  $K+1$  个数, 因为多包含了一类“背景类”, 输出中的每个数值都代表该区域为某类物体的类别概率。对于框回归, 其要做的是对 RPN 输出的候选检测框进行某种程度的细化。因为 RPN 网络获得的候选框有时存在一定偏差。认为通过 RPN 网络得到的候选框的 4 个参数为  $(x, y, w, h)$ , 其中,  $(x, y)$  表示候选框左上顶点的坐标;  $(w, h)$  表示候选框的宽和高。而正确的目标框位置用  $(x', y', w', h')$  表示, 那么框回归的作用就是学习参数  $\left(\frac{x'-x}{w}, \frac{y'-y}{h}, \ln \frac{w'}{w}, \ln \frac{h'}{h}\right)$ 。

## 1.4 改进 Faster R-CNN 模型的高压线缆目标检测

### 1.4.1 特征提取部分的改进

传统的卷积神经网络大多是由卷积层、池化层和全连接层组成深层网络, 对于图像分类等计算机视觉任务, 不断加深网络层数一般可以提高性能, 但当网络加深到一定程度时, 会导致梯度弥散或梯度消失, 网络变得难以训练, 同时导致提取的特征的语义属性减弱, 对类别的判别能力变差<sup>[18-19]</sup>。为了解决这个问题, 引入跳转连接, 将卷积层的输入特征图谱加到输出部分, 使网络具有能有效进行反传的梯度信息。同时增加  $1 \times 1$  的卷积操作, 在减少参数的同时使网络具有更强的特征提取能力, 多增加的激活层也使网络具有更好的非线性表达能力。此外, 调整模块中各操作的顺序为批标准化、ReLU, 再进行卷积操作, 这种完全预激活的方式可以起到更好正则化的作用, 同时使梯度信息可以无障碍传递。本文改进的特征提取模块如图 7。

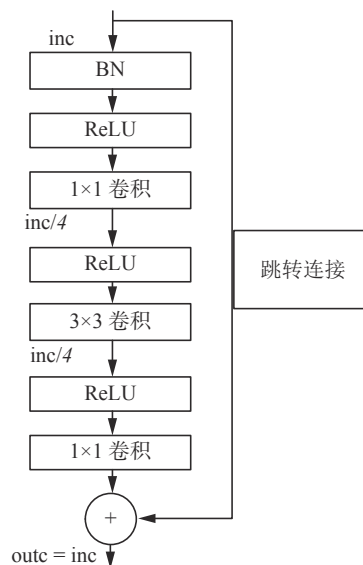


图7 改进的深层网络特征提取模块

Fig. 7 Improved deep network feature extraction module

改进的特征提取模块降低参数的同时, 也提高了网络在推理时的速度, 节省了测试时间, 更有利于高压线缆目标检测实际应用的需求。当使用图 7 的模块时, 参数量为  $17/16 inc^2$ , 参数量的计算公式如式 (1), 其中输入的通道数是  $inc$ , 输出的通道数是  $outc$ 。使用传统的两层卷积核大小为  $3 \times 3$  的卷积层时, 参数量为  $18/16 inc^2$ , 参数量减少了  $1/16$ , 计算复杂度减少了  $9/16$ 。

$$parameters = inc(k_h k_w) outc \quad (1)$$

### 1.4.2 区域候选框机制的改进

区域候选网络中候选框的尺寸和长宽比是非常重要的超参数, 它直接影响生成候选框的数

量,也影响着网络对不同尺寸目标的识别能力。原始 Faster R-CNN 针对特征图谱上的每个“锚”,由3种长宽比和3种尺寸生成9个候选框,然后对获得的原始图片候选区域使用非极大值抑制算法删除冗余的候选区域,其中最小的尺寸为 $128 \times 128$ 。但是在高压线缆识别中存在着连接点等尺寸较小的目标,为了提升网络对小目标的检测性能,本文增加了一组 $64 \times 64$ 大小的候选框,其长宽比依然有3种。在训练过程中,RPN部分使用的候选框尺寸则由9种变为12种,分别为 $64 \times 64$ 、 $128 \times 128$ 、 $256 \times 256$ 、 $512 \times 512$ ,3个长宽比分别为2:1、1:2和1:1。

在整个模型结构中,多任务分类器的损失函数定义为

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{\text{cls}}} \sum L_{\text{cls}}(p_i, p_i^*) + \lambda \frac{1}{N_{\text{reg}}} \sum p_i^* L_{\text{reg}}(t_i, t_i^*) \quad (2)$$

式中: $p_i$ 、 $t_i$ 是目标包围框和类别的预测值; $p_i^*$ 、 $t_i^*$ 是目标包围框和类别的实际值; $N_{\text{cls}}$ 、 $N_{\text{reg}}$ 分别为分类和回归的归一化参数; $\lambda$ 是平衡分类与回归的系数。回归损失函数和分类损失函数如公式(3)和式(4)所示:

$$L_{\text{reg}} = R(t_i - t_i^*) \quad (3)$$

$$L_{\text{cls}}(p_i, p_i^*) = -\log(p_i^* p_i + (1 - p_i^*)(1 - p_i)) \quad (4)$$

式(3)中 $R$ 是Fast R-CNN中定义的鲁棒的损失函数,如式(5)所示:

$$R(x) = \begin{cases} 0.5x^2, & |x| < 1 \\ |x| - 0.5, & \text{其他} \end{cases} \quad (5)$$

## 2 仿真实验与结果分析

### 2.1 高压线缆图像数据集

为了验证本文提出的基于Faster R-CNN的高压线缆目标检测方法的有效性和先进性,首先本文生成了具有大量高压线缆图片的数据集,并将数据处理成VOC2007数据形式,主要包括采集高压线缆图像数据和标注需检测目标两个部分。

#### 2.1.1 采集高压线缆图像数据

本文使用高清摄像机采集了不同场景、天气环境下的高压线缆图像,并对其中感兴趣的目标进行了人工标签。天气环境主要分为晴天、阴天和强光3种情况,之所以将强光单独作为一类是因为强光会导致拍摄的线缆图像出现断裂的情况。高压线缆图像中的杆塔结构较复杂,有单回路结构和多回路结构,上有瓷柱,有多根线缆,线缆有时会出现重叠现象;图像背景一般都较为复杂,主要包含房屋、树木等。

#### 2.1.2 标注检测目标

对采集到的线缆图像中需要检测的目标进行了人工标签,使用LabelImg标注工具。图像中需要检测的目标有缆线与瓷柱的连接处(连接点,corner)、向右倾斜的线缆(右倾线,rline)、向左倾斜的线(左倾线,lline)和带线夹的线(线夹线,clampline)。表1中所列的数据集Data<sub>1</sub>和Data<sub>4</sub>的场景近似,Data<sub>1</sub>中的连接点只有一种情况,即横向连接;而数据集Data<sub>4</sub>中连接点有两种情况,包括横向连接和纵向连接,如图8(a)和图8(d)所示,图8(d)展示了Data<sub>4</sub>中的两种连接点。

表1 不同场景的数据集  
Table 1 Test sets of five scenarios

数据集	总数	训练	测试	类别	场景特点
Data <sub>1</sub>	1 660	1 494	166	3	晴天、重叠多
Data <sub>2</sub>	2 937	2 643	294	4	阴天、背景复杂
Data <sub>3</sub>	2 557	2 301	256	3	强光、背景复杂
Data <sub>4</sub>	2 056	1 850	206	3	晴天、重叠多
Data <sub>5</sub>	1 790	1 611	179	3	晴天、背景复杂

### 2.2 训练设置及评价指标

所有场景的数据均训练迭代10万次,采用衰减学习率,前5万代为0.001,后5万代为0.000 1。为了对比本文所提方法的有效性,对比方法为原始的Faster R-CNN。高压线缆图像尺寸为 $640 \times 480$ 。

评价指标选择目标检测中常用的平均准确率(average precision, AP),其值为P-R曲线围成的面积,其中 $P$ 表示准确率、 $R$ 表示召回率。 $P$ 、 $R$ 的计算方法如式(6)和式(7)所示:

$$P = \text{TP}/(\text{TP} + \text{FP}) \quad (6)$$

$$R = \text{TP}/(\text{TP} + \text{FN}) \quad (7)$$

式中:TP是被正确划分为正例的个数,FP是被错误划分为正例的个数,FN是被错误划分为负例个数。mAP(mean average precision)即是所有类别的平均准确率的平均值。

实验环境为Ubuntu16.04系统,Intel Core i7-8700K和英伟达GTX 1080Ti显卡,使用TensorFlow框架进行提出算法的训练和测试。

### 2.3 实验结果及分析

使用5种场景数据的训练集分别训练原始Faster R-CNN模型和提出的改进Faster R-CNN模型,然后在测试集上进行测试,得到各类别的平均准确率。原始Faster R-CNN的测试结果如表2所示,改进的Faster R-CNN的测试结果如表3所示。

表2 原始 Faster R-CNN 在 5 种场景下测试结果

Table 2 The test results average of original Faster R-CNN in the five scenarios

数据集	连接点	左倾线	右倾线	线夹线	mAP
Data <sub>1</sub>	0.881	0.898	0.813		0.864
Data <sub>2</sub>	0.907	0.901	0.907	0.907	0.906
Data <sub>3</sub>	0.898	0.899	0.907		0.901
Data <sub>4</sub>	0.890	0.818	0.813		0.840
Data <sub>5</sub>	0.907	0.903	0.908		0.906

表3 改进 Faster R-CNN 在 5 种场景下测试结果

Table 3 The test results average of improved Faster R-CNN in the five scenarios

数据集	连接点	左倾线	右倾线	线夹线	mAP
Data <sub>1</sub>	0.939	0.939	0.895		0.924
Data <sub>2</sub>	0.928	0.905	0.944	0.961	0.935
Data <sub>3</sub>	0.962	0.882	0.921		0.922
Data <sub>4</sub>	0.982	0.829	0.947		0.920
Data <sub>5</sub>	0.992	0.883	0.970		0.948

从测试集的实验结果可以看出,在不同天气条件环境下,原始的 Faster R-CNN 模型对高压线缆图像中的各类别目标检测的平均准确率均在 80% 以上,而 mAP 值均在 84% 以上。说明 Faster R-CNN 模型可以解决高压线缆目标检测任务。而改进的 Faster R-CNN 模型对高压线缆目标具有更好的检测性能, mAP 值均在 92% 以上,相较原始的 Faster R-CNN 模型有明显提升,说明提出改进方法的有效性,且在阴天、强光和晴天背景复杂的场景下都可以较好地完成高压线缆目标检测任务。如图 8 是 5 种场景下,高压线缆目标检测结果的样例。每种场景展示了两张图像,图 8(a)~(e) 依次对应表 1 中的 5 种场景。

Data<sub>1</sub>、Data<sub>4</sub> 和 Data<sub>5</sub> 数据集均为晴天下的场景中,且出现线缆重叠的情况较多,对比表 2 和表 3 中 Data<sub>1</sub>、Data<sub>4</sub> 两个数据集的测试结果可以看出,改进的方法对线缆的检测有较大的提升, mAP 值分别提升了 6% 和 8%,但在 Data<sub>5</sub> 数据集上,改进算法对左倾线的检测 AP 值有些小的降低。说明改进的 Faster R-CNN 模型可以较好解决目标遮挡的检测问题,但个别较严重的重叠情况可能无法检测。结合带电作业机器人的实际操作看,当发现重叠的线缆较多时,可以调节摄像头的方位使拍摄的图像中重叠的线缆尽量少,从而提升对线缆目标检测准确率。而且在实际应用中带电作业机器人每次总是处理单根线缆,所以重叠问题对实际操作的影响并不是很大。

在实际操作中更常见的场景是背景复杂的时候,如图 8 中 Data<sub>2</sub> 和 Data<sub>5</sub> 所示的场景,背景中会有大面积的房屋、树木等干扰识别的目标。但在这两个场景下,改进的 Faster R-CNN 模型的 mAP 均值 93% 以上。说明改进的 Faster R-CNN 模型在进行高压线缆目标检测时,可以较好地克服背景中包含大面积的房屋、树木等复杂场景。这一点在实际中的应用更有价值,提高了系统的适用范围,同时也验证了改进方法在进行高压线缆目标检测时的有效性。

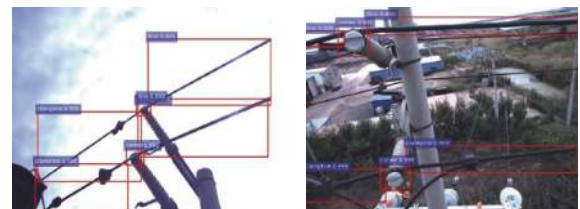
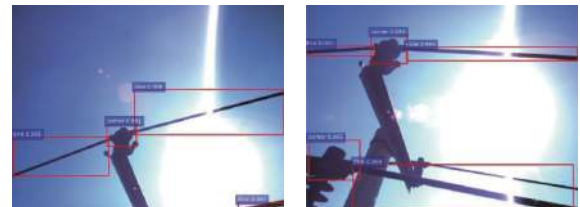
(a) Data<sub>1</sub> 晴天的目标检测结果(b) Data<sub>2</sub> 阴天的目标检测结果(c) Data<sub>3</sub> 强光的目标检测结果(d) Data<sub>4</sub> 晴天的目标检测结果(e) Data<sub>5</sub> 晴天的目标检测结果

图 8 5 种场景下测试集的目标检测结果

Fig. 8 Detection results of test sets in five scenarios

改进的 Faster R-CNN 模型进行高压线缆目标检测时的速度与使用的硬件环境密切相关,当使用高性能的 GPU 如 1080Ti 测试时,15 f/s 的处理



速度,相较于未改进算法的 12 f/s 处理速度有提升,完全能够满足带电作业机器人的实时性需求。

### 3 结束语

针对带电作业机器人需要自动识别高空高压线缆的实际课题需求,本文结合当前在视觉领域最有效的深度学习方法,实现了一种自动识别线缆目标的方法。首先引入了 Faster R-CNN 模型进行特征提取,针对高压线缆图像的特点提出了两个方面的改进:在特征提取时引入跳转连接并调整卷积层、激活层的顺序,提升网络对高层语义特征的提取能力,改进候选框生成策略提升网络对连接点等小目标的检测能力。本文改进的 Faster R-CNN 模型展现出了针对不同天气环境、较复杂背景下的鲁棒、实时、高精度检测的优良性能,可以满足带电作业机器人视觉智能识别的技术指标。本文将深度学习方法应用于带电作业机器人的视觉识别任务中,取得了较好的精度和实时的性能,具有很高的工程应用价值。

### 参考文献:

- [1] 赵玉良, 戚晖, 陈凡明, 等. 高压带电作业机器人专用遥控剥皮器的研制 [J]. 微计算机信息, 2010, 26(32): 146-147, 119.  
ZHAO Yuliang, QI Hui, CHEN Fanming, et al. Design on the remote controlled electric-driving remover for live working robot[J]. Microcomputer information, 2010, 26(32): 146-147, 119.
- [2] 王振利, 鲁守银, 李健, 等. 高压带电作业机器人视觉伺服系统 [J]. 制造业自动化, 2013, 35(7): 69-72.  
WANG Zhenli, LU Shouyin, LI Jian, et al. Vision servo system for high-voltage live working robot[J]. Manufacturing automation, 2013, 35(7): 69-72.
- [3] 于进勇, 丁鹏程, 王超. 卷积神经网络在目标检测中的应用综述 [J]. 计算机科学, 2018, 45(S2): 17-26.  
YU Jinyong, DING Pengcheng, WANG Chao. Overview: application of convolution neural network in object detection[J]. Computer science, 2018, 45(S2): 17-26.
- [4] CAI Zhaowei, VASCONCELOS N. Cascade R-CNN: delving into high quality object detection[C]//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA, 2018: 6154-6162.
- [5] 宋焕生, 张向清, 郑宝峰, 等. 基于深度学习方法复杂场景下车辆目标检测 [J]. 计算机应用研究, 2018, 35(4): 1270-1273.  
SONG Huansheng, ZHANG Xiangqing, ZHENG Baofeng, et al. Vehicle detection based on deep learning in complex scene[J]. Application research of computers, 2018, 35(4): 1270-1273.
- [6] HINTON G E, OSINDERO S, TEH Y W. A fast learning algorithm for deep belief nets[J]. Neural computation, 2006, 18(7): 1527-1554.
- [7] 毕晓君, 冯雪赟. 基于改进深度学习模型 C-GRBM 的人体行为识别 [J]. 哈尔滨工程大学学报, 2018, 39(1): 156-162.  
BI Xiaojun, FENG Xueyun. Human action recognition based on improved depth learning model C-GRBM[J]. Journal of Harbin Engineering University, 2018, 39(1): 156-162.
- [8] 龙慧, 朱定局, 田娟. 深度学习在智能机器人中的应用研究综述 [J]. 计算机科学, 2018, 45(S2): 43-47, 52.  
LONG Hui, ZHU Dingju, TIAN Juan. Research on deep learning used in intelligent robots[J]. Computer science, 2018, 45(S2): 43-47, 52.
- [9] 张慧, 王坤峰, 王飞跃. 深度学习在目标视觉检测中的应用进展与展望 [J]. 自动化学报, 2017, 43(8): 1289-1305.  
ZHANG Hui, WANG Kunfeng, WANG Feiyue. Advances and perspectives on applications of deep learning in visual object detection[J]. Acta automatica sinica, 2017, 43(8): 1289-1305.
- [10] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA, 2014: 580-587.
- [11] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE transactions on pattern analysis and machine intelligence, 2015, 37(9): 1904-1916.
- [12] GIRSHICK R. Fast R-CNN[C]//Proceedings of 2015 IEEE International Conference on Computer Vision. Santiago, Chile: IEEE, 2015: 1440-1448.
- [13] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA, 2016: 779-788.
- [14] 莫宏伟, 汪海波. 基于 Faster R-CNN 的人体行为检测研究 [J]. 智能系统学报, 2018, 13(6): 967-973.  
MO Hongwei, WANG Haibo. Research on human behavior detection based on Faster R-CNN[J]. CAAI transactions on intelligent systems, 2018, 13(6): 967-973.
- [15] 曹宇剑, 徐国明, 史国川. 基于旋转不变 Faster R-CNN 的低空装甲目标检测 [J]. 激光与光电子学进展,

2018, 55(10): 101501.

CAO Yujian, XU Guoming, SHI Guochuan. Low altitude armored target detection based on rotation invariant faster R-CNN[J]. Laser and optoelectronics progress, 2018, 55(10): 101501.

- [16] 魏湧明, 全吉成, 侯宇青阳. 基于 YOLO v2 的无人机航拍图像定位研究 [J]. 激光与光电子学进展, 2017, 54(11): 111002.

WEI Yongming, QUAN Jicheng, HOU Yuqingyang. Aerial image location of unmanned aerial vehicle based on YOLO v2[J]. Laser and optoelectronics progress, 2017, 54(11): 111002.

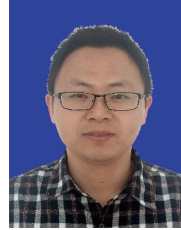
- [17] REN Shaoqing, HE Kaiming, GIRSHICK R, SUN J. Faster R-CNN: towards real-time object detection with region proposal networks[J]. [IEEE transactions on pattern analysis and machine intelligence](#), 2017, 39(6): 1137–1149.

- [18] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA, 2016: 770–778.

- [19] XIE Saining, GIRSHICK R, DOLLÁR P, et al. Aggregated residual transformations for deep neural networks

[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA, 2017: 5987–5995.

#### 作者简介:



刘召, 男, 1979 年生, 博士, 主要研究方向为机器人及其自动化装备。作为项目负责人和技术骨干完成研究课题 10 余项, 其中, 863 项目 1 项, 985 项目 1 项, 国防科技课题 4 项, 日本学术振兴会科学研究补助金基础研究项目 1 项。发表学术论文 10 余篇。



张黎明, 男, 1969 年生, 高级技师, 主要研究方向为智能配电网。个人获 24 项国家专利, 带领团队为公司实现技术革新 400 多项, 获国家专利 158 项, 20 余项填补智能电网建设空白。



耿美晓, 女, 1990 年生, 硕士, 主要研究方向为图像处理、机器人学习与人工智能。