

# 基于 Gromov-Wasserstein 最优传输的输电线路小目标检测方法

索大翔,李波\*

(天津大学管理与经济学部,天津 300072)

**摘要:**针对输电线路无人机线路巡检场景中目标检测算法在处理线路缺陷、零部件缺失等小目标时性能严重下降的问题,从标签分配角度提出新的损失函数,提高小目标检测的准确性和效果。区别于传统目标检测方法,将每个目标预测框视为高斯感受野,将真实值视为高斯热图,通过计算2个高斯分布之间的距离进行标签分配;提出利用 Gromov-Wasserstein 最优传输引导模型学习,该方法可以建立在现有的检测模型之上。对多个输电线路目标检测数据集进行试验,结果表明,采用高斯感受野和最优传输的标签分配方案在输电线路巡检中的小目标检测方面具有良好的效果。

**关键词:**输电线路;小目标检测;深度学习;最优传输;标签分配

中图分类号:TP391

文献标志码:A

引用格式:索大翔,李波. 基于 Gromov-Wasserstein 最优传输的输电线路小目标检测方法[J]. 山东大学学报(工学版),2024,54(3):22-29.

SUO Daxiang, LI Bo. Transmission line small object detection based on Gromov-Wasserstein optimal transport[J]. Journal of Shandong University (Engineering Science), 2024, 54(3):22-29.

## Transmission line small object detection based on Gromov-Wasserstein optimal transport

SUO Daxiang, LI Bo\*

(College of Management and Economics, Tianjin University, Tianjin 300072, China)

**Abstract:** In order to address the issue of severe performance degradation of target detection algorithms in the scenario of unmanned aerial vehicle (UAV) line inspection for power transmission lines, specifically when dealing with small targets such as line defects and missing components, a new loss function was proposed from the perspective of label assignment to improve the accuracy and effectiveness of small target detection. Different from traditional target detection methods, each predicted bounding box was treated as a Gaussian receptive field, and the ground truth value was treated as a Gaussian heat map. Label assignment was performed by calculating the distance between two Gaussian distributions. A Gromov-Wasserstein optimal transport-guided model learning method was introduced, which could be built upon existing detection models. Experimental results on multiple power transmission line target detection datasets demonstrated that the label assignment scheme using Gaussian receptive fields and optimal transport had achieved good performance in small target detection during power transmission line inspection.

**Keywords:** transmission line; small object detection; deep learning; optimal transport; label assignment

## 0 引言

随着现代社会经济的发展,高压输电线路成为我国能源传输的主要方式之一。输电线路一旦发

生事故导致停电,会对工业生产和生活产生极大影响。因此,对输电线路进行定期巡检非常重要。当前,对输电线路的运行维护多依赖于人力巡查,以及依靠固定摄像头传回的图像进行监视。然而输电线路数量庞大,横跨地域广,人工巡检耗时耗力、

收稿日期:2023-05-19

基金项目:国家社科基金资助项目(21&ZD102);国家自然科学基金资助项目(72132007)

第一作者简介:索大翔(1986—),男,山东泰安人,博士研究生,主要研究方向为人工智能及智慧物流。E-mail:sduhit@163.com

\*通信作者简介:李波(1965—),女,山西长治人,教授,博士生导师,博士,主要研究方向为供应链管理。E-mail:libo0410@tju.edu.cn

成本高昂且存在高空危险。因此,最近几年借助无人机航拍并通过计算机视觉方法对高压输电线路进行检测成为研究热点<sup>[1-2]</sup>。

输电线路目标识别属于图像检测技术。目前的图像检测技术在输电线路的缺陷识别中已经得到广泛研究。一些早期工作利用传统机器学习算法(如支持向量机)实现缺陷识别。文献[3]建立了基于电力线位置的滑动窗口检测方法,用于输电线路防振锤的检测和识别。在深度学习时代,在大数据和人工智能算法推动下,基于卷积神经网络的目标检测算法实现了高精度和高效的推理,比如快速区域卷积神经网络(faster region based convolutional neural networks, Faster R-CNN)、YOLO(you only look once)算法系列以及检测变换器(detection transformers, DETR)系列。因此,大量研究者尝试将目标检测的先进算法应用到输电线路目标检测中。

然而,输电线路除了跟通用目标检测存在域差别外,还广泛存在小目标问题。输电线路输电电压高,无人机拍摄时需要离目标一定距离,导致很多目标过小(小于16像素×16像素),得到的信息较少,非常容易导致漏检和误检。

区别于传统的增强小目标策略(如样本数据增强<sup>[4]</sup>、多尺度学习<sup>[5-6]</sup>、特征增强<sup>[7]</sup>等),本研究从标签分配的角度处理输电线路小目标检测。对于小目标,当前主流方法的标签分配策略不是最优的,将进一步阻碍标签分配过程。小目标的目标框较小,不论是用交并比 $I_{ou}$ 还是中心度计算均是基于均匀分布的假设,然而目标预测往往符合高斯分布。这种分布之间的差别导致小目标很难实现样本标签分配,而大尺寸目标不受影响。

为了缓解上述问题,本研究引入一种基于高斯分布的先验,构建基于高斯感受野的标签分配策略,有利于微小物体检测;提出直接衡量高斯感受野和真值的相似度作为标签分配的度量,在度量计算中提出 Gromov-Wasserstein 最优传输作为度量函数。

## 1 相关工作

在深度学习时代,随着特征提取及流程优化的不断进步,基于卷积神经网络的目标检测的性能获得显著提升。这些目标检测算法可以大体分为一阶段目标检测算法和两阶段目标检测算法。

### 1.1 一阶段目标检测算法

一阶段目标检测算法直接预测图像中目标的位置和相关类别,不需要候选框生成及细化过程,分为基于锚点的算法<sup>[8-10]</sup>和无锚点算法<sup>[11-12]</sup>。

单次多边框检测(single shot multi-box detector, SSD)系列<sup>[8-9]</sup>和 YOLO 系列<sup>[10]</sup>是早期的主流一阶段目标检测算法。文献[1]提出用 focal loss 处理一阶段目标检测算法中样本不均衡问题。从2019年开始,无锚点算法成为一阶段目标检测算法新的研究方向。全卷积一阶段目标检测(fully convolutional one-stage object detection, FCOS)首先提出一种基于全卷积网络(fully convolutional network, FCN)的逐像素目标检测算法<sup>[11]</sup>,实现了无锚点、无提议的解决方案,提出中心度,在性能上接近甚至超过目前很多先进主流的基于锚框目标检测算法。

与此同时,一阶段目标检测算法广泛应用于电网线路目标检测。文献[5]提出一种改进的 YOLOv3 目标检测算法,缓解输电线路场景从高空拍摄图片中小目标检测精度不高的问题。此外,文献[7]将设计的新特征融合模块植入到 SSD 模型,实现对输电线路异物附着故障的识别。

### 1.2 两阶段目标检测算法

两阶段目标检测算法利用区域生成网络(region proposal network, RPN)生成固定个数的候选区域,将这些候选区域输入检测网络,用于细化检测位置和目标类别。代表性的两阶段目标检测算法包括 Faster R-CNN<sup>[13]</sup>、区域全卷积网络(region-based fully convolutional networks, R-FCN)<sup>[14]</sup>及轻量头区域卷积神经网络(light-head region based convolutional neural networks, Light-head R-CNN)<sup>[15]</sup>。

两阶段目标检测算法也广泛用于输电线路目标检测:文献[16]采用两阶段深度卷积网络,利用特征金字塔网络(feature pyramid networks, FPN)提取多尺度特征,使主干网更好地适应目标尺寸变化,并通过特征增强获得更清晰、更具有代表性的多尺度目标特征;文献[17]将级联区域卷积神经网络(cascade region based convolutional neural network, Cascade R-CNN)应用到输电线路小目标缺陷检测中,实现了比 YOLOv3 更好的性能;文献[4]对 Faster R-CNN 进行特征增强并应用到输电线路铁锈检测。

### 1.3 标签分配

在目标检测任务中,标签分配用于决定每一个锚点是前景真值还是背景,有3种常用的方法。第

1种是  $I_{ou}$  阈值的方法,即计算每个锚点生成的候选区域与真值之间的重叠程度:

$$I_{ou} = \frac{D_1 \cap D_2}{D_1 \cup D_2}, \quad (1)$$

式中,  $D_1$  为当前目标真值区域,  $D_2$  为生成的候选框。根据  $I_{ou}$  与预设的阈值  $u$  决定目标标签

$$y = \begin{cases} c_y, & I_{ou} \geq u \\ 0, & \text{其他} \end{cases}, \quad (2)$$

式中  $c_y$  为目标检测的类别。Faster R-CNN 使用 0.7 和 0.3 作为正负样本的阈值<sup>[13]</sup>, R-CNN 使用 0.5 作为正负样本的阈值<sup>[15]</sup>。

第2种是无锚点的检测算法(如 FCOS<sup>[11]</sup>),在进行标签分配时,通常将目标中心点周围的锚点作为正样本,其他看作负样本。

第3种是基于 Transformer 的目标检测算法<sup>[18-19]</sup>,采用序列化输入和输出设计全新的标签分配策略。代表性的 DETR<sup>[18]</sup> 将标签分配看作图匹配过程,利用匈牙利算法<sup>[20]</sup> 实现标签跟预测框之间的匹配。

## 2 基于最优传输的标签分配

### 2.1 基本定义

最优传输(optimal transport, OT)用于描述以下问题:假设有  $m$  个供应商和  $n$  个需求商,其中第  $i$  个供应商持有  $s_i$  个单位的商品,第  $j$  个需求商需要  $d_j$  个单位的商品。从供应商  $i$  到需求商  $j$  的每单位货物运输成本为  $c_{ij}$ 。OT 的目标是寻求最优规划

$$\tau^* = \{\tau_{i,j} | i=1,2,\dots,m, j=1,2,\dots,n\}, \quad (3)$$

式中  $\tau_{i,j}$  为第  $i$  个供应商和第  $j$  个需求商之间的运输量。求解的约束条件是供应商的所有货物都可以以最低的运输成本运输到需求商。该传输过程可以描述为以下规划函数

$$\begin{aligned} \min_{\tau} & \sum_{i=1}^m \sum_{j=1}^n c_{ij} \tau_{ij}, \\ \text{s.t.} & \sum_{i=1}^m \tau_{ij} = d_j, \quad \sum_{j=1}^n \tau_{ij} = s_i, \quad \sum_{i=1}^m s_i = \sum_{j=1}^n d_j, \end{aligned} \quad (4)$$

$$\tau_{ij} \geq 0, \quad i=1,2,\dots,m, j=1,2,\dots,n.$$

求解该问题常用的方法是快速迭代的 Sinkhorn-Knopp 方法<sup>[21]</sup>,得到的最优解称为最优传输。

### 2.2 最优传输下的标签分配

针对目标检测任务,假设图像  $I$  产生  $M$  个预测框和对应的  $N$  个真值框。标签分配的任务是根据

预测框与真值的相似度判断预测框是正样本还是负样本。该标签分配可以用最优传输进行描述,对于每一幅图像,可以将真值看作供货商,预测框看作需求商。标签分配的代价就是将每一个预测框标记成正样本或负样本,代价函数就是预测框和真值框的相似度。

无人机航拍巡检时机位的变化会造成目标拍摄角度、姿态等不同,使目标尺寸分布具有波动性,小目标在图像中出现的尺度较小且变化波动一般符合高斯分布<sup>[22]</sup>,将目标视为高斯分布能够更好地模拟真实小目标的尺度分布规律。考虑到小目标广泛存在,传统的  $I_{ou}$  或中心度不适合用于计算代价函数,原因在于传统的均匀分布先验区域对先验区域内的每个位置进行平等对待,但实际上将特征点的感受野重新映射回输入图像时,有效感受野通常是高斯分布的,这会导致物体尺度特征与分配的特征点的感受野不匹配。而高斯感受野能够更好地匹配目标的尺度特征。文献[22-23]也指出,目标检测中的有效感受野理论上是高斯分布。基于以上因素,本研究将预测的检测结果表示成高斯感受野,如图1所示。

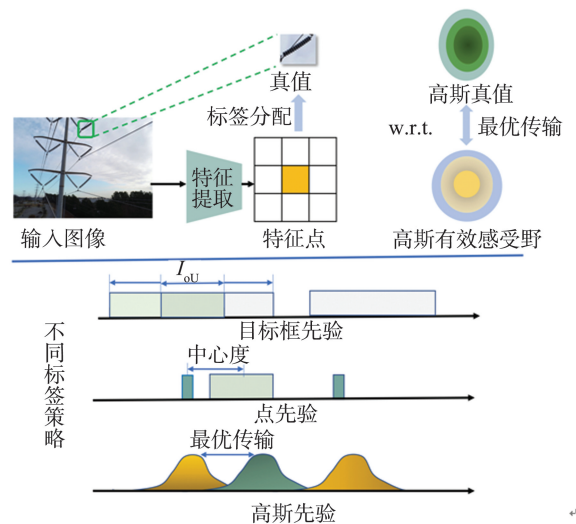


图1 利用高斯感受野进行目标表示

Fig.1 Target representation based on Gaussian receptive fields

针对特征图上的每一个点  $(x_n, y_n)$ , 都可以用一个二维高斯分布进行表达,高斯分布的均值矩阵  $\mu_e$  和协方差矩阵  $\Sigma_e$  分别为:

$$\mu_e = \begin{bmatrix} x_n \\ y_n \end{bmatrix}, \quad \Sigma_e = \begin{bmatrix} r_{en}^2 & 0 \\ 0 & r_{en}^2 \end{bmatrix}, \quad (5)$$

式中  $r_{en}^2$  为感受野半径的平方和。

得到高斯感受野后,需计算标签分配的距离,

即高斯感受野与真值之间的匹配程度。一个真值框用  $(x_g, y_g, w_g, h_g)$  表示,其中  $(x_g, y_g)$  为中心点坐标,  $w_g$  和  $h_g$  分别为真值框的宽和高,将其建模为标准二维高斯分布  $N_g(\mu_g, \Sigma_g)$ 。其中,每一个真值框的中心点作为高斯分布的均值,对角线的平方和作为协方差矩阵,分别表示为:

$$\mu_g = \begin{bmatrix} x_g \\ y_g \end{bmatrix}, \Sigma_g = \begin{bmatrix} w_g^2/4 & 0 \\ 0 & h_g^2/4 \end{bmatrix} \quad (6)$$

### 2.3 Gromov-Wasserstein 距离

区别于传统的  $I_{ou}$  或利用中心度衡量标签和预测框之间的距离,本研究采用前述的高斯分布进行描述,需要利用全新的距离函数进行度量,一种直观的度量函数是 Wasserstein 距离。给定高斯感受野  $n_e = N_e(\mu_e, \Sigma_e)$  和高斯分布的真值  $n_g = N_g(\mu_g, \Sigma_g)$ , 2个分布之间的 Wasserstein 距离

$$C_2^2(n_e, n_g) = \left\| \left( [x_n \ y_n \ r_{en} \ r_{en}]^T, [x_g \ y_g \ w_g/2 \ h_g/2]^T \right) \right\|_2^2 \quad (7)$$

Wasserstein 距离用于计算 2 个不重叠分布之间的距离,这一特性对小目标检测具有关键作用。对于小目标,训练过程中产生的预测框很难与真值框有任何重合,因此产生的高斯分布与真值分布之间重叠极少。

假设有  $m$  个预测目标框,  $n$  个真值,整个标签分配可以用最优传输建模为:

$$D_\omega(\mu, \nu) = \min_{T \in \tau(\mu, \nu)} \sum_{i=1}^m \sum_{j=1}^n T_{ij} c(h_{1i}, h_{2j}) \quad (8)$$

式中:  $\tau(\mu, \nu)$  为任意 2 个节点直接组合;  $h_{1i}, h_{2j}$  为节点特征(即目标框提取的特征); 矩阵  $T$  为传输规划;  $T_{ij}$  为  $T$  中的元素,表示第  $i$  个预测框和第  $j$  个真值框之间的传输量;  $c(h_{1i}, h_{2j})$  为预测分布节点图的节点  $i$  跟真值分布节点图的节点  $j$  之间的传输代价,  $c(h_{1i}, h_{2j}) = \exp(-\varphi(h_{1i}, h_{2j})/\lambda)$ , 其中  $\varphi$  为相似度函数,  $\lambda$  为规范化参数。

Wasserstein 距离可以用于计算检测图像中所有目标框与真值的距离。然而,很多时候输电线路中的部件排列是密集分布的,因此实际任务中,对每个目标进行标签分配时,需要考虑目标周围的上下文信息。预测目标框的距离跟真值高斯分布的距离,可以建模成 2 个图之间的距离度量。将  $m$  个预测目标框看作一个含有  $m$  个节点图  $\hat{\mu} = \{\mu_1, \mu_2, \dots, \mu_m\}$ ,  $n$  个真值集合看作另一个图  $\hat{\nu} = \{\nu_1, \nu_2, \dots, \nu_n\}$ 。

因此,在执行标签分配时,不仅需要考虑到 2 个图中两两节点之间的距离,还需要考虑到不同节点的边信息。基于此,衡量 2 个图之间的相似度可以用 Gromov-Wasserstein 距离

$$D_{gw}(\mu, \nu) = \min_{T \in \tau(\mu, \nu)} \sum_{i, i', j, j'} T_{ij} T_{i'j'} \hat{c}(h_{1i}, h_{2j}, h_{1i'}, h_{2j'}) \quad (9)$$

式中:  $\hat{c}(h_{1i}, h_{2j}, h_{1i'}, h_{2j'})$  表示不同子图之间边的传输代价度量,  $\hat{c}(h_{1i}, h_{2j}, h_{1i'}, h_{2j'}) = \|c(h_{1i}, h_{1i'}) - c(h_{2j}, h_{2j'})\|_2$ 。

### 2.4 实现细节

结合 Gromov-Wasserstein 距离的标签传输可以用现有的一阶段或两阶段目标检测模型(图 2)。对于利用  $I_{ou}$  作为标签分配的检测器(如 Faster R-CNN<sup>[13]</sup>),可以直接用 Gromov-Wasserstein 距离进行替换。对于利用中心度作为标签分配的检测器(如 FCOS<sup>[11]</sup>),可以对中心度函数进行替换,得到如下目标函数进行模型学习:

$$L_{\text{centerness}} = \sqrt{\frac{\varepsilon \left[ \frac{\min(l, r)}{\max(l, r)} \right] \min(l, r) + c}{\max(l, r)} \times \frac{\varepsilon \left[ \frac{\min(t, b)}{\max(t, b)} \right] \min(t, b) + c}{\max(t, b)}}, \quad (10)$$

式中:  $l, t, r, b$  为 FCOS 中的回归参数;  $\varepsilon(\cdot)$  为步长函数;  $c$  为一个因子,一般定义为 0.01。

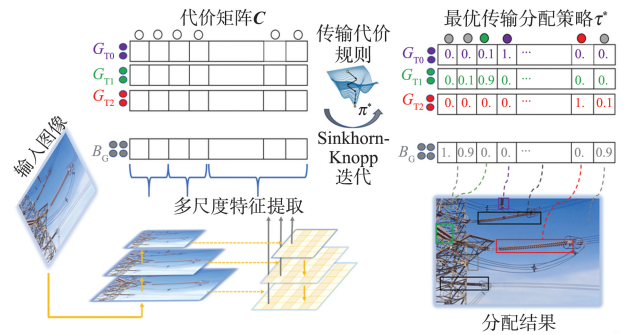


图 2 本研究技术路线  
Fig.2 The technical procedure

## 3 试验

### 3.1 试验数据集和试验环境

为验证本研究方法的有效性,试验主要在 3 个不同类型的输电线路数据集上进行,包括 TTPLA 数据集<sup>[24]</sup>、输电线路金属器具数据集<sup>[25]</sup>及绝缘子 SFID 数据集<sup>[26]</sup>。TTPLA 数据集包含输电塔和电缆 2 个类别,共 1 242 张图像、8 987 个实例,其中 1 000 张图像用于训练,剩余 242 张图像作为测试

集。输电线路金属器具数据集共 2 511 张图像,包含输电线路绝缘子、螺栓销钉、防震锤、线夹等 7 个类别,使用 1 600 张图像用于训练,600 张图像用于测试,剩余 311 张图像作为验证集。绝缘子 SFID 数据集由 13 718 张图像组成,其中训练集图像 10 975 张,测试图像 1 743 张,将测试图像按 3:1 划分为测试集和验证集,该数据集包含正常绝缘子和缺陷绝缘子 2 种类别。

所有试验均在同一 NVIDIA RTX 3090 GPU 上进行,核心代码基于 MMDetection 构建,用 ImageNet 的预训练模型作为主干,以 Faster R-CNN<sup>[13]</sup> 作为基准。优化器为随机梯度下降(stochastic gradient descent, SGD),训练 12 个轮次,动量设为 0.9,权重衰减设为 0.000 1,批尺寸设为 2,初始的学习率设为 0.005,在第 8 和第 11 个轮次时下调学习率,RPN 的目标框数量为 3 000。

### 3.2 评价指标

为了定量比较所提算法和对标算法对输电线路小目标缺陷检测的检测效果与精度,本研究采用精确率  $P$ 、召回率  $R$  和所有类别的平均准确率  $A_{mp}$  作为衡量指标,分别为:

$$P = \frac{T_p}{T_p + F_p}, \quad (11)$$

$$R = \frac{T_p}{T_p + F_N}, \quad (12)$$

$$A_{mp} = \frac{1}{C} \sum_{C=1}^C A_{PC}, \quad (13)$$

式中, $T_p$  为真实正样本正确识别成正的个数, $F_p$  为背景目标识别为正的个数, $F_N$  为真实目标识别为背景的个数, $A_{PC}$  为平均准确率, $C$  为类别个数。

### 3.3 试验结果

本研究方法与其他最优检测算法的性能比较如表 1 所示,包括一阶段目标检测算法 Yolact 和两阶段目标检测算法 Faster R-CNN、R-FCN 等。对于其他通用目标检测算法(如 Faster R-CNN),利用与文献[27-28]中相同的参数设定。如表 1 所示,本研究方法的  $A_{p50}$  和  $A_{mp}$  分别为 34.6% 和 21.4%,性能有所提升,表明本研究方法在输电线路小目标检测中具有更高的  $R$  和  $P$ 。由于输电线路数据集的样本量不大,采用更大规模的基础模型(如 ResNet101)并不能提升目标检测算法的性能,反而会因过拟合问题导致检测性能下降。

表 1 TTPLA 数据集检测结果  
Table 1 Detection results on the TTPLA dataset

模型	主干网络	$A_{p50}$	$A_{p75}$	$A_{mp}$	$A_{ps}$	$A_{pm}$	$A_{pl}$	单位:%
Yolact <sup>[24]</sup>	ResNet50	34.3	11.2	14.5	6.2	21.7	31.1	
Yolact <sup>[24]</sup>	ResNet101	32.6	10.1	14.1	7.3	20.7	26.4	
Faster R-CNN <sup>[13]</sup>	ResNet50	31.7	9.9	13.7	6.9	21.3	26.0	
R-FCN <sup>[14]</sup>	ResNet50	32.3	10.0	13.9	8.4	22.7	27.5	
本研究方法	ResNet50	34.6	21.4	21.4	18.1	22.2	27.6	
本研究方法	ResNet101	34.4	21.0	20.8	18.3	20.8	26.1	

注: $A_{p50}$  和  $A_{p75}$  分别为  $I_{o0}$  阈值取 0.50 和 0.75 时的平均精度, $A_{ps}$ 、 $A_{pm}$  和  $A_{pl}$  分别为针对小尺寸、中尺寸和大尺寸目标的检测准确率。

输电线路金属器具数据集上不同检测算法的性能比较如表 2 所示。由表 2 可以看出,在该数据集上本研究方法也实现了最优的检测结果。由于

该数据集存在一些小的螺栓销钉及线夹,本研究方法在该数据集上相比于其他算法有较大性能优势,进一步表明本研究方法的有效性和通用性。

表 2 输电线路金属器具数据集检测结果

Table 2 Detection results on dataset of metal appliances for transmission lines

模型	主干网络	$A_{p50}$	$A_{p75}$	$A_{mp}$	$A_{ps}$	$A_{pm}$	$A_{pl}$	单位:%
SSD <sup>[9]</sup>	ResNet50	22.3	20.1	18.7	-100	11.9	22.6	
Faster R-CNN <sup>[13]</sup>	ResNet50	23.1	21.5	19.6	-100	17.1	23.7	
YOLOv4 <sup>[10]</sup>	DarkNet	26.8	24.9	21.1	-100	18.2	24.4	
R-FCN <sup>[14]</sup>	ResNet50	24.5	23.7	20.1	-100	18.4	24.7	
本研究方法	ResNet50	28.4	26.2	22.0	-100	21.5	28.0	

在 SFID 数据集上的验证结果如表 3 所示,本研究方法的精确度为 99.0%, $A_{mp}$  为 86.3%。值得注意的是,本研究方法与所比较算法实现了相当的性能,其原因有 2 个:SFID 数据集的目标尺寸都比较

大,无法体现出本研究算法对小目标检测的优势;本研究算法采用的主干网络是 ResNet50 而不是更先进的 DarkNet,且没有采用 FINet<sup>[26]</sup> 中的数据增强方法。

表3 SFID数据集检测结果  
Table 3 Detection results on the SFID datasets  
单位:%

模型	主干网络	$A_{P50}$	$A_{mp(0.50\sim0.95)}$
Faster R-CNN	ResNet50	98.4	80.1
Mask R-CNN	ResNet50	98.3	82.0
YOLOX	ResNet50	99.4	86.0
Swin_transformer	Swin-T	99.0	86.4
YOLOv5	DarkNet	99.3	87.0
FINet	DarkNet	99.5	88.3
本研究方法	ResNet50	99.0	86.3

注: $A_{mp(0.50\sim0.95)}$ 为 $I_{ou}$ 阈值在0.50~0.95(步长0.05)下的平均准确率。

本研究方法在TTPLA数据集上进行消融试验的结果如表4所示,包括高斯感受野及Gromov-Wasserstein距离对最终检测性能的影响。所有消融试验的基线方法是Faster R-CNN,主干网络是ResNet50。由表4可以看出,高斯感受野的引入使本研究方法在基线方法基础上获得了性能提升, $A_{P50}$ 、 $A_{P75}$ 、 $A_{mp}$ 及 $A_{PS}$ 分别提升2.3、6.4、5.8及7.3百分点,说明使用该方法进行目标的标签表达能更有效地处理小目标的标签分配,提升目标检测性能。Gromov-Wasserstein距离的引入使本研究方法在进行标签分配学习时,充分考虑到不同目标之间的上下文关系,使密集排列的目标更好共现,更有利于小目标的检测,在基准方法的基础上,本研究方法的 $A_{P50}$ 、 $A_{P75}$ 、 $A_{mp}$ 及 $A_{PS}$ 分别提升了0.6、5.1、1.9及1.9百分点。

表4 消融试验结果  
Table 4 Results of ablation study  
单位:%

高斯感受野	Gromov-Wasserstein距离	$A_{P50}$	$A_{P75}$	$A_{mp}$	$A_{PS}$
		31.7	9.9	13.7	8.9
✓		34.0	16.3	19.5	16.2
✓	✓	34.6	21.4	21.4	18.1

注: $A_{PS}$ 为针对小尺寸目标的检测准确率,“✓”表示采用对应成分。

本研究进一步分析了高斯感受野对标签分配的影响。通过在图像上随机生成不同尺寸的真值框,模拟标签分配过程,统计Faster R-CNN<sup>[13]</sup>、FCOS<sup>[11]</sup>及本研究方法得到正样本的数量。真值尺寸范围为0~64像素,将其分成16个间隔,计算分配到每个间隔的正样本数,如表5所示,部分标签分配策略存在明显的样本不均衡问题。对于基于锚点的检测算法(Faster R-CNN),采用 $I_{ou}$ 作为分配标准,小尺寸目标很难获得正样本。无

锚点检测算法虽然在一定程度上缓解了这个问题,但对于极小目标仍然无法获得正样本,存在比较明显的样本不均衡问题。而本研究方法通过引入有效感受野的方式,获得了分布更均衡的正样本。

表5 不同算法的尺度-样本分配均衡性对比  
Table 5 Comparison of scale-sample allocation fairness for different algorithms

真值尺寸/像素	Faster R-CNN获得的正样本数/个	FCOS获得的正样本数/个	本研究方法获得的正样本数/个
2	0	0	3
6	0	1	3
10	0	3	3
14	0	5	3
18	0	8	3
22	1	8	3
26	4	8	3
30	7	7	4
34	7	8	4
38	8	8	4
42	6	8	3
46	3	7	3
50	3	7	3
54	4	7	3
58	5	7	3
62	6	8	3

### 3.4 检测结果呈现

使用新标签分配策略与普通方案的检测结果可视化对比如图3所示,其中绿色框代表 $T_p$ ,红色框代表漏检的目标。由图3可看出,本研究的检测方法在光线不足、小目标、目标长宽比悬殊等特殊情况下取得了较好的检测效果,特别是对于小目标的检测相较于普通方法具备明显优势。

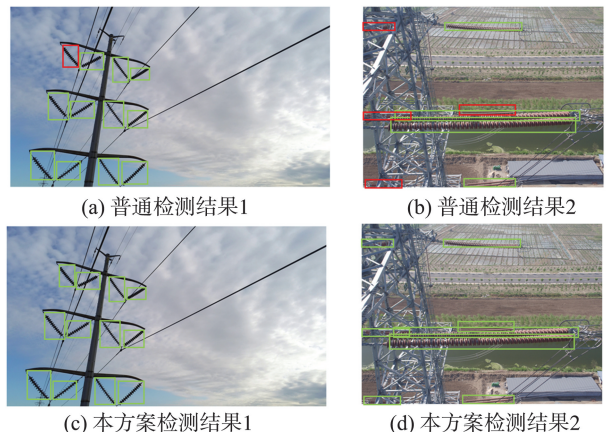


图3 检测结果可视化比较

Fig.3 Visual comparison of detection results

## 4 结论

为解决利用无人机航拍进行输电电路目标检测时分辨率低、遮挡、小目标漏检等问题,本研究从标签分配的角度重新思考小目标检测,提出基于 Gromov-Wasserstein 的全新度量和基于最优传输的标签分配方法,以提高目标检测精度。该方法能有效利用高斯分布描述样本,进而更好处理小目标标签分配时的监督信号。在后续研究中,将通过裁减主干网络、优化特征融合的方式,进一步提高对小目标的检测精度和检测实时性,促进无人机巡检核心技术的发展。

### 参考文献:

- [1] 裴伟, 许晏铭, 朱永英, 等. 改进的 SSD 航拍目标检测方法[J]. 软件学报, 2019, 30(3): 738-758.  
PEI Wei, XU Yanming, ZHU Yongying, et al. The target detection method of aerial photography images with improved SSD[J]. Journal of Software, 2019, 30(3): 738-758.
- [2] 周霞, 陈旦, 李霞. 我国供电可靠性的现状分析与展望[J]. 供用电, 2021, 38(1): 49-54.  
ZHOU Xia, CHEN Dan, LI Xia. Analysis and prospect of power supply reliability in China[J]. Distribution & Utilization, 2021, 38(1): 49-54.
- [3] WU H B, XI Y P, FANG W M, et al. Damper detection in helicopter inspection of power transmission line[C]//2014 Fourth International Conference on Instrumentation and Measurement, Computer, Communication and Control. Harbin, China: IEEE, 2014: 628-632.
- [4] GUO Z, TIAN Y, MAO W. A robust faster R-CNN model with feature enhancement for rust detection of transmission line fitting[J]. Sensors, 2022, 22: 7961-7977.
- [5] 李海峰, 武剑灵, 李立学. YOLOv3-ad 在输电线路监控中提高小目标物体检测准确度的研究[J]. 广东电力, 2021, 34(2): 92-100.  
LI Haifeng, WU Jianling, LI Lixue. Research on YOLOv3-ad in improving detection accuracy of small objects in transmission line monitoring[J]. Guangdong Electric Power, 2021, 34(2): 92-100.
- [6] 威银城, 江爱雪, 赵振兵, 等. 基于改进 SSD 模型的输电线路巡检图像金具检测方法[J]. 电测与仪表, 2019, 56(22): 7-12.  
QI Yincheng, JIANG Aixue, ZHAO Zhenbing, et al. Fittings detection method in patrol images of transmission line based on improved SSD[J]. Electrical Measurement & Instrumentation, 2019, 56(22): 7-12.
- [7] 邓天华, 赵曙光, 刘西钉. 基于改进 SSD 算法的输电线异物附着故障检测识别技术研究[J]. 智能电网, 2022, 12(2): 36-42.  
DENG Tianhua, ZHAO Shuguang, LIU Xiding. Research on detection and recognition technology of foreign object adhesion fault in transmission line based on improved SSD algorithm[J]. Smart Grid, 2022, 12(2): 36-42.
- [8] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE, 2017: 2999-3007.
- [9] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//Proceedings of the 14th European Conference on Computer Vision. Amsterdam, Netherlands: Springer, 2016: 21-37.
- [10] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA: IEEE, 2016: 779-788.
- [11] TIAN Z, SHEN C, CHEN H, et al. FCOS: fully convolutional one-stage object detection [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea: IEEE, 2019: 9626-9635.
- [12] KONG T, SUN F, LIU H, et al. Foveabox: beyond anchor-based object detection[J]. IEEE Transactions on Image Processing, 2020, 29: 7389-7398.
- [13] REN S, HE K, GIRSHICK R, 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.
- [14] DAI J, LI Y, HE K, et al. R-FCN: object detection via region-based fully convolutional networks[C]// Proceedings of the 30th International Conference on Neural Information Processing Systems. Barcelona, Spain: NIPS, 2016: 29-38.
- [15] LI Z, PENG C, YU G, et al. Light-head R-CNN: in defense of two-stage object detector [EB/OL]. (2017-11-23)[2023-4-15]. <https://arxiv.org/abs/1711.07264>.
- [16] 李辉, 董燕, 刘祥, 等. 基于两阶段深度网络的输电线路异常目标检测方法[J]. 控制与决策, 2022, 37(7): 1873-1882.  
LI Hui, DONG Yan, LIU Xiang, et al. Transmission line abnormal object detection method based on deep network of two-stage[J]. Control and Decision, 2022, 37(7): 1873-1882.
- [17] 吴军, 白梁军, 董晓虎, 等. 基于 Cascade R-CNN 算法的输电线路小目标缺陷检测方法[J]. 电网与清洁能源, 2022, 38(4): 19-27.

- WU Jun, BAI Liangjun, DONG Xiaohu, et al. Transmission line small target defect detection method based on Cascade R-CNN algorithm [J]. Power System and Clean Energy, 2022, 38(4): 19-27.
- [18] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers[C]//Proceedings of the 16th European Conference on Computer Vision. Glasgow, UK: Springer, 2020: 213-229.
- [19] ZHU X, SU W, LU L, et al. Deformable DETR: deformable transformers for end-to-end object detection [EB/OL]. (2021-03-18) [2023-04-17]. <https://arxiv.org/abs/2010.01459v1>.
- [20] GE Z, LIU S, LI Z, et al. OTA: optimal transport assignment for object detection [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN, USA: IEEE, 2021: 303-312.
- [21] CUTURI M. Sinkhorn distances: lightspeed computation of optimal transport[C]//Proceedings of the 26th International Conference on Neural Information Processing Systems. Stateline, USA: NIPS, 2013: 2292-2300.
- [22] LUO W, LI Y, URTASUN R, et al. Understanding the effective receptive field in deep convolutional neural networks[C]//Proceedings of the 30th International Conference on Neural Information Processing Systems. Barcelona, Spain: NIPS, 2016: 4905-4913.
- [23] XU C, WANG J, YANG W, et al. RFLA: Gaussian receptive field based label assignment for tiny object detection[C]//Proceedings of the 17th European Conference on Computer Vision. Tel Aviv, Israel: Springer, 2022: 526-543.
- [24] ABDEFATTAH R, WANG X, WANG S. TTPLA: an aerial-image dataset for detection and segmentation of transmission towers and power lines[C]//Proceedings of the 15th Asian Conference on Computer Vision. Kyoto, Japan: ACCV, 2020: 17-33.
- [25] 飞桨. 输电线路金具数据集[DS/OL]. (2020-12-29) [2023-02-15]. <https://aistudio.baidu.com/aistudio/datasetdetail/66310>.
- [26] ZHANG Z D, ZHANG B, LAN Z C, et al. FINet: an insulator dataset and detection benchmark based on synthetic fog and improved YOLOv5[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 6006508.
- [27] ZHAI Y, YANG X, WANG Q, et al. Hybrid knowledge R-CNN for transmission line multifitting detection [J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 5013312.
- [28] ZHAI Y, WANG Q, GUO C, et al. CGRM: a context-based graph reasoning model for fitting detection [J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 2518213.

(编辑:孙亚彤)

(上接第21页)

- [24] 徐晓, 丁世飞, 孙统风, 等. 基于网格筛选的大规模密度峰值聚类算法[J]. 计算机研究与发展, 2018, 55(11):2419-2429.
- XU Xiao, DING Shifei, SUN Tongfeng, et al. Large-scale density peaks clustering algorithm based on sparse screening[J]. Journal of Computer Research and Development, 2018, 55(11):2419-2429.
- [25] 谢娟英, 高红超, 谢维信.  $K$  近邻优化的密度峰值快速搜索聚类算法[J]. 中国科学:信息科学, 2016, 46(2):258-280.
- XIE Juanying, GAO Hongchao, XIE Weixin.  $K$ -nearest neighbors optimized clustering algorithm by fast search and finding the density peaks of a dataset [J]. Scientia Sinica Informationis, 2016, 46(2):258-280.
- [26] XU T F, JIANG J H. A graph adaptive density peaks clustering algorithm for automatic centroid selection and effective aggregation[J]. Expert Systems with Applications, 2022, 195:116539.
- [27] ZHANG Y F, CHEN S M, YU G. Efficient distributed density peaks for clustering large data sets in MapReduce [J]. IEEE Transactions on Knowledge and Data Engineering, 2016, 28(12):3218-3230.
- [28] MEHMOOD R, ZHANG G Z, BIE R F, et al. Clustering by fast search and find of density peaks via heat diffusion [J]. Neurocomputing, 2016, 208:210-217.
- [29] DING S F, LI C, XU X, et al. A sampling-based density peaks clustering algorithm for large-scale data [J]. Pattern Recognition, 2023, 136:109238.
- [30] RASOOL Z, ZHOU R, CHEN L, et al. Index-based solutions for efficient density peak clustering [J]. IEEE Transactions on Knowledge and Data Engineering, 2020, 34(5):2212-2226.
- [31] JIANG J, YAN X, YU Z T, et al. A Chinese expert disambiguation method based on semi-supervised graph clustering[J]. International Journal of Machine Learning & Cybernetics, 2015, 6(2):197-204.

(编辑:孙亚彤)