

两级上下文卷积网络宽视场图像小目标检测方法

王海涛^{1,2}, 姜文东³, 程远^{1,2}, 严碧武^{1,2}, 张宗峰⁴, 李涛^{1,2}, 张森海⁵

(1. 国网电力科学研究院武汉南瑞有限责任公司, 武汉 430073; 2. 南瑞集团有限公司, 南京 211106;
3. 国网浙江省电力有限公司, 杭州 310007; 4. 国网山东省电力公司日照供电公司, 山东日照 276826;
5. 国网浙江省电力有限公司嘉兴供电公司, 浙江嘉兴 314599)

摘要: 目标检测和识别已经在输电线路巡检中被广泛采用; 然而由于宽视场图像数据量大, 小目标相对宽视场较小, 分辨率低, 现有的图像金字塔、特征金字塔和多异构特征融合等方法虽能准确地检测大目标, 但小目标的检测精度低, 处理非常耗时, 因而快速、准确地检测宽视场图像中小目标仍是一个挑战; 提出一个两个 Faster-RCNNs 级联的上下文宽视场小目标检测卷积网络, 首先, 针对降分辨率的宽视场图像, 利用一个 Faster R-CNN 来检测目标的上下文区域, 然后, 针对上下文区域对应的高分辨率原始图像, 利用 Faster R-CNN 来检测小目标; 用航拍输电线路图像数据集进行了多尺度目标的检测试验, 试验结果表明, 文章提出的目标检测方法达到了 88% 的检测精度, 检测精度明显优于单级 Faster R-CNN 检测方法。

关键词: 小目标检测; 无人机图像; 输电线路巡检

Two-stage Context Convolutional Network for Small Target Detection In Wide-view-field Images

Wang Haitao^{1,2}, Jiang Wendong³, Cheng Yuan^{1,2}, Yan Biwu^{1,2},
Zhang Zongfeng⁴, Li Tao^{1,2}, Zhang Senhai⁵

(1. Wuhan NARI Limited Liability Company, State Grid Electric Power Research Institute, Wuhan 430074, China;
2. NARI Group Corporation Ltd. Nanjing 211106, China; 3. State Grid Zhejiang Electric Power Co., Ltd.,
Hangzhou 310007, China; 4. Rizhao Power Supply Company, State Grid Shandong Electric Power Company,
Rizhao 276826, China; 5. State Grid Jiaying Power Supply Company, Jiaying 314003, China)

Abstract: Object detection and recognition has been widely applied to power transmission line inspection. Existing methods, such as multi-scale image pyramid, multi-scale feature pyramid and multiple heterogeneous feature fusion, etc. can detect small objects accurately, but usually require heavy computational burden, thus fast and precise target detection in wide-view-field images is still challenging due to large amount of image data and low resolution of small targets. In this paper, we propose a two-stage context convolutional network for small target detection in wide-view-field images, which consists of two cascaded Faster R-CNNs, the first Faster R-CNN is used to locate context regions in a low resolution image, and another Faster R-CNN to detect small targets in high-resolution images of detected context regions. We test the proposed method is test on our data sets captured by unmanned aircraft, experimental results show that the proposed method could lead to 88.0% accuracy for small target detection and is higher than that of the one-stage Faster R-CNN.

Keywords: small object detection; UAV images; power transmission line

0 引言

由于高压和超高压输电线路大都处于崇山峻岭、交通死区、通讯盲区, 甚至无人区, 人工巡视不仅工作量大、周期长, 而且巡视结果受视点、遮挡及天气、地形等客观因素影响大。近年来, 随着无人机和计算机视觉技术的发展, 提出了基于无人机的输电线路巡检方法。它首先利用无人机搭载的高清相机, 对绝缘子污秽、裂痕、爆片及塔材锈蚀等线路关键部位及缺陷的细节进行巡视, 然后, 将

视频实时传输至地面站, 最后再由计算机取代人工来进行线路关键部件检测和故障识别。由于观察视点、光照、天气和关键部件与缺陷及其所处背景的复杂变化, 使得巡视图像具有复杂的多变性和多样性, 另外, 高清图像数据量大, 图像大小高达 $4\ 800 \times 2\ 704$, 而防震锤、间隔棒、挂点, 分压环和鸟巢等关键部件面积较小, 其图像常低于 300×100 像素, 因而, 如何准确地检测输电线路关键部位及其缺陷已成为无人机巡检技术研究的热点。

输电线路视觉巡检主要有基于手工设计特征的方法和基于深度特征学习的方法两类。早期的目标检测的特点是采用 SIFT^[1]、HOG^[2]、LBP^[3] 等人工设计特征提取算子。其基本思路是针对特定目标与场景的特点, 首先手动设计特征算子, 然后, 再进行特征的分类与识别。这类方法的

收稿日期: 2018-11-20; 修回日期: 2018-12-18。

基金项目: 国家电网公司科技项目(521104180025)。

作者简介: 王海涛(1977-), 男, 青海西宁人, 高级工程师, 主要从事输电线路智能监测与线路运维方向的研究。

缺点是目标检测对视点、光照、遮挡以及目标与场景变化的适应性差,不能满足输电线路巡检应用的需要。近年来提出的深度学习方法的基本思路是通过已标注的目标图像训练数据集,学习出深度特征网络的参数,然后,利用训练好的卷积网络,来检测和识别目标的位置和类型。深度学习检测方法主要有基于区域建议的方法和基于回归的方法两类。基于候选区域的方法主要有 R-CNN^[4], Fast R-CNN^[5]和 Faster R-CNN^[6]等,其优点是目标检测的精度高,缺点是目标检测的处理速度较慢;基于回归的方法主要有 SSD^[16-17]和 YOLO^[18]等,其优点是目标检测的处理速度快,缺点是小目标检测的精度较低。总的来说,深度学习检测的缺点是大量训练样本图像制备难度大,高分辨率图像深度网络的训练和目标的检测计算复杂度高;其优点是不需手工设计特征抽取算法,通过捕获光照、位置、尺度、旋转以及目标与场景等各种变化的训练样本,就可自动训练出目标特征的深度网络参数,极大地提高了较大尺度目标检测对视点、光照及背景变化的适应性,深度学习为输电线路关键部位与缺陷检测提供了一种新的技术途径。

1 相关工作

内容随着深度学习技术的不断发展,目标检测和识别技术取得了跨越式的发展,国内外学者已经在 TPAMI、IJCV、TIP、CVIU 以及 CVPR、ICCV、ECCV 等国际顶级期刊或会议上发表了许多重要的研究成果,主要有基于滑动窗口的方法和基于候选区域的方法两种。基于滑动窗方法的思路是针对输入图像或特征图,采用滑动窗搜索的方法,判断各个窗口是否包含目标。Felzens-walb 提出一种显著 Deformable Parts Model (DPM) 检测算法^[8],首先利用 HOG 特征金字塔滑动窗搜索方法,通过优化部件形变参数和匹配分数,实现形变目标的检测,其缺点是计算量大。Pierre Sermanet 利用多尺度的滑动窗口,提出了一种 OverFeat 特征提取算法^[9],用卷积网络将分类、定位和检测统一起来,提高了目标特征提取的适应性,其缺点是滑动窗生成的候选区域特征提取存在重复计算;基于候选区域方法的思路是先提取候选区域,然后,对候选区域进行分类和识别,主要有 R-CNN^[4], Fast R-CNN^[5]和 Faster R-CNN^[6]等。R-CNN 是先利用选择性搜索方法产生候选区域,然后,采用卷积网络提取目标的特征,再利用 SVM 进行特征分类^[4,12],其缺点是候选区域相互重叠,卷积特征存在重复计算,处理速度慢。Kaiming He^[13]提出了一个 SPPnet 深度卷积网络空间金字塔池化目标检测方法^[13],在最后一个卷积层和其后的全连接层之间,增加一个空间金字塔池化层,提取任意大小候选区域的固定大小特征图,其特点是输入任意大小的图像,处理速度和精度比 R-CNN 高。Girshick, Ross 提出了一种 Fast R-CNN 检测方法^[5],首先将选择性搜索得到的目标候选框投影到特征图,并池化采样成固定大小的特征,然后,再利用回归分析检

测目标的位置和类型。其特点是通过共享卷积网络计算,提高了目标检测的处理速度。Shaoqing Ren 提出了区域建议网络 Faster R-CNN 检测方法^[6],首先,通过卷积特征图产生 RPN 区域建议,然后结合 Fast R-CNN 网络,实现目标检测,其特点是 RPN 生成共享卷积特征计算,提高了目标的检测精度和速度。PASCAL VOC^[10]和 MS COCO^[11]等数据集上的检测试验表明,基于候选区域的 R-CNN 方法在处理的速度和精度方面要优于基于滑动窗口的方法。由于深度卷积网络在过滤掉光照、视点、遮挡以及尺度变化特征的同时,也丢失了小目标的细节特征,导致小目标的检测精度较低。宽视场高分辨率图像虽然提高了小目标检测的精度,但同时也增加了计算的复杂度。因此,如何快速、准确、可靠地检测宽视场中的小目标仍是输电线路关键部位与缺陷检测研究的难点。

为了提高小目标特征表述的精度,Kong Tao 提出了 HyperNet 检测方法^[14],将高层的语义信息和底层高分辨率信息结合起来,提高多尺度目标的表述精度。Lin, T. Y. 提出了特征金字塔检测方法^[15],在每个特征层做预测,利用多尺度特征信息,提高了多尺度目标的检测精度,其缺点是检测速度慢。Liu Wei 提出了基于多特征融合和 SSD 的小目标检测方法^[18],引入小目标的上下文信息,提高了目标检测的速度和精度。另外,文献 [19-20] 还提出通过视图聚合和对抗网络,来提高小目标特征的精度,来改善小目标的检测精度,但其增大的输入图像尺度会导致训练和测试的计算和硬件成本的增加。为了解决小样本引起的目标检测精度低问题,Goodfellow^[20]提出了 GAN 对抗生成网络方法^[20],并将 GAN 应用于图像生成,来扩充训练样本集;Li Jianan 提出了一种 Perceptual GAN^[22],通过学习每个类别不同尺度物体之间特征图的内在结构相关性,再利用变换后的特征,来提高小目标检测的准确率。其缺点是对抗生成模型通常难以训练且收敛。此外,文献 [23-24] 和 [7] 还提出将上下文信息和语义分割引入目标检测,利用上下文信息改善了目标检测的精度和速度。由于航拍图像中防震锤、间隔棒、绝缘子挂点和分压环等小目标虽然具有稳定的上下文背景,但是,小目标在上下文背景中的位置是随机变化的,因而现有小目标检测方法的漏检率较高。

受上述研究工作的启发,本文提出一种两级 Faster R-CNN 级联的目标检测方案,它根据小目标上下文背景特点,采用两个 Faster-RCNN 来分别学习和建模小目标及其上下文背景特征,首先执行上下文背景区域检测,获取大尺度上下文背景,然后,将检测出的上下文候选区域映射回原始图像,提高小目标特征图的分辨率。图 1 给出了提出的两级 Faster-RCNN 级联上下文小目标检测的处理框图,第一级针对降分辨率的宽视场图像,检测小目标的上下文区域。然后,将上下文对应的原始图像送给第二级的 Faster R-CNN,进而实现宽视场中的小目标精确检测。图 2 分别给出了将分辨率图像中小目标特征图和上下文背景中

小目标的特征图。不同图像分辨率的特征表达是完全不同的, 可以看出上下文候选区域中小目标包含更多的细节特征, 小目标特征表述的精度更高, 本文提出的方法能够先检测小目标所在的语义背景, 然后, 映射回原图提高候选区域的分辨率, 这样, 可以提高小目标的分辨率, 从而提升其检测性能。本文第 III 节将针对航拍图像中分压环、防震锤、间隔棒和悬挂装置等小目标的上下文背景特点, 详细阐述提出的两级 Faster-RCNN 级联上下文小目标检测算法, 第 IV 节给出了实验结果。

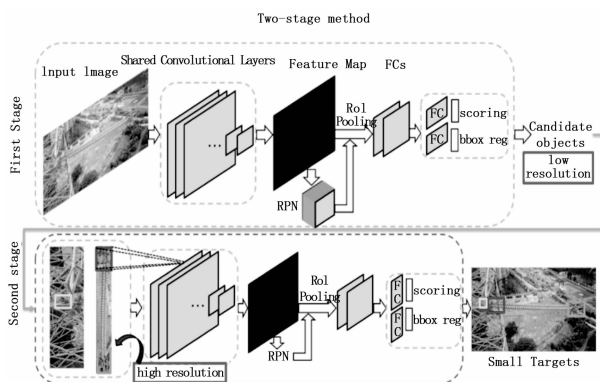


图 1 两级级联 Faster R-CNN 的联合框架
(橙色线表示两级 Faster R-CNN 之间的级联关系)

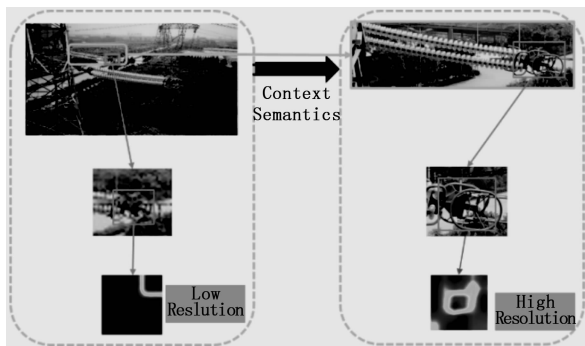


图 2 高层卷积层中提取的特征图所做的可视化效果

2 两级 Faster R-CNN 级联的航拍图像小目标检测算法

研究表明, Faster R-CNN 减少了生成候选区域的数量计算量, 因而在提高检测速度的同时, 也保证了检测的准确性。Faster R-CNN 由 RPN 候选区域生成 RPN 网络和 Fast R-CNN 检测器两个模块组成。RPN 网络先采用 3 种长宽比的锚框生成候选区域, 然后, 将 RPN 的输出作为 Fast R-CNN 的输入, 并通过细化微调来做预测。Faster R-CNN 特点是 RPN 和 Fast R-CNN 共享卷积特征, 可用 VGG16^[25], ResNet152^[26] 等非常深的基础网络来进行高质量的预测。在无人机巡检实验中, 航拍原始图像为 $4\,800 \times 2\,704$ 像素, 小目标的尺寸通常低于 300×100 像素。由于深度卷积网络在过滤掉光照、视点以及尺度变化特征的同时, 也丢失了小目标的细节特征, 在网络的深层, 小目标会变

得更小, 在深层特征图中可能还不到 1 个像素, 小目标的特征图变得十分粗糙, 导致线路关键部件及其缺陷等小目标的检测精度较低。因此, 直接用 Faster R-CNN 进行检测, 航拍图像中的小目标漏检率较高。

为了提高无人机航拍宽视场高清图像中小目标的检测精度, 我们引入小目标的上下文信息, 将检测分为两个阶段。首先, 利用语义分割方法^[7], 生成可能包含小目标的上下文候选区域。传统的语义分割方法主要由自适应阈值分割^[27], 自适应区域分割^[28], 超像素分割^[29] 和 SIFT 特征匹配^[1] 等传统的语义分割方法, 这些方法十分依赖于人工设计的特征, 不能自动的从训练数据中学习得到特征表示, 对不同检测应用的适应性较差。本文采用一个 Faster R-CNN 方法来检测可能包含有小目标的几类上下文大目标候选区域。然后, 针对提取到的上下文大目标候选区域对应的原始图区域, 利用第二级的 Faster R-CNN 来检测和定位小目标, 因而在提高宽视场图像中小目标检测精度的同时, 也保持了较高的处理速度。

2.1 训练 RPN

本节简要概述 RPN 结构和锚。RPN 可以接收任意大小的输入图像并输出可能包含目标的一组区域。为了生成候选区域, 我们用一个小窗网络在共享卷积层的最后一层滑动, 将 256 维的 ZFnet 或 512 维的 VGG16 卷积特征映射到低维, 并将该向量输入到两个子全连接层, 然后, 再连到边界框回归层和边界框分类层。由此预测区域含有目标的概率以及预测框相对于真实框的偏移。考虑到目标的尺度和旋转变换, Faster R-CNN 采用了 3 种不同的长宽比 (1: 1, 1: 2, 2: 1) 组合和 3 个尺度组合的 9 个 anchor 来预测目标的多尺度窗口, 使得区域建议更加准确, 再利用随机梯度下降 (SGD) 方法进行端到端训练 RPN 分类和回归分支。

2.2 训练 Fast R-CNN

首先, 使用 VGG16 预训练模型来初始化基础网络。将 RoI 降采样到 7×7 特征图, 与 VGG16 的第 13 层卷积层特征结合起来, 然后, 通过全连接层 (FCs) 映射成 4 096 维特征向量。对于每个 RoI 有两个输出向量: softmax 类别概率和边界框的回归偏移。模型训练时, 使用 SGD 端到端的优化多任务损失, 并对检测网络进行微调。图 3 给出了近似训练的过程, 即使用 VGG16 预训练模型来初始化共享卷积层。当训练 Fast R-CNN 时, RPN 网络的参数被固定。

2.3 近似联合训练

Faster R-CNN 中讨论了三种方法共享卷积特征来训练网络。在本文中, 我们采用近似联合训练方法。虽然它不像非近似联合训练 (考虑边界框坐标可微分的 RoI 池化层) 那样精确, 但是实验结果与其他两种方法非常接近。首先, 将候选目标输入第二级 Faster R-CNN 训练网络。当训练 Faster R-CNN 检测器时, 由 RPN 生成的候选区域被固定。在反向传播时, 同时优化 RPN 损失和 Fast R-CNN 损失。详细的训练过程如图 3 所示。网络整个训练过

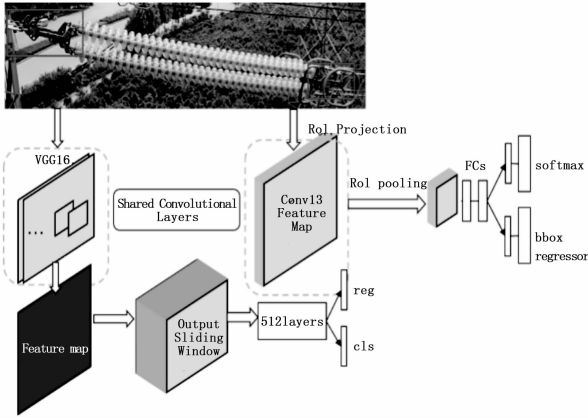


图 3 近似训练的过程

程的概况如图 4 所示。

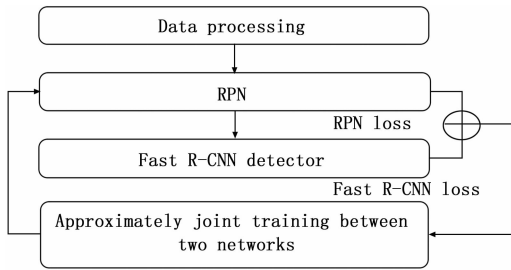


图 4 前向和反向传播网络训练过程示意图

2.4 多尺度输入

我们在第一阶段先检测大目标，以降低输入尺度并消除对小目标的冗余干扰。具体而言，首先通过 Faster R-CNN 检测电力塔架和绝缘子区域，然后用第二级 Faster R-CNN 检测鸟巢、挂点和分压环。两个大目标都是任意尺度的。电力塔大概是原始图像的三分之一大小；绝缘子通常是矩形，其长宽比大于 10。因此，我们需要预处理在第二阶段中富含上下文信息的候选区域，来适应 RPN 的输入要求。对于在数据集中这两个大目标，我们采用随机裁剪的方法来解决这个问题。绝缘子候选框在真实框附近裁剪，保证长宽比保持在 1:2，并且短边缩放到 600 像素。

3 实验

本节将针对高分辨率宽视场中小物体图像，评估和比较两级 Faster R-CNN 的级联和其他方法的检测精度和速度性能。

3.1 数据集

实验中，使用的国网南瑞公司提供的无人机拍摄的数据集，该数据集有 3 700 张图片。每个图像的大小为 4 800 × 2 704。图 5 展示了无人机采集的一些图像。所有图像都是高分辨率的，包含了多类小尺寸的目标。在不同图像中，这些目标不仅尺度、位置以及所处背景差异很大，而且，观察视角也具有很大的不确定性和随机性，在航拍图像中这些目标可能被遮挡，因而，这使得我们的数据集非常具有挑战性。我们进行数据清洗，筛掉不包含目标的图像以

及差异非常小的图像。



图 5 数据集的示例。蓝绝缘子、鸟巢、塔号牌、挂点和分压环等检测结果

实验中检测的小目标有：鸟巢，塔号牌（表示电塔的数量或标号），挂点（悬挂绝缘子的装置）和分压环。如在第 3 节中所述，在第一阶段需要获取包含上下文信息的候选区域，所以我们还标注了 2 个电气设备：电力塔架和绝缘子。

3.2 实验设置

我们用预训练的 ImageNet [30]，VGG16 [29] 来提取特征，训练两阶段 Faster R-CNN。使用 PyTorch 框架来实现文中 CNN 架构。每批随机抽取 300 张图片进行训练。为了合理使用 GPU 内存，如果图像宽度大于 2 000 像素，则根据 $1\ 200/w$ 的比例大小进行调整， w 和 h 保持不变，其中 w 和 h 是图像的宽度和高度。使用 SGD 求解器迭代 10 000 次，基本学习率为 0.001。此外，anchor 与真实框的 IoU 重叠大于 0.7 的设为正样本。考虑到正样本十分稀缺的情况，为了均衡正负样本，我们采样更加偏向于正样本。

3.3 第一阶段：大目标检测结果

在第一阶段，用 Faster R-CNN 在原始图像上做语义分割，从而提取可能包含小目标的主要对象。表 1 给出了第一阶段和整个阶段的平均准确率。我们在电力塔和绝缘子检测中获得了非常高的精度，其主要原因是塔架和绝缘子尺度比较大，容易学习到稳定的可辨识特征。

3.4 第二阶段：小目标检测结果

在第二阶段，采用相同的检测过程来评估小目标检测性能。在第一阶段之后，将预先检测到的上下文区域输进到第二检测网络。表 1 给出了单级检测的结果。

表 1 单级检测方法的结果

项目	大目标		小目标			
	电力塔	绝缘子	鸟巢	塔号牌	挂点	均压环
AP	0.951	0.923	0.752	0.732	0.923	0.914
mAP	0.937		0.830			
时间/ms	142		142			

3.5 与单级 Faster R-CNN 的比较

为了验证本文提出的小目标检测方法的有效性，我们对单级的 Faster-RCNN 和级联的 Faster-RCNN 做了对比实验。即一种是用单级 Faster R-CNN 直接检测原图中的小目标，另一个实验是利用本文的两级 Faster-RCNN 级联检测方法。如图 6 所示，给出了这两种不同方法的部分结果图像。表 2 给出了每类目标的平均准确率。实验结果表明，对于近场拍摄的目标图像，这两种方法都可以检

测到小目标, 但本文级联方法的回归框更加准确。此外, 由于鸟巢和塔号牌的样本数量有限, 并且尺度比挂点和分压环要小得多, 故挂点和分压环的准确率明显要比鸟巢和塔号牌高。对于远景拍摄的宽视场鸟巢、塔号牌等小目标图像, 本文提出的方法的结果明显优于直接用单级 Faster R-CNN 检测方法, 其检测的精度提升了 13%。

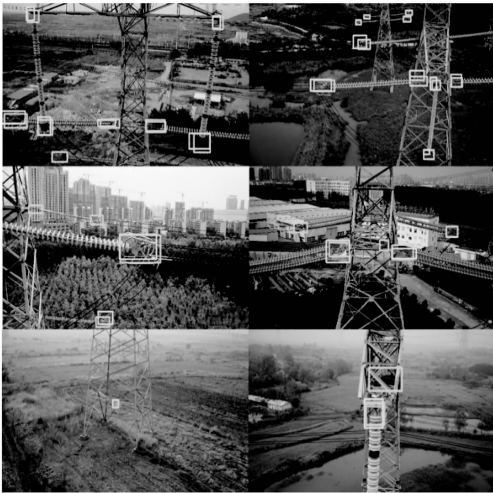


图 6 单级和两级级联 Faster R-CNN 方法的检测结果

表 2 每类目标的准确率

测试项目	间隔棒	防震锤	绝缘子	鸟巢	塔号牌	电力塔	精度 mAP	时间 ms
单个 Faster-RCNN	0.851	0.875	0.923	0.752	0.732	0.951	0.847	142
级联 Faster-RCNN	0.853	0.875	0.922	0.851	0.835	0.946	0.88	180

3.6 不同 anchor 长宽比的对比实验

由于场景图像中不同目标的长宽比变化很大, 测试中采用了不同的 anchor 进行小目标检测实验。下面列举了 3 种 anchor 的参数: 1:1, 1:2, 2:1; 1:1, 1:3, 3:1; 1:1, 1:5, 5:1。表 3 给出了这 3 种纵横比的实验结果。从实验结果可以看出, 长宽比为 1:1, 1:2, 2:1 的 anchor 比其它长宽比的 anchor 有着更好的检测精度。

表 3 不同 anchor 纵横比的结果

纵横比	1:1 1:2 2:1	1:1 1:3 3:1	1:1 1:5 5:1
间隔棒	0.851	0.851	0.844
防震锤	0.875	0.864	0.836
绝缘子	0.923	0.924	0.918
鸟巢	0.752	0.746	0.741
塔号牌	0.732	0.711	0.704
电力塔	0.951	0.950	0.948
mAP	0.847	0.841	0.831

4 结语

本文针对宽视场输电线路巡检现场图像, 提出了一种

两级 Faster R-CNN 级联的上下文小目标检测方法, 采用两个 Faster-RCNN 深度学习卷积神经网络分别对上下文背景和小目标的特征检测进行建模, 提高了候选区域内小目标图像的分辨率和处理速度。首先, 采用一个 Faster-RCNN 检测上下文候选区域, 然后, 针对上下文候选区域映射到原始图像, 利用另一个 Faster-RCNN 检测上下文候选区域中的小目标, 提高了宽视场图像中小目标检测的精度和速度。针对输电线路巡检的现场宽视场图像数据库, 目标检测试验表明, 与单级 Faster R-CNN 相比, 本文提出的方法不仅具有更高的小目标检测精度, 同时, 对于 $4\ 800 \times 2\ 704$ 这样大幅面的图像也保持了较高的处理速度, 满足了输电线路快速巡检应用的要求。

感谢武汉南瑞有限责任公司无人机巡检组提供的视频数据。

参考文献:

- [1] Lowe D G. Object recognition from local scale-invariant features [A]. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR) [C]. 1999, 2: 1150-1157.
- [2] Dalal N, Triggs B. Histograms of oriented gradients for human detection [A]. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR) [C]. 2005, 1: 886-893.
- [3] Ojala T, Pietikainen M, Maenpaa T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2002, 24 (7): 971-987.
- [4] Girshick, Ross, et al., Rich feature hierarchies for accurate object detection and semantic segmentation [A]. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR) [C]. 2014, 580-587.
- [5] Girshick, Ross, Fast R-CNN [A]. IEEE international conference on computer vision (ICCV) [C]. 2015: 1440-1448.
- [6] Ren S Q, et al. 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.
- [7] García G M, et al. Semantic segmentation priors for object discovery [A]. The International Conference on Pattern Recognition (ICPR) [C]. 2016: 549-554.
- [8] Felzenszwalb P F, et al. Object detection with discriminatively trained part-based models [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32 (9): 1627-1645.
- [9] Sermanet, Pierre, et al., Overfeat: integrated recognition, localization and detection using convolutional networks [A]. International Conference on Learning Representations (ICLR) [C]. 2014, arXiv: 1312. 6229, 2014.
- [10] Everingham M, et al. The PASCAL visual object classes (VOC) challenge [J]. International Journal of Computer Vision. 2010, 88 (2): 303-338.

[11] Lin T Y, et al. Microsoft COCO: common objects in context [A]. European Conference on Computer Vision (ECCV) [C]. 2014, arXiv: 1405. 0312, 2014.8693: 740 - 755.

[12] Uijlings J R R, et al. Selective search for object recognition [J]. International Journal of Computer Vision, 2013, 104 (2): 154 - 171.

[13] Kaiming He, 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.

[14] Tao Kong, et al. , HyperNet: Towards Accurate Region Proposal Generation and Joint Object Detection [A]. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR) [C]. 2016, 845 - 853.

[15] Lin T Y, et al. Feature pyramid networks for object detection [A]. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR) [C]. 2017: 936 - 944.

[16] Liu W, et al. SSD: single shot multibox detector [A]. European Conference on Computer Vision (ECCV) [C]. 2016, Part I, LNCS 9905, 2016, 21 - 37.

[17] Cao G M, et al. Feature-fused SSD: fast detection for small objects [Z]. arXiv: 1709.05054, 2017.

[18] Redmon J, et al. You only look once: unified, real-time object detection [A]. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR) [C]. 2016: 779 - 788.

[19] Ku, Jason, et al. Joint 3D proposal generation and object detection from view aggregation [Z]. arXiv: 1712.02294, 2017.

[20] Goodfellow, Ian J, et al. Generative adversarial nets [A]. International conference on neural information processing systems (NIPS) [C]. 2014, arXiv: 1406.2661, 2014.

[21] Mathieu M, Couprie C, Lecun Y. Deep multi-scale video prediction beyond mean square error [A]. International Conference on Learning Representations (ICLR) [C]. 2015, arXiv: 1511.05440.

[22] Li J N, et al. , Perceptual generative adversarial networks for small object detection [A]. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR) [C]. 2017: 1951 - 1959.

[23] Sean B, et al. Inside-outside net: detecting objects in context with skip pooling and recurrent neural networks [A]. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR) [C]. 2016, 2874 - 2883.

[24] Li B, et al. Object detection via end-to-end integration of aspect ratio and context aware part-based models and fully convolutional networks [A]. CoRR [C]. abs/1612. 00534, 2016.

[25] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition [A]. International Conference on Learning Representations (ICLR) [C]. 2015, arXiv: 1409.1556, 2015.

[26] He, K M, et al. , Deep residual learning for image recognition [A]. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR) [C]. 2016: 770 - 778.

[27] Wang M, et al. An adaptive threshold segmentation algorithm for gesture segmentation [J]. Applied Mechanics and Materials. 2014, Vols. 513 - 517: 457 - 460.

[28] Wani M A, Batchelor B G. Edge-region-based segmentation of range images [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994, 16 (3): 314 - 319.

[29] Freeman, William T, Jones T R, et al. Example-based super-resolution [J]. IEEE Computer Graphics and Applications, 2002, 22 (2): 56 - 65.

[30] Zeiler, Matthew D, Fergus R. Visualizing and understanding convolutional networks [A]. European Conference on Computer Vision (ECCV) [C]. 2014: 818 - 833.

[31] Deng J, et al. ImageNet: A large-scale hierarchical image database [A]. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR) [C]. 2009: 248 - 255.

（上接第 198 页）

DA 转换，采用直接由 FPGA 输出合成信号后经反向放大，再加入直流偏置的方法，可降低部分硬件成本。

同时，此方法也可推广到其他类型的发动机转速与扭矩的仿真试验中，为用户提供高性能和低成本解决方案，保证实验的准确性和有效性，提高工作效率，缩短研制周期。

参考文献:

[1] 章伟, 张跃明. 一种航空涡轴发动机输出扭矩转速信号测量方法 [J]. 测控技术, 2014, 33: 133 - 135.

[2] 何宾, 张艳辉. Xilinx Zynq-7000 嵌入式系统设计与实现 [M]. 北京: 电子工业出版社, 2016.

[3] 符晓, 张国斌, 朱洪顺. Xilinx ZYNQ-7000 AP SoC 开发实战指南 [M]. 北京: 清华大学出版社, 2016.

[4] 谢尚豪, 张硕, 许波. 一种高精度频率测量电路设计 [J]. 电子测量技术, 2018 (18): 115 - 121.

[5] 吴厚航. 勇敢的芯一伴你玩转 Xilinx FPGA [M]. 北京: 清华大学出版社, 2017.

[6] Xilinx. LogiCORE IP DDS Compiler v6.0 [EB/OL]. https://www.xilinx.com/support/documentation/ip_documentation/dds_compiler/v6_0/pg141-dds-compiler.pdf, 2017-12-20/2018-12-6.

[7] Texas Instruments. DAC8811 16-Bit, Serial Input Multiplying Digital-to-Analog Converter [EB/OL]. <http://www.ti.com/cn/cn/lit/ds/symlink/dac8811.pdf>, 2016-2-1/2018-12-6.

[8] Analog Device. Quad Precision, High Speed Operational Amplifier OP467 [EB/OL]. <https://www.analog.com/media/en/technical-documentation/data-sheets/OP467.pdf>