水电水利
Hydropower and Water Resources
第 9 卷◆第 2 期◆版本 1.0◆2025 年
文章类型：论文|刊号（ISSN）：2529-7821 /（中图刊号）：868GL002

基于深度学习的架空输电线路监拍图像识别

杨学航
国网河北省电力有限公司超高压分公司
DOI:10.12238/hwr.v9i2.6062

[摘　要] 为解决输电线路通道的可视化系统日抓拍图片分辨率低,图像自然环境背景复杂,图片中小目标像素占比小,信息少,线路上搭挂的小目标受天气及光照干扰大,难识别等问题。本文提出了一种基于RFBS-SRGAN(Receptive Field Block and Spectral Normalization-SRGAN)和DH-YOLOv5(Decoupled Head-YOLOv5)的深度学习架空输电线路监拍图像检测方法,该方法可以充分利用历史监拍图像数据集,通过训练学习到图像中目标物体的特征,且经过优化后的模型不仅能提高图像中目标物的识别精度,还能提高识别速度,对提高超/特高压线路可视化通道运维具有重要意义。
[关键词] RFBS-SRGAN；DH-YOLOv5；输电线路图像识别
中图分类号：TM726　文献标识码：A

Monitoring image recognition of overhead transmission lines based on deep learning

Xuehang Yang
State Grid Hebei Electric Power Co., Ltd. Ultra High Voltage Branch

[Abstract] In order to solve the problems of the visualization system of transmission line channels, such as low resolution of daily captured pictures, complex background of natural environment of images, small pixel proportion of small targets in pictures, little information, and large interference of weather and light on the small targets on the line, which is difficult to identify. This paper proposes a novel approach based on RFBS-SRGAN (Receptive Field Block and Spectral Normalization-SRGAN) and DH-YOLOv5 (Decoupled) Head-YOLOv5 deep learning monitoring image detection method for overhead transmission lines. This method can make full use of historical monitoring image data set and learn the features of target objects in images through training. Moreover, the optimized model can not only improve the recognition accuracy of target objects in images, but also improve the recognition speed. It is of great significance to improve the visual channel operation and maintenance of ultra/UHV lines.
[Key words] RFBS-SRGAN; DH-YOLOv5; transmission line image recognition

引言

我国拥有广袤的陆地面积,地形地貌复杂多变,超特高压输电线路不可避免的要建立在高山大陵,跨越森林湖泊。为降低人工巡视线路的人身风险,缩短线路巡检周期,提高线路巡视效率,国网公司落地了输电线路通道的可视化系统,实现了全天候监控各个区域,逐步由“人防”转换到“技防”。

1 输电线路通道监拍图像概述

输电线路通道的可视化系统每隔10分钟抓拍一张输电线路走廊照片,每天会产生海量的图片,这些图片通过后台传输到监控坐席,需要坐席人员判定隐患类型及隐患等级,造成人员劳动强度高,而且,输电线路通道可视化系统的抓拍照片,会出现分辨率过低、背景复杂多变、目标种类多样、噪声干扰等问题,人工容易出现误识别,导致错报,甚至漏报,无法提供精准有效的预警信息。这就需要图像智能分析技术辅助人工判别,提升系统的精准预警能力,指导运维人员差异化巡视,提高工作人员的利用率,从而实现输电线路运维管理水平的提高。

传统的超分辨率算法主要依靠基本的数字图像处理技术进行复原低分辨率图像中的高频信息。基于插值的超分辨率[1,2]重建如最近邻插值法、双线性插值法和双立方插值法等,重建过程中,仅根据一个事先定义的转换函数来计算超分辨率图像,不考虑图像的降质退化模型,往往会导致复原出的图像出现模糊、锯齿等现象。基于退化模型[3,4]的超分辨率重建迭代反投影法、凸集投影法和最大后验概率法等,通过人为定义先验知识和观测模型对其成像过程进行全方位的建模,随着放大倍数的增大,超分辨率重建细节特征也越来越模糊,图像复原效果,无论从算

Copyright © This work is licensed under a Commons Attibution-Non Commercial 4.0 International License.

文章类型：论文|刊号（ISSN）：2529-7821 /（中图刊号）：868GL002

法精准度方面,还是从人眼视觉方面,传统方法离高清化的超分辨率重建目标还有不小的距离。

2 构建RFBS-SRGAN

本文提出一种基于RFBS-SRGAN[5]和DH-YOLOv5[6]的深度学习架空输电线路监拍图像检测识别方法。该方法总体上为两个步骤,首先通过RFBS-SRGAN把输电线路通道的可视化系统中所抓拍的照片进行重构,使其分辨率得到提高,图片中线上搭挂物等小目标的纹理特征和细节特征就会更明显,通过提升小目标的分辨率使得小目标的特征表示更明显,更容易被目标检测模型DH-YOLOv5感知到；然后将高分辨率的图片输入DH-YOLOv5网络,对具有复杂自然背景及噪声干扰的图片中的各种目标进行定位和分类。最后采用目标检测模型中常用的平均精度均值mAP来验证所提模型的目标检测精度,采用FPS(Frame Per Second)指标验证其目标检测速度。

2.1建立训练数据集及测试数据集

在2020年1月1日至12月31日时间段内,从某省输电线路通道的可视化系统随机选取抓拍照片1500张照片作为训练集,然后从中任意选取500张照片作为测试集。根据输电线路通道隐患识别目标将训练样本中的目标分为吊车、钩机等大型机械和风筝、塑料布、防尘网等小目标物体,通过人工手动的方法按照不同的场景、天气、光照等进行样本标记,准备训练。这样训练出来的模型鲁棒性会更高。

2.2构建RFBS-SRGAN训练模型

由于照片中会包含建筑物、树木、施工机械、线上异物等不同形状不同大小的物体,会出现线上搭挂物等小目标难识别等问题,因此本文提出了RFBS-SRGAN。

2.2.1构建生成器

为了有效获得图中多尺度物体精细的细节和纹理,增强物体特征的识别性能,通过在SRGAN中引入感受野模块Receptive Field Block (RFB)生成高分辨率图片,感知驱动的方法有助于提升图像的视觉效果,并且为了能更好的获得细节特征,快速准确识别图片中的小目标,同时降低模型的计算复杂度,在感受野模块中采用小尺度卷积,恢复高频细节同时保持内容一致性。本文所提RFBS-SRGAN的生成器如上图所示。该网络结构主要由五部分组成,

特征提取模块由内核为3×3大小的卷积层组成,原始低分辨率图片 I_{LR} 输入 f_{conv} 卷积后输出为式(1)：

$$x_{conv} = f_{conv}(I_{LR}) \tag{1}$$

深度残差可以进行高效的特征提取,削弱图像中的噪点。Trunk-A模块由20个残差密集块级联组成,每个模块由6层卷积组成,该模块中的第 n 个残差密集块定义为 f^{n}_{RRDB} ,式(2)是Trunk-A模块的输出。

$$x_{RRDB} = f^{n}_{RRDB}(f^{n-1}_{RRDB}(\cdots f^{0}_{RRDB}(x_{conv}))\cdots) \tag{2}$$

Trunk-RFB模块由不同尺度的卷积滤波器组成,恢复丰富的图像细节,Trunk-RFB模块中第 m 个感受野残差密集块定义为 f^{m}_{RRFB} ,式(3)是Trunk-RFB模块的输出。Trunk-RFB由10个Residual of Receptive Field Dense Blocks(RRFDB)堆叠而成,每个RRFDB包含6个RFB

$$x_{RRFDB} = f^{m}_{RRFDB}(f^{m-1}_{RRFDB}(\cdots f^{0}_{RRFDB}(x_{RRDB}))\cdots) \tag{3}$$

Trunk-RFB模块的输出 x_{RRFDB} 反馈到单个的感受野模块和上采样模块,为了感受输电线路通道可视化监拍图片的不同空间背景,不同像素深度信息细节,交替使用最近邻插值和子像素卷积方式进行上采样。

上采样模块的输出 x_s 可以表示为公式(4)。其中, f_{RFB} 表示RFB函数, f_{inter} 表示最近邻域插值函数, f_{sub} 表示子像素卷积的函数。

$$x_s = f_{sub}(f_{inter}(f_{sub}(f_{inter}(f_{RFB}(x_{RRFDB}))))) \tag{4}$$

最后的两个卷积层的内核大小为3×3。 f_{c1} 和 f_{c2} 表示最后两个卷积层的函数,最终的超分辨率输出结果 I_{SR} 可以用公式(5)

$$I_{SR} = f_{c2}(f_{c1}(x_s)) \tag{5}$$

为了更好的重建输电线路通道可视化系统抓拍图中小目标的图像,本文的感受野模块,将原来卷积层3×3尺寸的大内核,替换为多个小核1×1、1×3、3×1分支池组合,根据输入图像中各种尺寸目标,应用不同的尺寸的卷积多分支池内核进行图像偏移量控制,经过感受野模块不仅能够提取细粒度的多尺度特征,尤其是线条特征,例如导线、地线及边缘轮廓等,还可以有效降低参数量和时间复杂度。

2.2.2判别器构建

判别器由多个由卷积层-激活函数层-谱归一化层组成的残差模块级联组成,其用Leaky Relu作为激活函数,同时,通过引入光谱正则化,约束各个残差模块的权重矩阵,能有效解决对抗网络训练不稳定,锐化过度及伪影问题,最后通过全连接层并使用Sigmoid激活函数,获得输入图像的识别概率。

令判别器的输出为式(6)

Copyright © This work is licensed under a Commons Attibution-Non Commercial 4.0 International License.

文章类型：论文|刊号（ISSN）：2529-7821 /（中图刊号）：868GL002

$$D(x)=Wx+b \tag{6}$$

其中W为对抗网络的参数矩阵，b为偏置项，在邻域δ内，当输出满足式(7)时：

$$\frac{\|D(x+\delta)-D(x)\|}{\|\delta\|}=\frac{\|W\delta\|}{\|\delta\|}\le\sigma(W) \tag{7}$$

σ（W）即W的谱归一化。

2.2.3定义损失函数

用预训练的VGG模型从图像的特征图谱中对生成图像I^{SR}和真实图像I^{HR}进行特征提取，R为训练数据集，网络的对抗性损失为L_{adv}，控制网络训练稳定性，像素损失为L_{pix}，避免过度生成高频细节。

I^{LR}表示输入的低分辨率图像。像素损失L_{pix}定义为生成图像I^{SR}和真实图像I^{HR}之间的曼哈顿距离，如等式(8)所示：

$$L_{pix}=\sum_{R}\left\|I^{SR},I^{HR}\right\| \tag{8}$$

VGG损失是生成图像I^{SR}的特征图谱与真实图像I^{HR}特征图谱之间的曼哈顿距离，如等式(9)所示：

$$L_{VGG}=\sum_{R}\left\|VGG_{conv36}(I^{SR}),VGG_{conv36}(I^{HR})\right\| \tag{9}$$

其中36表示预训练VGG模型中第36层的特征图。

生成对抗网络的对抗性训练损失为：

$$L_{adv}=-E\left[\log(1-\Delta_{\mathrm{Re}al})\right]-E\left[\log(\Delta_{Fake})\right] \tag{10}$$

其中$\Delta_{\mathrm{Re}al}$与Δ_{Fake}衡量生成图像I^{SR}和真实图像I^{HR}之间的差异性：

$$\begin{aligned}\Delta_{\mathrm{Re}al}&=\sigma(D(I^{HR})-E(D(I^{SR})))\\ \Delta_{Fake}&=\sigma(D(I^{SR})-E(D(I^{HR})))\end{aligned} \tag{11}$$

σ为sigmoid函数，函数E表示训练过程中每个小批量中所有数据取平均，即期望。

综上，所提模型的生成器损失函数如下，其中λ与η分别为像素损失与对抗损失系数，用来平衡各个损失项：

$$L_G=\lambda L_{pix}+L_{VGG}+\eta L_{adv} \tag{12}$$

判别器损失定义为：

$$L_D=L_{\mathrm{Re}al}+L_{Fake} \tag{13}$$

其中$L_{\mathrm{Re}al}$使得真实图像与生成图像尽可能相似，L_{Fake}生成图像与真实图像尽可能不相似。

$$L_{Fake}=-E\left[1-\log(\Delta_{Fake})\right] \tag{14}$$

$$L_{\mathrm{Re}al}=-E\left[\log(\Delta_{\mathrm{Re}al})\right] \tag{15}$$

3 构建DH-YOLOv5模型

输入端在原Mosaic增强[7]的基础上融合了Mixup增强[8]，在达到最后训练结束的20代之前使用两种增强方式，对照片做色彩，翻转，缩放，随机两张样本按比例混合等图像增强操作，进一步提高模型对输电线路可视化系统图片中小目标的识别能力。Backbone在DH-YOLOv5的Darknet53基础上，增加了SPP金字塔池化组件。主要用于将全卷机以后的数据，使用最大值池化，不管输入尺寸多少，都压缩到一个固定的尺寸的输出层，Neck层由自上而下的特征金字塔结构FPN上采样和自下而上的金字塔注意力网络PAN 下采样结构组成，进行多尺寸预测，大幅度增强了检测精度，特别是提高小目标的检测水平，Prediction层加入了解耦头，提高识别精度，加快网络收敛速度，并对损失函数进行重新定义。DH-YOLOv5检测器在3个不同尺度的特征图上来检测物体，其中下采样倍数分别为8倍、16倍和32倍，分别适用小、中大物体的检测。

$$\begin{aligned}L(object)=&\lambda_{coord}\sqrt{\sum_{i=0}^{K\times K}\sum_{j=0}^{M}I_{ij}^{obj}\left[(x_i-\hat{x}_i)^2+(y_i-\hat{y}_i)^2\right]}\\&+\lambda_{coord}\sqrt{\sum_{i=0}^{K\times K}\sum_{j=0}^{M}I_{ij}^{obj}\left[(\frac{w_i-\hat{w}_i}{\hat{w}_i})^2+(\frac{h_i-\hat{h}_i}{\hat{h}_i})^2\right]}\\&+\sum_{i=0}^{K\times K}\sum_{j=0}^{M}I_{ij}^{obj}\left[\hat{C}_i\log(C_i)+(1-\hat{C}_i)\log(1-C_i)\right]\\&+\lambda_{noobj}\sum_{i=0}^{K\times K}\sum_{j=0}^{M}I_{ij}^{noobj}\left[\hat{C}_i\log(C_i)+(1-\hat{C}_i)\log(1-C_i)\right]\\&+\sum_{i=0}^{K\times K}I_{ij}^{obj}\sum_{c\in class}\left[\hat{p}_i(c)\log(p_i(c))+(1-\hat{p}_i(c))\log(1-p_i(c))\right]\end{aligned} \tag{16}$$

所提DH-YOLOv5模型损失函数由定位分支损失、类别分支损失及目标判别损失三部分构成。定位分支损失采用预测框与真实框的欧式距离损失函数训练框的位置与大小信息，类别分支

Copyright © This work is licensed under a Commons Attibution-Non Commercial 4.0 International License.

与目标存在性判别分支采用每个目标单独进行二分类交叉熵损失函数训练。

式中，K 为边框的尺寸，M 为边框数量。I_{ij}^{obj} 和 I_{ij}^{noobj} 表示第 i 个网格的第 j 个边框是否为目标检测对象。(x_i, y_i, w_i, h_i) 和 $(\hat{x}_i, \hat{y}_i, \hat{w}_i, \hat{h}_i)$ 为所预测目标框的位置与大小信息坐标和真实目标框的坐标。C_i 和 $\hat{C}_i$ 为第 i 个框存在物体的预测置信度和真实置信度。$p_i(c)$ 和 $\hat{p}_i(c)$ 为物体属于 c 类的预测概率和真实概率，λ_{noobj} 为目标物体存在性判别损失系数，λ_{coord} 为目标框判别坐标损失系数。

4 实验与分析

经过以上模型的构建，分别对监拍图像中的吊车、钩机等大型机械，线上搭挂风筝、塑料布、防尘网等小目标异物等进行模型训练后，在测试集中测试得到如下结果：

表1 不同方法平均精度均值与识别速度对比

方法	mAP(%)	FPS(帧/s)
YOLOv5	89.72	40.6
DH-YOLOv5	92.08	45.2
RFBS-SRGAN+DH-YOLOv5	96.28	58.6

由上表得出结论，经过优化的模型RFBS-SRGAN+DH-YOLOv5在平均精度均值比传统YOLOv5高6.56%，比DH-YOLOv5高4.2%，而且识别速度也比传统YOLOv5快了26.6帧，比DH-YOLOv5快18.3帧。通过以上评估指标对模型进行评估和对比，表明本文所提模型能快速且准确识别线路通道内物体，该技术能够辅助人工判别，提升系统的精准预警能力，指导运维人员差异化巡视，提高工作人员的利用率，从而实现输电线路运维管理水平的提高。

5 结语

输电线路监拍图像识别比较复杂，本文所提方法可以不仅能提高图像中目标物的识别精度，还能提高识别速度。对输电线路运检可视化精准快速预警提升，具有明显效果，对保障超/特高压电力大动脉安全稳定运行具有重要意义。未来，伴随着人工智能的发展和管理的不断创新，输电线路监拍图像识别算法的研究，进一步为我国电力系统稳定运行提供有力支持。

[参考文献]

[1]黄健,赵元元,郭苹,等.深度学习的单幅图像超分辨率重建方法综述[J].计算机工程与应用,2021,57(18):13-23.

[2]符祥,郭宝龙.图像插值技术综述[J].计算机工程与设计,2009,30(001):141-144,193.

[3]王成志.计算机智能图像识别算法的探讨[J].软件,2021,42(06):136-138.

[4]张艳青,马建红,韩颖,等.真实场景下图像超分辨率重建研究综述[J].计算机工程与应用,2023,59(8):28-40.

[5]张毅,沈可心,刘凯丽,等.基于SRGAN模型的监控图像增强应用研究[J].电脑编程技巧与维护,2024(7):144-146,173.

[6]于立君,孙超,王辉,等.航拍图像小目标检测算法设计[J].实验室研究与探索,2024,43(10):21-25.

[7]陈翠琴,范亚臣.基于改进Mosaic数据增强和特征融合的Logo检测[J].计算机测量与控制,2022,30(10):188-194,201.

[8]李明峰,邵琳钰,蔡昌利.基于混合特征提取的细粒度图像识别方法[J].南昌大学学报(工科版),2023,45(3):297-306.

作者简介：

杨学航(1993--),男,汉族,河北井陉人,硕士研究生,电力工程师,主要研究从事工作：无人机输电线路巡检、输电线路的运行与维护。

Copyright © This work is licensed under a Commons Attibution-Non Commercial 4.0 International License.