

引用格式:王万国,田兵,刘越,等.基于RCNN的无人机巡检图像电力小部件识别研究[J].地球信息科学学报,2017,19(2):256-263. [Wang W G, Tian B, Liu Y, et al. 2017. Study on the electrical devices detection in UAV images based on region based convolutional neural networks. Journal of Geo-information Science, 19(2):256-263. ] DOI:10.3724/SP.J.1047.2017.00256

# 基于RCNN的无人机巡检图像电力小部件识别研究

王万国<sup>1,2</sup>, 田兵<sup>3</sup>, 刘越<sup>1,2</sup>, 刘良<sup>1,2</sup>, 李建祥<sup>1,2</sup>

1. 国网山东省电力公司电力科学研究院 国家电网公司电力机器人技术实验室, 济南 250002;  
2. 山东鲁能智能技术有限公司, 济南 250101; 3. 国网山东省电力公司, 济南 250000

## Study on the Electrical Devices Detection in UAV Images based on Region Based Convolutional Neural Networks

WANG Wanguo<sup>1,2\*</sup>, TIAN Bing<sup>3</sup>, LIU Yue<sup>1,2</sup>, LIU Liang<sup>1,2</sup> and LI Jianxiang<sup>1,2</sup>

1. Electric Power Robotics Laboratory of SGCC, Shandong Electric Power Research Institute, Jinan, 250002, China; 2. Shandong Luneng Intelligence Technology Co., Ltd. Jinan, 250101, China; 3. State Grid Shandong Electric Power Company, Jinan, 250000, China

**Abstract:** With the wide application of Unmanned Aerial Vehicle (UAV) in the inspection of power transmission line, the demand for objects detection and data mining from images acquired by UAV also grows significantly. Traditional detecting methods use some classical machine learning algorithms, such as support vector machine (SVM), random forest or adaboost etc. and combine the low level features such as gradient, colors or texture to detect electrical devices. These image features must be carefully designed and changed a lot from various object kinds. Thus, they are not suitable for UAV images with complex background and multiple kinds of object. On the other hand, the disadvantages of these methods are that they cannot take advantage of the high quantity and large coverage of UAV acquired images, and cannot get a satisfactory accuracy. The recent developing Deep Learning method brings light to this problem. Convolutional neural network (CNN) performs excellently in object recognition area and outstand many other methods used in the past. Without the need of extracting images' features, CNN becomes the many state-of-the-art methods in object recognition rapidly. In object detection, Region-based convolutional neural networks (RCNN) retrieves the region that may contain the object from the images to detect and recognize the object. However, the computation is so expensive that it cannot meet the requirement of detecting massive UAV's images and cannot be used in practical projects. Fast R-CNN and Faster R-CNN solve this problem by changing the way of object retrieval. They use features produced by CNN network layers and apply a region proposal network layer behind to locate the object. After that, fully connected layers and softmax layer follow to classify the features corresponding to object into special kinds. Using this strategy, Fast R-CNN and Faster R-CNN save lots of time to produce region proposal and can perform object detection at nearly real time. The principle and processes of Faster R-CNN and several other object detection methods are described in this paper, and they are tested for electrical devices detection from images of the power transmission line obtained by UAV. We analyzed the influence of several key parameters to the device detection results, such as the dropout ratio, non-maximum suppression (nms) and batch size. Then, we gave some constructive advice of tun-

收稿日期 2016-03-17; 修回日期: 2016-06-21.

基金项目 2014年国家电网公司发展项目“无人机巡检实用化关键技术及检测体系研究”。

作者简介 王万国(1984-), 男, 硕士, 中级工程师, 主要从事基于数字图像的变电站和输电线路设备识别、目标跟踪、图像去雾等研究。E-mail: wangwanguo03@qq.com

ing parameters in Faster R-CNN. We also analyzed the advantages and weakness of three advanced detection algorithms, including Deformable Part Models (DPM) and two deep learning-based methods named Spatial pyramid pooling networks (SPPNet) and Faster R-CNN. Finally, we constructed image datasets of power transmission line inspection obtained by UAV and tested the three methods above. The recall ratio and accuracy ratio of them are compared and the superiority of the Faster R-CNN is validated. Testing results showed that Faster R-CNN method can detect various electrical devices of different categories in one image simultaneously within 80 milliseconds and achieve an accuracy of 92.7% on a standard test set, which is of great significance in real-time power transmission line inspection. These results also showed the advantages of the Faster R-CNN and we apply Faster R-CNN in our practical projects to detect electrical devices.

**Key words:** deep learning; RCNN; Region-based convolutional neural networks; UAV images; electrical devices detection

**\*Corresponding author:** WANG Wanguo, E-mail: wangwanguo03@qq.com

**摘要** 随着无人机(UAV)在电力巡线作业中的应用推广,对无人机巡检图像的信息挖掘或目标识别需求也越来越强烈。传统的电力部件识别流程常使用经典的机器学习算法,如支持向量机(SVM)、随机森林或adaboost,结合梯度、颜色或纹理等浅层特征来对电力部件进行识别,难以充分利用无人机巡检图像的信息,并且难以达到较高的准确率。卷积神经网络(CNN)在目标识别中表现优异,在很多目标识别场景之中成为首选算法。基于区域的卷积神经网络(RCNN)通过使用CNN从图像中提取可能含有目标的区域来检测并识别目标,但是计算复杂,难以满足识别海量电力巡检图片的需求。Fast R-CNN和Faster R-CNN利用CNN网络提取图像特征,后接一个区域提议层,优化了提取可能含有目标区域的方式并改进识别目标的分类器,使目标的检测和识别几乎实时。本文详细描述了Faster R-CNN算法流程,并在无人机电力巡线巡检图像部件检测中使用,然后分别对DPM、SPPnet和Faster R-CNN识别方法进行了对比分析,利用实际采集的电力小部件巡检数据构建的数据集对3种方法进行试验验证,并讨论了不同参数对识别结果的影响。实验结果表明,基于深度学习的识别方法实现电力小部件的识别是可行的,而且利用Faster R-CNN进行多种类别的电力小部件识别定位可以达到每张近80 ms的识别速度和92.7%的准确率。

**关键词** 深度学习;RCNN;卷积神经网络;无人机巡检图像;电力部件识别

## 1 引言

近几年随着无人机(Unmanned Aerial Vehicle, UAV)应用的逐渐普及,电力巡线无人机受到各大电网公司的广泛关注,具有广阔的应用前景。一方面,无人机巡线具有野外作业风险低、成本低以及作业灵活的特点;另一方面,产生的海量数据需要经过人工判读才能得到最终的巡检报告,因此采用图像识别的方法对这些数据进行部件检测识别具有十分重要的意义。

与传统的小部件识别相比,无人机巡检得到的图像具有背景复杂、小部件与背景对比度低、不同地区不同季节背景差异较大、存在大量干扰等难题。传统的电力部件识别算法主要采用人工设计的特征,如SIFT(Scale-invariant feature transform)<sup>[1]</sup>、边缘检测符<sup>[2]</sup>、HOG(Histogram of Oriented Gridients)等,不能很好地适用于电力部件,采用的分割算法主要基于部件外围轮廓骨架<sup>[3]</sup>、自适应阈值<sup>[4]</sup>等进行图像分割。但这些方法在设计原则上往往是基于

特定类别来实现的,其准确率低,不具有可扩展性;而且方法结构松散,缺少对低层特征进行综合利用进而达到全局最优识别的目的。与之相比,Malik团队的轮廓检测及层次图像分割方法<sup>[5]</sup>和多尺度组合聚合(Multiscale Combinatorial Grouping, MCG)<sup>[6]</sup>方法,以及Uijlings等提出的基于选择性搜索(Selective Search)的目标识别方法<sup>[7]</sup>给出了将多种低层特征进行全局优化并构建层次结构模型的范式,提高了准确率,但这些方法尚不具备随样本数量增多提升识别准确率的能力。

2012年起,深度学习引起了人们的广泛关注,并在图像识别与检测中取得了良好的识别效果。本文研究了深度学习在电力部件识别中的应用,以及采用优化算法对参数进行调优,并比较了DPM(Deformable Part Models)<sup>[8]</sup>、基于RCNN(Region based Convolutional Neural Network)<sup>[9]</sup>的SPPnet(Spatial pyramid pooling networks)<sup>[10]</sup>和Faster R-CNN<sup>[11]</sup>3种算法,针对电力小部件识别问题分析了不同算法的效果和性能。

## 2 经典DPM方法和RCNN

对于目标识别主要包括目标位置的确定和目标类别的判别。根据确定目标位置的方式不同,可以分为2类:①采用滑动窗口的方式,逐窗口判断是否存在目标对象;②采用区域提议的方式,先集中生成可能包含目标对象的区域框,再逐一判断每个候选框是否包含目标对象。滑动窗口方式识别方法的典型算法是可变形部件模型DPM;区域提议方式识别的典型算法是基于区域的卷积神经网络RCNN。

### 2.1 可变形部件模型DPM

可变形部件模型DPM<sup>[12]</sup>方法是由P. Felzenszwalb提出的经典目标识别算法。在检测阶段,DPM在图像特征金字塔上作为一个滑动窗口运行,图像特征金字塔通常由HOG特征建立。DPM通过优化一个综合部件变形代价函数和图像匹配得分的得分函数来给每个滑动窗口赋予一个得分<sup>[8]</sup>。

### 2.2 基于区域提议的卷积神经网络RCNN

Ross等2014年提出的基于区域的卷积神经网络方法RCNN<sup>[9]</sup>,成为基于区域提议方式进行识别的典型方案。在检测阶段,RCNN分为4个步骤:①使用视觉方法(如Selective Search)生成大量候选区域;②对每个候选区域用CNN进行特征提取,形成高维特征向量;③将这些特征量送入一个线性分类器计算属于某个类别的概率,用于判断所包含对象;④对目标外围框的位置和大小进行一个精细的回归。

与DPM使用滑动窗口进行遍历搜索的方式相比,RCNN第一步的区域提议是选择性搜索,使用得分最高的前2000个区域可以有效减少后面特征提取的计算量,能很好地应对尺度问题;CNN在实现上采用GPU进行并行计算,计算效率明显优于DPM方法(实现上采用单CPU计算);外围框回归使目标定位的精确性进一步提升。在训练阶段,RCNN也有4个步骤:①使用选择性搜索集中生成每张图片的候选区域,并对每个候选区域用CNN提取特征,本文CNN采用的是训练好的ImageNet网络;②利用候选区域和提取出的特征对ImageNet网络进行调优,调优依据标准的反向传播算法进行,从特征层开始向后调整各层权重;③以特征层输出的高维特征向量和目标类别标签为输入,训练

支持向量机;④训练对目标外围框位置和大小进行精细回归的回归器。

RCNN方法在准确率和效率上远远超过DPM方法,成为基于深度学习进行识别的典型方案。2014年和2015年,Ross和微软亚洲研究院的研究者陆续提出了改进的RCNN方法:SPPnet<sup>[10]</sup>首次引入空间金字塔池化层从而放宽了对输入图片尺寸限制并提高准确率;Fast R-CNN<sup>[11]</sup>采用自适应尺度池化能够对整个网络进行调优从而提高深层网络识别的准确率;Faster R-CNN<sup>[11]</sup>通过构建精巧的区域提议网络来代替时间开销大的选择性搜索方法,从而打破了计算区域提议时间开销大的瓶颈问题,使实时识别成为可能。本文主要研究了利用Faster R-CNN方法对电力部件进行识别。

## 3 基于Faster R-CNN方法的电力部件识别定位

与SPPNet和Fast R-CNN相比,Faster R-CNN<sup>[11]</sup>方法既突破了计算区域提议的时间瓶颈,又能保证理想的识别率。因此,本文以Faster R-CNN识别方法为主,提取电力小部件的识别特征并进行目标识别验证。

### 3.1 电力部件识别的网络训练

Faster-RCNN方法包含2个CNN网络:区域提议网络RPN(Regional Proposal Network)和Fast R-CNN检测网络。训练阶段的主要步骤如图1所示。对RPN网络和Fast R-CNN检测网络进行联合训练,如图2所示。

#### (1)预训练CNN模型

RPN网络和检测网络都需要对预训练的ImageNet网络进行初始化<sup>[11]</sup>,通常采用的网络主要有

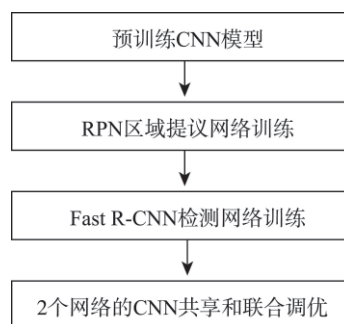


图1 部件识别的联合网络训练过程

Fig. 1 Flowchart of devices detection in joint network training



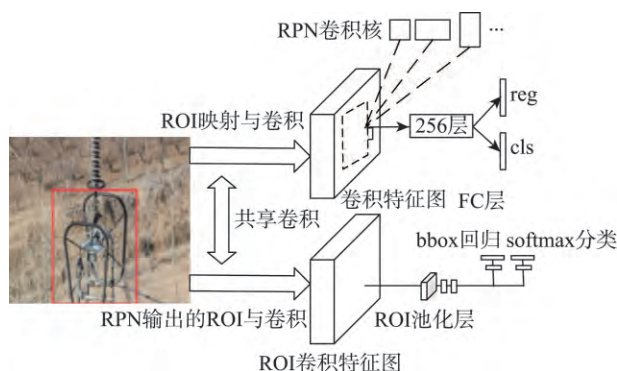


图2 网络训练过程示意图

Fig. 2 The schematic diagram of network training

ZFnet网络(Zeiler and Fergus)<sup>[14]</sup>和VGG16<sup>[15]</sup>网络(Simonyan and Zisserman moda)。因本文数据集规模较小,故选用ZFnet网络。ZFnet包含5个卷积层,有些卷积层后面添加池化层和3个完全连接的特征层。利用ILSVRC 2012图像分类任务中的训练数据(120万张图像,1000类)对ZFnet模型进行预训练。区域提议网络和检测网络都是在ZFnet输出后添加特定的层得到。这些特定层可以对输入图片提取出可能含有目标的区域,并计算出以该区域为目标概率。

ZFnet的最后一个卷积层(即第5个卷积层)包含256个通道,被称为特征图(Feature Map)。特征图为输入图像的深层卷积特征,同类物体的深层特征十分接近;而不同类物体的深层特征差异很大,即在特征图上物体具有很好的可分性。

### (2)RPN网络训练

用电力部件图像构建图像训练集,但电力部件图像集与预训练图像集无论是类别数量还是图像样式都存在很大的差别。在用电力部件图像集训练RPN网络时,直接用上一步预训练的ZFnet模型初始化RPN,使用反向传播算法对区域提议网络进行调优。

RPN网络以任意大小的图像为输入,之后输出

一系列可能包含目标的区域框。如图3所示,在ZFnet的CONV5后面添加一个小的卷积层,这个小的卷积层采用滑动方式运作,对于特征图上的每一个位置(对应原始图像上一个位置),由小卷积层进行卷积运算,即在此位置开一个小窗口进行卷积运算,得到同一个位置对应的256维向量(由于有256个通道),该向量反映了该位置小窗口(对应原始图像上某一窗口)内的深层特征。由这个256维的特征向量可以预测:①该位置小窗口属于目标/背景的概率值,即得分;②该位置附近包含目标的窗口相对于该位置小窗口的偏差,用4个参数表示,2个平移,2个放缩。

采用3种不同尺寸和3种不同比例(1:1, 1:2, 2:1)组合成的9种基准小窗口对包含目标的窗口位置进行预测,可以使区域提议更准确。

### (3)Fast R-CNN检测网络训练

根据步骤(2)生成的区域提议结果是基于Fast-RCNN方法训练独立的检测网络,检测网络也利用ZFnet预训练模型初始化。

对输入图像进行5层卷积网络的特征提取,第5层特征图(CONV5)是一个256×256的特征图,取出CONV5上对应的深度特征,将256个通道内的全部特征串联成一个高维(4096维)特征向量,称为FC6特征层,后面添加另一个4096维的特征层,形成FC7,FC6和FC7之间采用完全连接。由FC7特征层可预测:①候选区域框属于每个类别的概率,即得分;②目标对象外围框的更合适的位置,用它相对于候选区域框的2个平移和2个放缩共4个参数表示。通过预先标记的信息利用反向传播算法对该检测网络进行微调。

### (4)2个网络的CNN共享和联合调优

将2个网络单独训练并未实现卷积网络的参数进行共享。

利用步骤(3)训练的检测网络来初始化RPN网

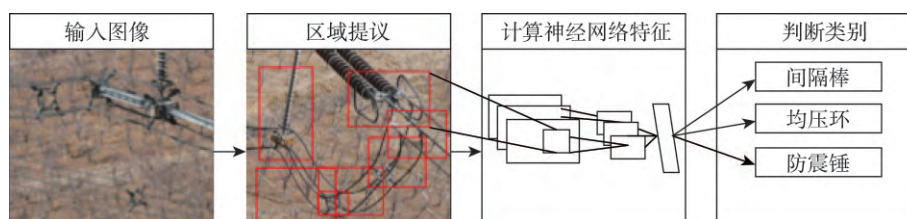


图3 检测识别过程

Fig. 3 Detection identification process

络,并固定共享的深度卷积层(如图2中红色双向箭头所指),对RPN网络的特殊部分进行调优,为了与检测网络对应,称此部分为RPN网络的FC层,这样2个网络就共享了深度卷积层。

最后,固定共享的卷积层,对Fast R-CNN的FC层进行调优。这样2个网络就共享了卷积层并形成了一个联合的网络。

### 3.2 检测识别过程

由上面的训练可知,2个网络最终可共用同一个5层的卷积神经网络,这使整个检测过程只需完成系列卷积运算即可完成检测识别过程,彻底解决了原来区域提议步骤时间开销大的瓶颈问题。

检测识别的过程如图3所示,其实现步骤为:

(1)对整个图像进行系列卷积运算,得到特征图CONV5;

(2)由区域提议网络在特征图上生成大量候选区域框;

(3)对候选区域框进行非最大值抑制<sup>[16]</sup>,保留得分较高的前300个框;

(4)取出特征图上候选区域框内的特征形成高维特征向量,由检测网络计算类别得分,并预测更合适的目标外围框位置。

## 4 结果与分析

无人机拍摄影像具有分辨率较高、包含目标较小的特点,拍摄影像的角度具有多样性和一定随机性。本文识别3类小型电力部件—间隔棒、防震锤和均压环。

### 4.1 训练样本处理

数据集来源于多旋翼无人机和直升机巡检图像,从季节上覆盖了春、夏、秋、冬4个季节。原始影像大小为5184像元×3456像元(图4(a)),截取以目标为主体的正方形小块图像,统一放缩至500像元×500像元(图4(b)),作为训练样本。

### 4.2 训练集和测试集构建

本次试验,对于间隔棒、均压环和防震锤的每一类部件,均使用1500张训练样本,共4500张样本构成训练集;每类500张测试影像,共1500张影像构成测试集。对训练集中每张图片里完整出现的没有被遮挡的小型电力部件标记其外围框(训练集

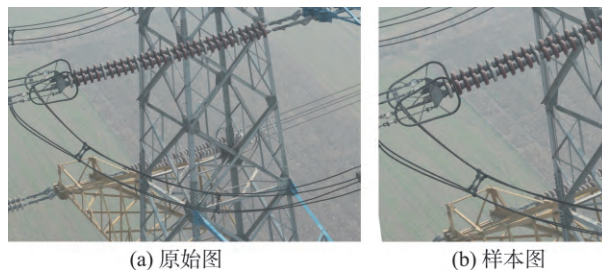


图4 原始数据和训练样本数据

Fig. 4 The original data and the training sample data

图片中不完整或被遮挡的电力部件不标记);而对测试集,要标出每张图片里出现的所有电力部件,包括不完整的和被遮挡的。

测试时,当识别出的外围框与标记的外围框重叠面积达到标记外围框的90%以上时,视为一次成功识别。本次试验中,用正确率和召回率来评判识别的准确性,其中正确率为目标类别标记正确的外围框个数除以所有标记出的外围框个数;召回率为目标类别标记正确的外围框个数除以所有标准的外围框个数。由于本次试验识别的类别仅有3种类型,因此分别对每一类电力部件识别的正确率和召回率做统计。

### 4.3 实验结果

本文使用Caffe框架实现卷积神经网络模型。使用3.2节构建的训练集和测试集,首先研究了多个Faster-RCNN参数对mAP(平均准确率均值)的影响,然后将Faster R-CNN方法与基于Selective Search方法进行区域提议的SPPnet方法和DPM方法进行对比。

Faster-RCNN涉及到一些参数,例如dropout比例、最大迭代次数、批处理尺寸、nms(非极大值抑制)前后区域保留个数,这些参数对mAP有较大影响。表1-3为改变参数时候的试验结果。

根据表1,当dropout的比例从0.2增大至0.8时,mAP总体上是下降趋势,但是在0.6时有一个最高值。目前并无相关理论解释dropout对mAP的影响,通常取经验值。

令dropout取值为0.6,改变nms前后候选区域数量,测试其对mAP的影响,结果如表2所示。根据表2,随着nms数量的减小,mAP也逐渐减小,这是因为经过nms后,保留的候选区域也随之减小,导致检测结果准确度下降。因此较高的nms可以得到较好的检测结果。

令nms前后候选区域个数分别取2000和300,

表1 不同dropout比例对mAP影响

Tab. 1 Different dropout rate effects on mAP

dropout 比例	最大迭代次数	区域提议阶段批尺寸	检测阶段批尺寸	nms 前候选区域个数	nms 后候选区域个数	mAP
0.2	8000	256	128	2000	300	0.827
0.3	8000	256	128	2000	300	0.811
0.4	8000	256	128	2000	300	0.817
0.5	8000	256	128	2000	300	0.791
0.6	8000	256	128	2000	300	0.829
0.7	8000	256	128	2000	300	0.781
0.8	8000	256	128	2000	300	0.775

表2 不同nms数目对mAP影响

Tab. 2 Different numbers of nms effects on mAP

最大迭代次数	区域提议阶段批尺寸	检测阶段批尺寸	nms 前候选区域个数	nms 后候选区域个数	mAP
8000	256	128	2000	300	0.829
8000	256	128	1500	250	0.818
8000	256	128	1000	200	0.806
8000	256	128	500	100	0.802

表3 不同的批尺寸对mAP影响

Tab. 3 Different batch sizes effects on mAP

最大迭代次数	区域提议阶段批尺寸	检测阶段批尺寸	nms 前候选区域个数	nms 后候选区域个数	mAP
8000	256	128	2000	300	0.829
8000	128	64	2000	300	0.83
8000	64	32	2000	300	0.847
8000	32	16	2000	300	0.848

改变批尺寸,测试其对mAP的影响,结果如表3所示。由表3可以看出,不同的批尺寸得到不同的mAP。随着批尺寸逐渐变小,mAP逐渐增大。有理论<sup>[17]</sup>表明,当batch\_size为1时,优化速度最快。

根据mAP最大时对应的参数,对测试集使用Faster-RCNN与SPPnet、DPM进行部件识别,对应的3类电力部件的正确率和召回率如表4所示。从表4可看出,Faster R-CNN方法识别的准确率明显高于SPPnet和DPM,而DPM方法准确率最低。这主要是由于区域提议网络可以产生比SPPnet更精准的候选框,而DPM方法采用滑动窗口进行检测,其特征为HOG特征,而不是深度训练特征。此外,Faster R-CNN在网络训练的第(2)步对全部特征层和卷积层的权重进行调优,而SPPnet仅调优特征层,从而限制了识别准确率。值得注意的是,Faster R-CNN采用的区域提议网络和检测网络具有很好的泛化能力,能够识别出部分被遮挡和中间穿过铁杆的间隔棒,且对各种不同方向的部件都可正确识

表4 在测试集上识别准确的对比

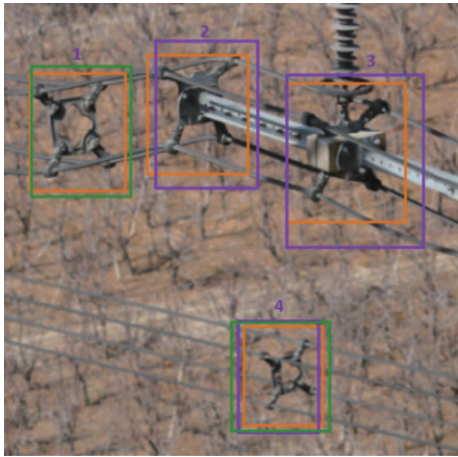
Tab. 4 Comparison of accuracy identified on the test set

识别能力/%	间隔棒	均压环	防震锤
Faster R-CNN-正确率	91.2	92.7	84.3
Faster R-CNN-召回率	88.5	84.3	79.3
SPPnet-正确率	85.6	86.4	79.1
SPPnet-召回率	80.4	78.6	73.7
DPM-正确率	52.2	60.7	51.5
DPM-召回率	51.9	55.2	49.8

别。图5是利用3种方法对同一张图片进行间隔棒识别的对比情况,其中绿色、紫色、橙色分别代表DPM、SPPnet和Faster RCNN的识别结果。以左上角第一个间隔棒为起点,按照顺时针方向定义图中4个间隔棒序号分别为1、2、3、4。表5表示每种方法将间隔棒识别为间隔棒的概率。

本文所有试验均基于同一台服务器而进行测试,测试集图片大小为5184像素×3456像素,DPM方法基于CPU而实现,Faster R-CNN方法和SPPnet





注:绿色为DPM;紫色为SPPnet;橙色为Faster R-CNN

图5 3种方法对间隔棒的识别结果

Fig. 5 The results of three methods of spacer recognition

都使用Nivdia Titan Black GPU(6G显存)进行卷积计算,识别过程耗用3G显存。另外,Faster R-CNN

表5 DPM、SPPnet、Faster RCNN对4个间隔棒的识别概率

Tab. 5 Detection probability of 4 spacers with regard to DPM, SPPnet and Faster RCNN

方法	间隔棒1	间隔棒2	间隔棒3	间隔棒4
DMP	0.84	—	—	0.89
SPPnet	—	0.81	0.73	0.90
Faster RCNN	0.98	0.86	0.75	0.96

的非最大值抑制也采用GPU实现。从表6可以看出,DPM的运算时间在分钟级,无法与其他两种方法进行时间效率的对比;对于SPPnet这样的典型的RCNN方法,区域提议占据了主要的计算时间;而Faster R-CNN中,由于卷积特征的共用(区域提议网络和检测网络的特殊层都添加在共用的特征图CONV5的后面),使得区域提议时间几乎可以忽略不计,检测时间可以在近80 ms内即可完成。结果表明,采用深度学习方法基于特定的图形加速卡可以实现巡检图像的实时检测。

表6 Faster R-CNN和SPPNet、DPM方法计算时间开销对比

Tab. 6 Comparison of computation expense of SPPNet method and DPM method

方法	平均时间	区域提议数量	关键步骤计算时间	关键步骤计算时间	关键步骤时间
Faster R-CNN	77 ms	平均17 000,取前2000	卷积+区域提议时间47 ms	非最大值抑制+外围框优化时间30 ms	非最大值抑制
SPPnet	27.3 s	3000至15000,取前2000	Selective Search区域提议24.1s	卷积特征提取2.5 s	0.7 s
DPM	224 s	—	—	—	—

5 结语

本文在总结分析当前几种典型的目标检测识别方法的基础上,验证了利用RCNN等深度学习算法对电力小部件识别的准确性和效率,并分析了不同参数对Faster R-CNN检测结果的影响。实验表明,利用特定的GPU计算单元可以实时目标检测和识别,可为后期无人机巡检图像的智能化处理及巡检无人机的精确拍摄奠定良好的基础。

此外,根据深度学习的特点,构建更大的样本库可能进一步提高准确率。下一步的工作是构建更精细的识别类别,包括某些部件的缺陷图像都可看做一种类型,这样不但可以实现目标部件的分类,还可实现部件的缺陷识别。

参考文献(References):

[1] 苑津莎,崔克彬,李宝树.基于ASIFT算法的绝缘子视频图像的识别与定位[J].电测与仪表,2015,7(7):106-112.  
[Yuan J S, Cui K B, Li B S. Identification and location of

insulator video images based on ASIFT algorithm[J]. Electrical Measure and Instrumentation, 2015,7(7):106-112. ]  
[2] 吴庆岗.复杂背景输电线路图像中部件边缘提取算法研究[D].大连:大连海事大学,2012. [Wu Q G. Study on the algorithm for edge extraction of components in power line images with complex backgrounds[D]. Dalian: Dalian Maritime University, 2012. ]  
[3] 金立军,胡娟,闫书佳.基于图像的高压输电线路间隔棒故障诊断方法[J].高电压技术,2013,39(5):1040-1045. [Jin L J, Hu J, Yan S J. Method of spacer fault diagnose on transmission line based on image process[J]. High Voltage Engineering, 2013,39(5):1040-1045. ]  
[4] 曹婧.航拍输电线路图像中绝缘子部件的提取[D].大连:大连海事大学,2012. [Cao J. Insulators extraction in aerial image of inspecting transmission line[D]. Dalian: Dalian Maritime University, 2012. ]  
[5] Arbelaez P, Maire M, Fowlkes C C, et al. Contour detection and hierarchical image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011,33(5):898-916.

- [ 6 ] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. Computer Vision and Pattern Recognition, 2014.
- [ 7 ] Uijlings J, De Sande K E, Gevers T, et al. Selective search for object recognition[J]. International Journal of Computer Vision, 2013,104(2):154-171.
- [ 8 ] Girshick R, Iandola F, Darrell T, et al. Deformable part models are convolutional neural networks[C]. Computer Vision and Pattern Recognition, 2015.
- [ 9 ] Girshick R, Donahue J, Darrell T, et al. Region-based convolutional networks for accurate object detection and segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016,38(1):142-158.
- [10] He K, Zhang X, Ren S, 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.
- [11] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [C]. Neural Information Processing Systems, 2015.
- [12] Felzenszwalb P F, Girshick R, Mcallester D, et al. Object detection with discriminatively trained part-based models [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010,32(9):1627-1645.
- [13] Girshick R. Fast R-CNN[C]. International Conference on Computer Vision, 2015.
- [14] Zeiler M D, Fergus R. Visualizing and understanding convolutional networks[J]. Springer International Publishing, 2014,8689:818-833.
- [15] Simonyan K, Zisserman A. Very Deep Convolutional networks for large-scale image recognition[C]. International Conference on Learning Representations, 2015.
- [16] Neubeck A, Van Gool L. Efficient non-maximum suppression[J]. International Conference on Pattern Recognition, 2006,3:850-855.
- [17] Bottou L. Large-scale machine learning with stochastic gradient descent[M]. Proceedings of COMPSTAT 2010. USA: Physica-Verlag HD, 2010:177-186.