

doi: 10.3969/j.issn.1007-290X.2020.009.020

深度学习在电力设备锈蚀检测中的应用

范海兵¹, 胡锡幸², 刘明一³, 肖俊⁴

(1. 国网浙江省电力有限公司检修分公司, 浙江 宁波 315000; 2. 国网浙江省电力有限公司杭州供电公司, 浙江 杭州 310007; 3. 国网浙江省电力有限公司金华供电公司, 浙江 金华 321000; 4. 浙江大学 计算机科学与技术学院, 浙江 杭州 310012)

摘要: 电力设备在长期运转中会产生形态不规则的锈蚀, 人工巡检工作存在难度大、效率低、成本高等问题, 而采用传统的图像处理方法对锈蚀进行检测的效率和准确率都不高。为此, 提出基于深度学习的电力设备锈蚀检测技术。该方法以变电站、输电网络等电力设备实际作业环境中的监控设备拍摄的图像数据作为训练样本, 采用基于深度学习的目标检测技术, 识别并定位电力设备上的锈蚀区域。以 Faster R-CNN 模型和 YOLOv3 模型作为神经网络的基础模型进行验证, 结果表明 2 个模型都能检测出较为明显的锈蚀, Faster R-CNN 模型在准确度和检测数量方面表现较好, 而 YOLOv3 模型检测速度更快。

关键词: 深度学习; 目标检测; 锈蚀检测; 电力设备; 卷积神经网络

中图分类号: TM769; TP18

文献标志码: A

文章编号: 1007-290X(2020)09-0154-12

Application of Deep Learning in Power Equipment Corrosion Detection

FAN Haibing¹, HU Xixing², LIU Mingyi³, XIAO Jun⁴

(1. Maintenance Company of State Grid Zhejiang Electric Power Co., Ltd., Ningbo, Zhejiang 315000, China; 2. Hangzhou Power Supply Company of State Grid Zhejiang Electric Power Co., Ltd., Hangzhou, Zhejiang 310007, China; 3. Jinhua Power Supply Company of State Grid Zhejiang Electric Power Co., Ltd., Jinhua, Zhejiang 321000, China; 4. College of Computer Science and Technology, Zhejiang University, Hangzhou, Zhejiang 310012, China)

Abstract: Irregular shape corrosion will occur in the long-term operation of power equipment, and there are some problems in manual inspections such as large difficulty, low efficiency, and high costs. The detection efficiencies and accuracy of traditional image processing methods for corrosion detection are not high. Thus, this paper proposes a technology based on deep learning for power equipment corrosion detection which uses images captured by the monitoring equipment in the actual operating environment of power equipment such as substations and transmission networks as training samples, adopts the objective detection technology based on deep learning to identify and locate corrosive areas on the power equipment. The Faster R-CNN model and YOLOv3 model are used as the basic model of the neural network for verification and the results show that both models can detect relatively obvious corrosion. The Faster R-CNN has better performance in accuracy and detecting numbers, and the YOLOv3 model performs better in detection speed.

Key words: deep learning; object detection; corrosion detection; electric power equipment; convolutional neural network (CNN)

随着社会的快速发展, 社会用电需求逐年上涨。据《浙江省电力发展“十三五”规划》披露, 在“十二五”末的 2015 年, 浙江全省最高负荷和用电量分别达到 62.9 GW、355.3 TWh 左右。此外, 依据《浙江省电力发展“十三五”规划》中推荐方案的

预测, 在“十三五”末的 2020 年, 全省最高负荷、用电量将分别达到 84.0 GW、453.0 TWh。随着电网规模的扩大, 对电力供应的可靠性的要求也在不断提高。电力设备在长期运转的过程中, 不可避免地会出现锈蚀问题。而作为基础设施的电力设备一旦出现严重的锈蚀缺陷问题, 极有可能导致设备故障, 影响整个相关区域的电力系统正常运行, 从而严重影响生产生活并带来难以预计的经济损失。

收稿日期: 2020-01-16 修回日期: 2020-03-19

基金项目: 国家自然科学基金项目(61976185)

这就需要对相应的电力设备进行定期巡检和及时的维护检修。及时发现锈迹并检修锈蚀问题，能够有效减少电力设备的恶劣性故障，从而保障电网安全稳定运行。

传统的电气设备状态查验主要是依靠人工的方式对电力设备进行定期的巡视检查，这种人力巡检方式存在许多问题^[1]：①巡检人员的安全难以保证。与其他普通巡检任务不同，对电力设备的检视通常需要工作人员暴露在大量高负荷运转的电气设备中，且这些设备随时可能发生故障，具有一定的安全风险。②巡检效率低。安排人力定期逐个检查设备，对管理人员和作业人员都是很大的考验，工作强度较大，如果遇到恶劣天气则更加影响工作效率。③巡检准确度不高。人力巡检的判断主观性比较强，尤其是长期重复作业，更容易降低该作业的稳定性。

2019 年，国家电网有限公司做出了加快构建“三型两网”的主要战略部署，提出了构建状态全面感知、信息高效处理、应用便捷灵活的泛在电力物联网。构建设备状态全感知、数据高度信息化的能源互联网已经成为必然的发展趋势，针对电力设备的锈蚀，有必要研究准确而又高效的智能巡检方法。

网络通信基础设施的建设，使视频图像采集任务得到保障。目前主要的图像采集方式是无人机巡检^[2-4]和核心设备间安装的全天候视频监控^[5]。这些智能化的设备会产生大量的巡检图像，并实时传输到后台系统中，再由监控中心的工作人员根据捕获到的视频图像资料判断是否存在异常。依靠人工查验这些采集到的视频图像信息并找出设备锈蚀的方案虽然存在可行性，但是效率低、反馈慢，而且人力成本昂贵；因此，对于智能化识别方法的需求十分迫切^[6]。常规的图像处理方法泛化能力不足，这些方法通过产生滑动窗口对整幅图像进行遍历，得到候选区域后，提取区域内的边缘、轮廓等特征，然后送入分类器进行目标物体的类别判断。这种基于常规图像处理的目标检测方法存在一些问题：①滑动窗口往往是设定固定的 1 个或者多个尺寸，没有灵活性，对图片中不同大小的物体检测任务冗余性太大；②边缘、轮廓这种人工预设的特征鲁棒性较差^[7]。

近年来，得益于单位成本计算能力的提高，机器学习、深度学习发展迅猛。尤其是深度学习下的目标检测技术，得到了学术界和工业界的广泛认可^[8-12]。本文基于目标检测领域中主流的 Faster R-CNN 算法和 YOLOv3 算法，提出基于深度学习

的电力设备锈蚀检测技术，并从高精度和高速度 2 个方面进行实例验证，实现从视频、图片数据流中检测并定位实际作业环境中电力设备上的锈蚀区域。

1 基于深度学习的电力设备锈蚀检测

1.1 研究现状

目标检测是采用卷积神经网络(convolutional neural network, CNN)等深度学习方法来快速定位图像中感兴趣目标的技术，它不仅要检测出目标所属的类别，还要定位出目标所在的位置和大小，该方向是当前计算机视觉的研究热点。2012 年之前，目标检测算法主要是基于手工提取目标特征，因此算法的实际效果很大程度上取决于人力在模型训练和参数调整上的经验，泛化能力差而且对复杂对象的特征学习效果不好。2012 年 AlexNet 模型^[13]的提出，标志着深度学习、目标检测乃至计算机视觉进入了全新的阶段。发展至今，深度学习下的目标检测技术主要分为单阶法和双阶法 2 类。双阶法是先产生候选框，然后将候选框中的图像特征送入分类器进行预测并得到最终的分类结果，代表的模型有区域卷积神经网络(region-based convolutional neural network, R-CNN)^[14]和它的改进版本 Fast R-CNN^[15]和 Faster R-CNN^[16]等；单阶法是将目标检测作为回归任务，对图像的网格直接进行回归操作，不生成候选框，代表的模型有 YOLO(you only look once)^[17]、SSD(single shot multi detector)^[18]和 RetinaNet^[19]等。

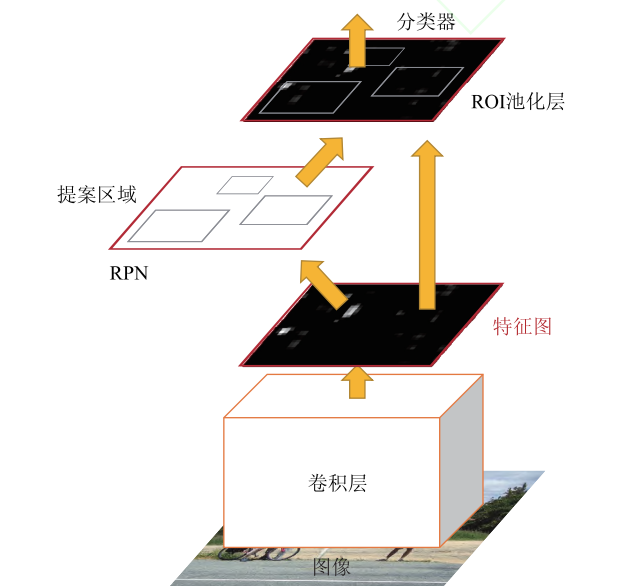
目前，基于深度学习的目标检测技术开始逐渐应用到电网电力系统中^[20-24]。双阶法凭借训练速度快、提取的特征可复用等优势，成为众多非实时性检测任务的首选方法。上海大学的汤踊、韩军^[25]等人利用 Faster R-CNN 来实现输电线路不同部件的识别与分类，并提出通过调整 CNN 模型的卷积核大小和通过对图像的旋转变换来扩充数据集的方法，更好地提高模型的识别准确率和缩短识别时间；上海电力大学的马静怡^[26]等人针对 Faster R-CNN 算法在电力设备状态图像处理过程中生成的候选框与目标设备不匹配而导致的设备识别率降低问题，提出依据像素值相近及位置相邻原则构建连通域，将连通域的长宽比作为输入信息对 Anchor 候选框进行修正；长沙理工大学的王超洋^[27]等人提出了基于融合特征金字塔网络(feature pyramid network, FPN)结构的 Faster R-CNN 配电网架空线

路异常状态检测系统,可以更好地识别变压器等小目标物体。在单阶法方面,樊绍胜等人提出基于改进的 SSD 电力设备红外图像异常自动检测方法^[28]和基于 YOLOv3 的输电线路故障检测方法^[29],不仅泛化性强,而且可以在极短的时间内对待检测图像实现物体的定位和分类,从而实现实时自动检测,使现有电力巡检设备实现“智能+”。

在锈蚀检测领域,难点主要在于如何解决锈迹不规则和数据集样本不足的问题。对此,中国石油大学的薛冰提出基于 Mask R-CNN 的电力设备锈蚀检测识别方法^[30],采用 Faster R-CNN 完成目标检测的功能,采用全卷积网络(fully convolutional network, FCN)完成语义分割的功能,最后实现像素级别的分类识别,一定程度上解决了识别不规则图像的问题。浙江大学的纪扬等人^[31]在电缆隧道铁锈识别这项任务中,引入迁移学习的概念,很好地解决了基于中小规模数据集训练的 CNN 难以获得较高识别精度的问题。

1.2 Faster R-CNN

R-CNN 是目标检测领域中十分经典的方法,相比于传统的手工特征提取方式,R-CNN 使用 CNN 来提取图像的深度特征,然后接 1 个分类器来预测候选区域的类别,并得到最终的结果。Faster R-CNN 是 R-CNN 的改进方法,网络框架如图 1 所示,该算法由区域提案网络(region proposal net, RPN)候选框提取模块和 Fast R-CNN 检测模块组成。



ROI—感兴趣区域, region of interest 的缩写。
图 1 Faster R-CNN 网络框架
Fig. 1 Faster R-CNN framework

1.2.1 RPN

以往的目标检测算法生成候选框都非常耗时,如 OpenCV 中的 AdaBoost 使用滑动窗口和图像金字塔来产生建议窗口,R-CNN 使用基于遍历的选择性搜索(selective search, SS)^[32]方法。Faster R-CNN 丢弃传统的滑动窗口和暴力遍历方法,使用 RPN 来生成建议窗口,明显地加速了候选框的区域提案。

RPN 可以通过图形处理器(graphics processing unit, GPU)来加速网络训练计算速度,其中 Anchor 是 RPN 的核心。在得到特征图后,使用滑动窗口来生成建议区域,特征图上滑动窗口的中心映射在原图的 1 个区域就是 Anchor,然后以此为中心,生成 3 种尺度和 3 种长宽比两两组合成的 9 种 Anchors。特征图每个位置都会对应 9 个 Anchors,如果特征图的大小为 $W \times H$,那么这张图中 Anchors 的数量就是 $W \times H \times 9$ 。这种滑动窗口的方式可以关联整张特征图,最后得到多尺度、多长宽比的 Anchors。

特征图上的每个点前向映射得到 k ($k=9$) 个 Anchors,后向输出 512 维度的特征向量,而 Anchors 是用来获取用于分类和回归的区域候选框,因此全连接层后需要连接 2 个子连接层,即分类层和回归层(网络框架如图 2 所示)。分类层用于判断 Anchor 属于前景(foreground,待预测)还是背景(background,需要舍弃),向量维数为 $2k$;回归层用于计算 Anchors 的偏移量和缩放量,实现输出结果的精准定位,共 4 个参数 $\{X, Y, W, H\}$,分别代表该矩形候选框的中心点坐标和矩形的长宽,向量维数为 $4k$ 。

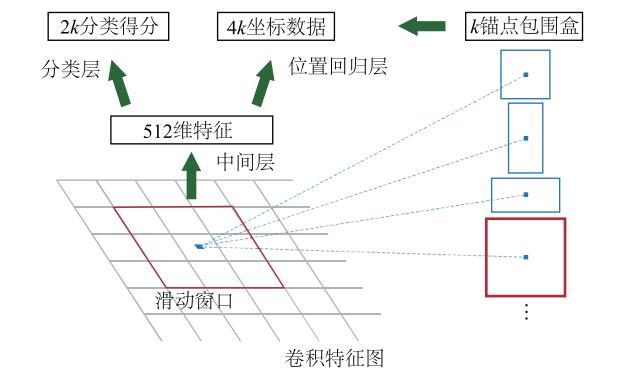


图 2 RPN 网络框架
Fig. 2 RPN framework

1.2.2 Fast R-CNN 分类网络

Fast R-CNN 用于对前面通过 RPN 生成的建

议区域进行分类和边框回归的计算, 从而获得精准的目标位置。如图 1 所示, 整个系统共享卷积特征图, 即通过 CNN 计算得到的特征图(feature map)不仅作为 RPN 的输入, 也作为 Fast R-CNN 的输入, 避免了重复计算, 加快了运行速度。

分类(classification)部分如图 3 所示, 首先网络结合 RPN 得到的 ROI 和主干网络输出的卷积特征(proposal feature), 通过全连接层(full connect layer)和 softmax 函数计算这些区域最终分类到哪个类别(cls_prob, 如人、车、树等), 输出的是一个概率向量; 同时, 再次利用矩形框回归(bounding box regression)计算每个建议区域的位置偏移量, 并不断修正, 得到更加精确的目标定位框(bbox_pred)。

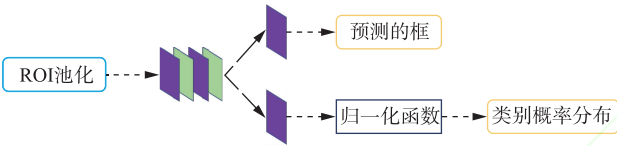


图 3 分类网络框架

Fig. 3 Classification network framework

1.2.3 模型训练

整个 Faster R-CNN 网络的损失函数 L 分为 2 个部分〔式(1)〕, 即分类损失 L_{cls} 和回归损失 L_{reg} 。分类损失〔式(2)〕是经典的二分类交叉熵损失, 回归损失〔式(3)〕是对 anchors 位置偏差的计算, 其中 f_{smooth} 为 smooth 函数表达式, 见式(4)。

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

$$L_{\text{cls}}(p_i, p_i^*) = -\log[p_i^* p_i + (1 - p_i^*)(1 - p_i)]; \quad (2)$$

$$L_{\text{reg}}(t_i, t_i^*) = \sum_i f_{\text{smooth}}(t_i - t_i^*); \quad (3)$$

$$f_{\text{smooth}}(x) = \begin{cases} 0.5x^2, & |x| < 1; \\ |x| - 0.5, & \text{其他}. \end{cases} \quad (4)$$

式(1)–(4)中: N_{cls} 用于归一化, 其值为该训练批次的大小; N_{reg} 用于归一化, 其值为 Anchor Location 的数量; λ 为常量, 用于平衡权重; i 为训练批次中 Anchor 的索引; p_i 表示对第 i 个 Anchor 作为前景(foreground)的预测概率, p_i^* 表示第 i 个 Anchor 作为前景的真实标签值(ground truth), 并使用 0 和 1 来标识; 向量 $t = (X, Y, W, H)$, t_i 表示对第 i 个 Anchor 的预测结果(t_i

$\in \{X, Y, X, H\}$), t_i^* 表示第 i 个 Anchor 的真实标签值。

1.3 YOLO

不同于 Faster R-CNN 这种 2 步检测算法, YOLO 这种单阶法是直接在待预测图像上计算候选框的类别概率和位置信息。YOLO 算法采用 1 个单独的 CNN 结构实现端到端(end-to-end)的目标检测, 不需要利用 RPN 来生成候选框, 因此 YOLO 算法的检测速度比 R-CNN 算法快。

1.3.1 网格划分下的回归和分类

YOLO 算法首先把 1 张图像拆分为 $S \times S$ 个方格(grid cell), 若 1 个物体的中心落在某个方格之内, 那么这个方格就负责预测这个物体。这样的话, 就不需要生成候选框, 直接对每个方格进行计算即可。比如, 图 4 中“狗”的中心坐落于 5 行 2 列的方格中, 所以这个方格就负责预测图像中的“狗”。具体来说, 每个方格会预测若干个边界框(Bounding Box)和每个边界框的置信度(confidence score)。其中置信度包括 2 个方面, 一个是考虑边界框含有“目标”的可能性, 另一个是边界框的准确度。前者记为 P_{object} , 当该边界框为背景时(即不包含任何物体), 此时 $P_{\text{object}} = 0$; 而当该边界框包含物体时, $P_{\text{object}} = 1$ 。边界框的准确度通过采用预测框和实际框(ground truth)的交并比(intersection over union, IOU)来表征, 记为 $I_{\text{truth}, \text{pred}}$ 。因此回归层面的置信得分定义为 $C = P_{\text{object}} \times I_{\text{truth}, \text{pred}}$ 。这样来看, 每个边界框的预测值包含 5 个元素 $\{X, Y, W, H, C\}$ 。

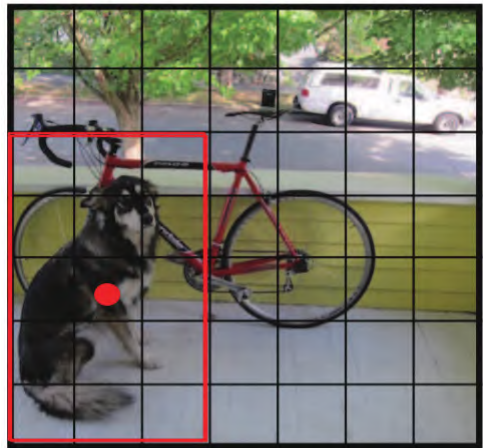


图 4 YOLO 算法网格划分示意图

Fig. 4 Schematic diagram of YOLO grid partitioning

对分类的处理和回归类似, 同样是基于拆分好的单元格, 模型网络对每个单元格预测 c 个类别的



图 6 “位置不敏感性”和“位置敏感性”的示例

Fig. 6 Examples of the dilemma between ‘translation-invariance’ and ‘translation-variance’

2 模型改进

2.1 空间变换网络

CNN 在图像分类中取得显著的成效，主要是得益于其深层结构具有空间不变性；因此，即使对图像中的目标物体作了平移变换或者旋转变换，CNN 模型依然可以将其识别出来。可以通过选择合适的降采样比例来保证准确率和空间不变性，但是池化层带来的空间不变性是不够的，它受限于预先选定的固定尺寸的感受野。STN 引进了一种可学习的采样模块 Spatial Transformer^[34]，该模块的学习不需要引入额外的数据标签，可以在网络中对特征图进行空间变换操作。

CNN 是尽可能让模型适应物体的形变，而 STN 是通过 Spatial Transformer 将形变后的物体变换回正常形态(比如将字摆正)，然后再送入分类器识别。如图 7 所示， U 为原始图像， V 为经由网络映射后的输出结果。图 7(a)中的图如果送入分类器进行识别的话，通常输出的结果是“字母 a”，但是经过仿射变换后分类器的输出结果是“数字 9”〔图 7(b)〕。

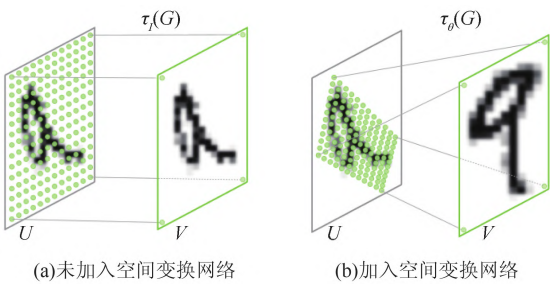


图 7 仿射变换示例

Fig. 7 Affine transformation examples

在实际生产作业环境中，监控探头等图像采集设备捕获的视频图像资料容易因光线干扰、设备安装位置移动和信号噪声等而产生畸变；因此，需要深度学习网络模型进行必要的修正，对获取到的 ROI 的特征进行变换，变为相对正确的姿态再送入分类器进行识别。

2.2 R-CNN 系列下的改进

2.2.1 主干特征网络的调整

Faster R-CNN 采用 VGG16^[35] 作为主干网络 (backbone) 来提取输入图像的特征，已经取得了不错的成绩。从经验来看，神经网络的深度对模型性能影响很大，当增加网络层数后，网络可以进行更深更复杂的特征提取，所以理论上网络层数越深，模型的效果应该越好；但是，大量实验结果表明深度的网络会出现退化现象，网络准确度出现饱和和甚至下降。深层网络存在梯度消失或者梯度爆炸问题，这使得深度学习模型很难训练。对此，带有“短路机制”的残差深度神经网络模型 ResNet^[36] 被提出，可以很好地解决退化问题。本文在实例验证部分，还会选用 ResNet 的升级版 ResNeXt 网络^[37] 进行对比实验。

2.2.2 更精准和精细的 RPN

对于 RPN 的改进，可以通过生成更多 Anchor 来实现。但是这样会增加计算量，降低模型预测速度。此外，针对电力设备锈蚀这一检测任务，分类单一而且数据集标注不够精确，刻意地增加 Anchor 未必能提高网络模型的效果，甚至会出现性能降低。

本文选用 FPN 对 Faster R-CNN 的 RPN 进行优化。FPN 先进行传统的自下而上 (bottom-up) 的特征卷积，然后对左侧横向对应特征图进行特征融合，并生成自顶而下 (top-down) 的特征，从而得到不同分辨率下的特征。可以发现，FPN 已经有不同大小的特征，无须像 RPN 中选用 3 种不同比率、尺寸的 Anchor，只需要 3 种比率不同的框，即每级 Anchor 的尺寸相同。

2.3 YOLO 系列下的改进

2.3.1 分类的置信得分计算

从前文对 YOLO 算法的介绍可知，对划分好的方格进行概率预测时，这些概率值其实是在各个边界框置信度下的条件概率，即 $P_{classi | object}$ 。因此，无论 1 个方格预测多少个边界框，其只预测 1 组类别的概率值。对此，可以把类别概率预测和边界框绑定在一起，实现方法如式(8)所示，同时可以计算出各个边界框类别的置信度。

$$P_{classi | object} \times P_{object} \times I_{truth, pred} = P_{classi} I_{truth, pred} \cdot \tag{8}$$

2.3.2 分类损失函数

YOLO 算法在初代版本使用的是 softmax 函数，这种损失函数会扩大最大类别概率值而抑制其他类别概率值。其后期版本中已更改为每一个类独立使用逻辑分类器，使用交叉熵(cross-entropy)损失，这样能够更好地处理标签重叠(包含)关系。

2.3.3 多尺度预测

YOLO 算法虽然检测速度快，但是初期最大的问题是对小尺寸物体的检测效果不理想。在 YOLOv3 中，通过对低分辨率的特征图与高分辨率的特征图进行特征融合的方法，形成新的特征图层，并对其进行单独预测，形成最终的多尺度预测。

值得注意的是，YOLO 系列中的多尺度有 2 种含义，一个是输入图像的多尺度，另一个是不同分辨率的特征图，本文选用后者。此外，原模型作者是先对图像进行归一化再缩放到统一尺寸。考虑到电力设备锈迹的不规则性，本文是先进行图像预处理直接缩放到统一尺寸，再送入 CNN，并在 YOLOv3 的特征金字塔结构(Darknet-53)的基础上额外增加 1 次下采样用于模拟类 FPN 结构。

3 实验验证

3.1 数据集的收集与预处理

通过对变电站、输电网络等电力设备实际作业环境的实地图像采集和对电网系统监控视频流的图片抽取，收集形成初步的图像数据集。在图像预处理阶段，对图片的文件格式和编码格式进行统一。本文实验中模型训练所采用的数据集格式统一为 RGB 三通道的 jpg 文件，通过编写程序进行格式转换和必要的数据清洗，图 8 为初筛之后数据集的部分样本图片。

3.2 数据集的标注

无论是 R-CNN 系列还是 YOLO 系列算法，都



图 8 数据集部分样本

Fig. 8 Dataset samples

属于全监督机器学习模型，所以需要对数据集样本进行标注，以便网络在训练过程中有可以作为参照基准的目标位置和类别标签。本文使用目前主流的一款开源标注软件 labelImg 对已经清洗完毕的图片数据集进行标注，这个软件可以帮助用户快速生成带有 Bounding Box(物体的目标位置)和 Label(物体类别)数据的 XML 文件，供程序在模型训练时读取。

3.3 数据增广

锈迹识别与其他目标检测任务的主要区别在于：不同设备上铁锈的形状和颜色相差较大，相同设备在不同生产环境下产生的锈迹也有所不同。此外，人力对数据集的标注主观性比较强，对于一些锈迹的边缘区域难以标注或者是选择性标注。因此，用于训练的锈迹图像数据集具有自身性的缺陷。对此需要采取一些额外的操作来尽可能降低这些缺陷造成的影响。如图 9 所示，本文对清洗完毕的数据集采取一些非形变的几何变换类和颜色变换类的数据增广方法，来实现对那些形状、颜色差异性较大而且样本数较少的锈蚀图像的数据扩充。

3.4 模型训练和结果分析

本文的实验环境为 Intel(R) Xeon(R) CPU E5-2667 v2 @3.30 GHz, 128 GB RAM, 单张 Nvidia GTX 1080Ti @12 GB, 以及 Ubuntu 16.04 LTS 操作系统。算法模型基于 PyTorch 机器学习框架，实验验证的指标包括准确率和召回率，准确率为正确检测

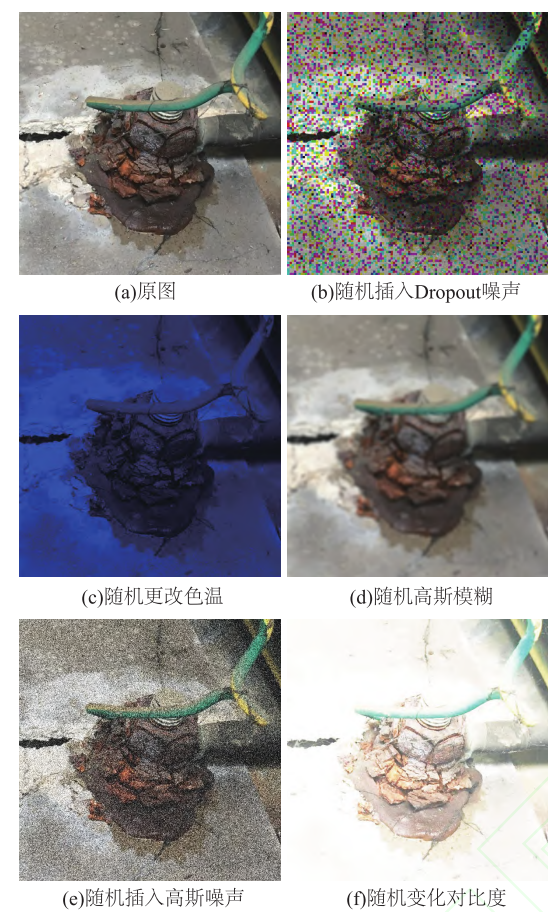


图 9 数据增广示例

Fig. 9 Examples of data augmentation

的目标数量与标记目标总数量之比，召回率为正确检测的目标数量与实际目标总数量之比。值得注意的是，本文对预测结果的阈值设置为 0.8，这是相对于应用场景过高的一个数值。阈值越低，最终得到的召回率和精准率越高。为了在消融实验中得到较好的对比效果，故采用较高的阈值。真实生产环境下的预测结果可参照后文的预测可视化结果。

3. 4. 1 实验结果的可视化

实验结果表明 Faster R-CNN 在不同主干网络下表现无明显差异，可视化结果上不作冗余展示。图像检测的部分可视化结果如图 10 所示。

从可视化角度来说，最终训练好的模型都能够检测出较为明显的锈蚀。对比来看，Faster R-CNN 的检测结果要优于 YOLOv3 的检测结果。这是因为 Faster R-CNN 中相对更加复杂的网络结构能够更深度地提取出锈蚀的图像特征，而且双阶法能够提案更多有意义的候选框。

3. 4. 2 实验数据及其分析

本文为了更具体地表现单阶法(以 YOLOv3 为例)和双阶法(以Faster R-CNN 为例)的区别及不

同改进方法对模型最终表现的影响，设计了消融实验，通过实验结果的对比，验证电力设备锈蚀检测技术的可行性和多场景应用的适应性。

如图 11 的损失曲线所示，以 Faster R-CNN 和 YOLOv3 的损失曲线为参照，模型基本在迭代数 64 500 左右收敛到局部最小值，故以下的消融实验中默认选用迭代数为 65 000 的模型进行测试。

本文从整套消融实验中选取了具有代表性的 9 组实验，表 1 为这 9 组实验模型的最终测试结果。其中，第 8 组和第 9 组中的“Darknet*”代表对原 YOLOv3 网络结构的修改版本，包括多尺度参数的调整、网络增加下采样层等。

下面详细分析表格中所展示的消融实验结果。首先，最明显的一点是 YOLO 系列的速度比 R-CNN 系列快得多，这是因为 R-CNN 系列在使用 RPN 产生候选框时消耗了大量的计算时间。本文的实验平台采用的是高性能 GPU 计算平台，如果在中央处理器(central processing unit, CPU)平台上运行网络模型，则两者的差距会更明显。同时，更多的计算消耗带来的是更精准的预测结果，从测试结果数据上来看，R-CNN 系列比 YOLO 系列准确度高约 10%。

接着，从 R-CNN 系列来看，模型性能基本会随着各个模块的优化而得到提升，召回率呈现增长趋势。由第 4 组和 5 组分别与第 3 组的数据对比可知，FPN 相比于 STN 的改进效果更好。这个现象可以从以下 2 个方面进行解释：①FPN 产生的多尺度特征比 RPN 产生的多尺度 Anchor 更具有价值，尤其是在电力设备锈蚀这种形状颜色不规则的使用场景下，优势更为明显；②在数据预处理阶段已经对数据集进行了空间变换的增广，这样就使得 STN 的优势难以体现出来。召回率的增长基本符合理论预期，但是对比第 5 组和第 6 组的实验数据可知，召回率提高，而精准率却略微降低。这一现象主要是因为在对网络结构改进时，为了追求检测小物体而对 PRN 产生候选框的惩罚系数设置不合理，也就是侧重于 Anchor 的产生而不重视边界框回归或分类失败的损失。

然后，从 YOLO 系列来看，模型性能整体提升的幅度并不明显。值得注意的是，STN 在 YOLO 下的优化效果要稍好于在 R-CNN 下的提升，这很大程度上取决于两者在特征金字塔模型设计上的差异。此外，Faster R-CNN 下的 RPN 是一个很成熟的网络，这也限制了 STN 提升的空间。

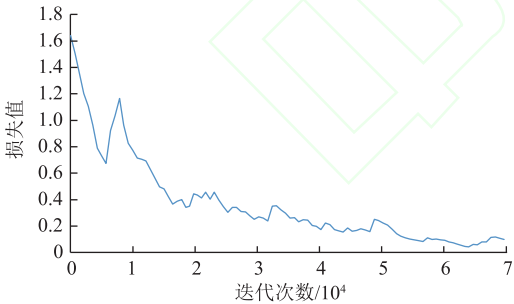


(a)YOLOv3的检测结果

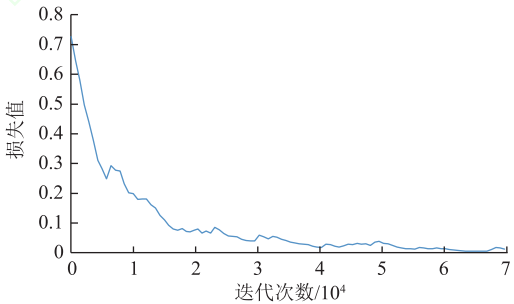


(b)Faster R-CNN的检测结果

图 10 图片检测结果
Fig. 10 Image detection results



(a)Faster R-CNN损失曲线



(b)YOLOv3损失曲线

图 11 模型训练损失曲线
Fig. 11 Training loss curves of models

表 1 模型测试结果
Tab. 1 Model test results

实验组号	网络模型	精准率	召回率	单张检测速度/s
1	Faster R-CNN + VGG16	0.660 3	0.732 3	0.049
2	Faster R-CNN + ResNet101	0.673 1	0.740 5	0.085
3	Faster R-CNN + ResNeXt	0.679 3	0.750 4	0.087
4	Faster R-CNN + ResNeXt + STN	0.682 2	0.756 2	0.088
5	Faster R-CNN + ResNeXt + FPN	0.706 6	0.774 4	0.086
6	Faster R-CNN + ResNeXt + STN + FPN	0.699 7	0.802 9	0.089
7	YOLOv3 + DarkNet	0.597 7	0.682 2	0.009
8	YOLOv3 + DarkNet *	0.615 8	0.708 4	0.010
9	YOLOv3 + DarkNet * + STN	0.621 1	0.711 7	0.010

3.4.3 实验结果总结

总的来说, 单阶法、双阶法这 2 种目标检测方法都可以有效地检测出电力设备的锈蚀。2 种方法的特点也非常明显: 双阶法精度更高, 模型训练时间快, 检测速度稍慢; 单阶法精度稍低, 模型训练时间慢, 检测速度快。通常情况下, 对于那些配有高性能服务器的大型数据中心, 对检测的实时性要求较低, 则可以选用双阶法的目标检测算法模型。此外, 由于拥有的海量数据可以及时更新模型, 双阶法对模型训练也足够友好, 训练成本低。而对于那些类似于无人机实时巡检任务的场景, 采用单阶法的目标检测算法则是更好的选择。在考虑模型实际应用时, 首要考虑的是在高速度和高精度之间的优先选择, 如果没有明确界限, 则需要再考虑模型后期升级成本或者模型运行的计算硬件条件, 综合各方面因素选用较为合适的方案。

4 结束语

本文将深度学习中的目标检测技术应用到电力设备锈蚀检测中, 并分别以 Faster R-CNN 和 YOLOv3 模型为基础模型进行实例验证。结果表明这 2 种模型都可以检测并定位较为明显的锈蚀, 而且从实验数据可以看出这 2 种深度学习模型有着各自的特点: Faster R-CNN 在准确度和小尺寸的锈蚀检测方面有着更好的表现; YOLOv3 在检测速度上略胜一筹。实际应用时, 可根据不同使用场景的需求选用不同的模型。该技术的应用可以很好地提升电力系统监控的自动化、信息化水平。

下一步的研究重心主要有以下几点: ①增加数据集的样本数, 初期考虑到数据资料保密性等原因, 能收集的图像数据有限, 而数据集规模的增大通常能够对模型的鲁棒性和精准度有所促进; ②增强对于小尺寸目标的检测能力, 加强小目标的特征表达, 优化模型的损失函数(如交叉熵损失^[38]); ③裁剪模型, 尽可能在不太影响预测结果的前提下减少网络层数和参数, 减小执行预测任务时的计算量, 实现边缘设备的模型部署^[39-40]。

参考文献:

[1] 钱金菊, 王柯, 王锐, 等. 变电站智能机器人巡检任务规划[J]. 广东电力, 2017, 30(2): 143-149.
QIAN Jinju, WANG Ke, WANG Rui, et al. Inspection mission planning for substation intelligent robot [J]. Guangdong Electric Power, 2017, 30(2): 143-149.

[2] 彭向阳, 王柯, 肖祥, 等. 大型无人直升机电力线路智能巡检宽带卫星通信系统[J]. 高电压技术, 2019, 45(2): 368-376.

PENG Xiangyang, WANG Ke, XIAO Xiang, et al. Broadband satellite communication system in the intelligent inspection of electric power line base on large scale unmanned helicopter[J]. High Voltage Engineering, 2019, 45(2): 368-376.

[3] 邵瑰玮, 刘壮, 付晶, 等. 架空输电线路无人机巡检技术研究进展[J]. 高电压技术, 2020, 46(1): 14-22.
SHAO Guiwei, LIU Zhuang, FU Jing, et al. Research progress in unmanned aerial vehicle inspection technology on overhead transmission lines[J]. High Voltage Engineering, 2020, 46(1): 14-22.

[4] 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 Wanguo, TIAN Bing, LIU Yue, et al. Study on the electrical devices detection in UAV images based on region based convolutional neural networks [J]. Journal of Geo-Information Science, 2017, 19(2): 256-263.

[5] 张春风, 徐文峰, 刘进峰. 基于机器视觉的变电站监控视频数据结构化方法研究[J]. 电力信息与通信技术, 2018, 16(5): 18-23.
ZHANG Chunfeng, XU Wenfeng, LIU Jinfeng. Research on surveillance video data structuring method based on computer vision in substation[J]. Electric Power ICT, 2018, 16(5): 18-23.

[6] 吴晖, 李铭钧, 杨英仪, 等. 面向变电站视频监控终端的目标检测方法及其优化[J]. 广东电力, 2019, 32(9): 62-68.
WU Hui, LI Mingjun, YANG Yingyi, et al. An object detection method and its optimization for substation video surveillance terminals[J]. Guangdong Electric Power, 2019, 32(9): 62-68.

[7] YANG Y, NIE M, MAI X, et al. An image recognition method of substation switch state based on robot vision[C]// 2018 International Conference on Power System Technology (POWERCON). Piscataway: IEEE Press, 2018: 4190-4195.

[8] 李旭冬, 叶茂, 李涛. 基于卷积神经网络的目标检测研究综述[J]. 计算机应用研究, 2017, 34(10): 2881-2886, 2891.
LI Xudong, YE Mao, LI Tao. Review of object detection based on convolutional neural networks [J]. Application Research of Computers, 2017, 34(10): 2882-2886, 2891.

[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 detections [J]. ACTA Automatica Sinica, 2017, 43(8): 1289-1305.

[10] 郭梦浩, 徐红伟. 基于 Faster RCNN 的红外热图像热斑缺陷检测研究[J]. 计算机系统应用, 2019, 28(11): 265-270.
GUO Menghao, XU Hongwei. Hot spot defect detection based on infrared thermal image and Faster RCNN [J]. Computer System & Applications, 2019, 28(11): 265-270.

[11] 卢树杰, 陈庆诚, 文祥, 等. 基于无人机的钢结构锈蚀视觉

- 检测[J]. 工业控制计算机, 2019, 32(7): 79-81, 87.
- LU Shujie, CHEN Qingcheng, WEN Xiang, et al. Visual inspection of steel structure corrosion based on UAV[J]. Industrial Control Computer, 2019, 32(7): 79-81, 87.
- [12] 潘卫国, 陈英昊, 刘博, 等. 基于 Faster-RCNN 的交通信号灯检测与识别[J]. 传感器与微系统, 2019, 38(9): 147-149, 160.
- PAN Weiguo, CHEN Yinghao, LIU Bo, et al. Traffic light detection and recognition based on Faster-RCNN [J]. Transducer and Microsystem Technologies, 2019, 38(9): 147-149, 160.
- [13] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[C]//Advances in Neural Information Processing Systems. Nevada, United States: NeurIPS, 2012: 1097-1105.
- [14] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA: IEEE, 2014: 580-587.
- [15] GIRSHICK R. Fast R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. Santiago, Chile: IEEE, 2015: 1440-1448.
- [16] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[C]//Advances in Neural Information Processing Systems. Montreal, Canada: NeurIPS, 2015: 91-99.
- [17] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016: 779-788.
- [18] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector [C]//European Conference on Computer Vision. Amsterdam, The Netherlands: Springer, Cham, 2016: 21-37.
- [19] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C]//Proceedings of the IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017: 2980-2988.
- [20] 杨可林, 许永盛, 李鹏. 深度学习在输电通道危物辨识技术中的应用[J]. 广东电力, 2019, 32(9): 49-55.
- YANG Kelin, XU Yongsheng, LI Peng. Application of deep learning in identification technology for dangerous objects in power transmission channels [J]. Guangdong Electric Power, 2019, 32(9): 49-55.
- [21] 张宇航, 邱才明, 杨帆, 等. 深度学习在电网图像数据及时空数据中的应用综述[J]. 电网技术, 2019, 43(6): 1865-1873.
- ZHANG Yuhang, QIU Caiming, YANG Fan, et al. Overview of application of deep learning with image data and spatio-temporal data of power grid [J]. Power System Technology, 2019, 43(6): 1865-1873.
- [22] 周秀, 朱洪波, 马云龙, 等. 基于深度学习的变压器局部放电模式识别研究[J]. 高压电器, 2019, 55(12): 98-105.
- ZHOU Xiu, ZHU Hongbo, MA Yunlong, et al. Partial discharge pattern recognition of transformer based on deep learning [J]. High Voltage Apparatus, 2019, 55(12): 98-105.
- [23] 林刚, 王波, 彭辉, 等. 基于改进 Faster-RCNN 的输电线巡检图像多目标检测及定位[J]. 电力自动化设备, 2019, 39(5): 213-218.
- LIN Gang, WANG Bo, PENG Hui, et al. Multi-target detection and location of transmission line inspection image based on improved Faster-RCNN [J]. Electric Power Automation Equipment, 2019, 39(5): 213-218.
- [24] 王立军, 张拓, 刘光伟, 等. 基于机器视觉技术的高压断路器机械特性诊断[J]. 高电压技术, 2020, 46(6): 2148-2154.
- WANG Lijun, ZHANG Tuo, LIU Guangwei, et al. Diagnostics on mechanical characteristics of high voltage circuit breaker based on machine vision technology[J]. High Voltage Engineering, 2020, 46(6): 2148-2154.
- [25] 汤踊, 韩军, 魏文力, 等. 深度学习在输电线路中部件识别与缺陷检测的研究[J]. 电子测量技术, 2018, 41(6): 60-65.
- TANG Yong, HAN Jun, WEI Wenli, et al. Research on part recognition and defect detection of transmission line indeep learning[J]. Electronic Measurement Technology, 2018, 41(6): 60-65.
- [26] 马静怡, 崔昊杨. 基于改进 RPN 网络的电力设备图像识别方法研究[J]. 供用电, 2020, 37(1): 57-61, 72.
- MA Jingyi, CUI Haoyang. Research on image recognition of power equipment based on improved RPN network algorithm[J]. Distribution & Utilization, 2020, 37(1): 57-61, 72.
- [27] 王超洋, 樊绍胜, 刘铮, 等. 基于改进 Faster RCNN 的配网架空线路异常状态检测[J]. 电力学报, 2019, 34(4): 322-329.
- WANG Chaoyang, FAN Shaosheng, LIU Zheng, et al. Abnormal state detection for distribution network based on improved Faster RCNN [J]. Journal of Electric Power, 2019, 34(4): 322-329.
- [28] 王旭红, 李浩, 樊绍胜, 等. 基于改进 SSD 的电力设备红外图像异常自动检测方法[J]. 电工技术学报, 2020, 35(增刊1): 302-310.
- WANG Xuhong, LI Hao, FAN Shaosheng, et al. Infrared image anomaly automatic detection method for power equipment based on improved single shot multi box detection [J]. Transactions of China Electrotechnical Society, 2020, 35(S1): 302-310.
- [29] 张迪, 樊绍胜. 基于 YOLO V3 的输电线路故障检测方法[J]. 自动化技术与应用, 2019, 38(7): 125-129.

ZHANG Di, FAN Shaosheng. Fault detection of transmission line based on YOLO V3[J]. Instrumentation and Measurement, 2019, 38(7): 125-129.

[30] 薛冰. 基于 Mask R-CNN 的电力设备锈迹检测[J]. 计算机系统应用, 2019, 28(5): 248-251.

XUE Bing. Rust detection of power equipment based on Mask R-CNN[J]. Computer Systems & Applications, 2019, 28(5): 248-251.

[31] 张重远, 罗世豪, 岳浩天, 等. 基于 Mel 时频谱-卷积神经网络的变压器铁芯声纹模式识别方法[J]. 高电压技术, 2020, 46(2): 413-422.

ZHANG Zhongyuan, LUO Shihao, YUE Haotian, et al. Pattern recognition of acoustic signals of transformer core based on Mel-spectrum and CNN [J]. High Voltage Engineering, 46(2): 413-422.

[32] UIJLINGS J R R, VAN DE SANDE K E A, GEVERS T, et al. Selective search for object recognition[J]. International Journal of Computer Vision, 2013, 104(2): 154-171.

[33] DAI J, LI Y, HE K, et al. R-FCN: object detection via region-based fully convolutional networks[C]//Advances in Neural Information Processing Systems. Barcelona, Spain: NeurIPS, 2016: 379-387.

[34] JADERBERG M, SIMONYAN K, ZISSERMAN A. Spatial transformer networks[C]//Advances in Neural Information Processing Systems. Montreal, Canada: NeurIPS, 2015: 2017-2025.

[35] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [C]//International Conference on Learning Representations. San Diego, USA: ICLR, 2015: 1556

[36] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016: 770-778.

[37] XIE S, GIRSHICK R, DOLLÁR P, et al. Aggregated residual transformations for deep neural networks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE, 2017: 1492-1500.

[38] 张福民, 魏永磊, 李占凯, 等. 基于改进交叉熵算法的概率最优潮流计算[J]. 智慧电力, 2020, 48(6): 79-84.

ZHANG Fumin, WEI Yonglei, LI Zhankai, et al. Probabilistic optimal power flow calculation based on improved cross-entropy algorithm[J]. Smart Power, 2020, 48(6): 79-84.

[39] 罗鸿轩, 肖勇, 杨劲锋, 等. 基于边缘计算与 MapReduce 的智能量测终端数据处理方法[J]. 智慧电力, 2020, 48(3): 22-29.

LUO Hongxuan, XIAO Yong, YANG Jinfeng, et al. Data processing method for smart metering terminal based on edge

computing and MapReduce[J]. Smart Power, 2020, 48(3): 22-29.

[40] 马富齐, 王波, 董旭柱, 等. 电力视觉边缘智能: 边缘计算驱动下的电力深度视觉加速技术[J]. 电网技术, 2020, 44(6): 2020-2029.

MA Fuqi, WANG Bo, DONG Xuzhu, et al. Power vision edge intelligence: power depth vision acceleration technology driven by edge computing[J]. Power System Technology, 2020, 44(6): 2020-2029.

项目简介:

申请单位 浙江大学

项目名称 视觉内容智能感知与生成方法研究

项目概述 主要研究内容为复杂视觉场景对象关系分析、场景图构建、外部知识库和视觉场景图支撑下的视觉问答、复杂视觉内容可控生成。在以上理论突破的基础上，形成一组可用于复杂视觉场景智能感知与可控内容生成的工具集与应用系统。

主要创新点 项目的研究试图将数据驱动与知识工程相结合，实现视觉内容的深入推理感知与可控生成，主要创新之处体现在以下几个方面：①拟将视觉关系当成是智能体之间的通信协议，将全局信息指导下的多物体语义特征融合与多尺度特征学习和表达相结合，进而把整个视觉场景中的视觉关系检测问题统一成一个标准的多智能体合作通信的决策问题，具有较好的创新性。②场景图不仅能够有效降低在特征表示过程中的信息丢失，还在表示形式上(例如三元组)能够自然与外部知识库相结合。项目将研究对场景图和知识库进行充分合理的融合并用于视觉问答，这也是研究的创新之处及研究重点之一。③所提出的基于线索及记忆生成对抗网络的高精度复杂视觉内容可控生成思路，试图将人类先验知识(即海量精准标注图像、图像线索)结合到当前输入(即输入文本、文本线索)，使得图像合成过程不仅仅受当前输入影响，也能调动“大脑”中存储的相关信息，同时系统具备与环境交互能力，学习经验知识并进一步更新知识表示，具有较好的创新性。

作者简介:

范海兵(1986)，男，江苏如皋人，工程师，主要从事变电运维工作，E-mail: fhbsdu@163.com。

胡锡幸(1982)，女，浙江慈溪人，高级工程师，博士，主要从事电气试验和带电检测等相关工作，E-mail: favorweal@163.com。

刘明一(1990)，男，黑龙江大庆人，工程师，硕士，主要从事变电运维及智能运检技术应用工作，E-mail: 753888298@qq.com。

肖俊(1979)，男，贵州铜仁人，教授，博士生导师，博士，研究方向主要集中在跨媒体智能、机器学习和计算机视觉领域，E-mail: junx@cs.zju.edu.cn。

(编辑 彭艳)