刘蓉,窦爱霞.LA-YOLO11:结合知识蒸馏技术的高精度输电塔检测算法[J].地震工程学报,2027,49(1):152-162.DOI:10.20000/j.1000-0844.20250702002

Liu Rong, Dou Aixia.LA-YOLO11: a high-precision transmission tower detection algorithm combined with knowledge distillation [J].China Earthquake Engineering Journal,2027,49(1):152-162.DOI:10.20000/j.1000-0844.20250702002

LA-YOLO11:结合知识蒸馏技术的高精度输电塔检测算法

刘 蓉,窦爱霞

(中国地震局地震预测研究所 地震预测与风险评估应急管理部重点实验室,北京 100036)

摘要:针对输电塔遥感影像检测中存在的背景复杂、目标被遮挡、成像模糊等问题,结合知识蒸馏技术,基于YOLO11算法提出一种高精度输电塔检测算法——Linear Attention-YOLO11(简称LA-YOLO11)算法。该算法在骨干网络中引入受Mamba启发的线性注意力模块,增强模型对全局特征的提取能力,以适应影像中被遮挡和成像模糊的目标。同时,在头部网络中引入基于最小点距离的交并比(IoU)损失函数,提升模型对相邻区域空间位置差异的敏感性,进一步增强模型在复杂场景中的泛化能力。在此基础上,利用掩码生成蒸馏技术,在不增加模型参数量及计算量的前提下显著提高检测精度。基于输电塔卫星影像数据集的测试结果表明,改进后算法的平均精度均值(IoU=0.5)和多阈值平均精度均值(IoU为0.5~0.95,步长0.05)分别提升了1.8%和4.7%。与经典算法相比,改进后的算法在背景复杂、目标被遮挡、成像模糊等场景中表现出更高的检测准确性。

关键词:输电塔;YOLO11;目标检测;掩码生成蒸馏

中图分类号:TP751;TM75 **文献标志码**:A **文章编号**:1000-0844(2027)01-0152-11

DOI:10.20000/j.1000-0844.20250702002

LA-YOLO11: a high-precision transmission tower detection algorithm combined with knowledge distillation

Liu Rong, Dou Aixia

(*Key Laboratory of Earthquake Forecasting and Risk Assessment, Ministry of Emergency Management, Institute of Earthquake Forecasting, CEA, Beijing 100036, China*)

Abstract: To address the challenges of complex backgrounds, target occlusion, and image blur in transmission tower detection using remote sensing imagery, this study proposes Linear Attention (LA)-YOLO11, a high-precision detection algorithm that integrates the YOLO11 framework with knowledge distillation. The algorithm integrates a Mamba-inspired linear attention module into the

收稿日期:2025-07-02

基金项目:中国地震局地震预测研究所基本科研业务费专项(CEAIEF20250105)

第一作者简介:刘 蓉(2001-),女,硕士研究生,主要从事遥感与地震灾害风险评估相关研究。E-mail:454440546@qq.com。

通信作者:窦爱霞(1977-),女,研究员,主要从事地震灾害风险评估与应急监测评估、空间技术综合减灾应用研究。E-mail:axdothy@163.com。

backbone network to enhance global feature extraction, thereby improving the feature-representation robustness of the module when encountering occluded and blurred targets. Additionally, a minimum point distance-based intersection-over-union loss function is incorporated into the head network to increase sensitivity to spatial position differences in adjacent regions, further strengthening the model's generalizability in complex scenarios. On this basis, masked generative distillation is implemented to significantly improve detection accuracy without increasing the number of model parameters or computational cost. Experimental results on a transmission tower satellite image dataset show that the optimized algorithm achieves 1.8% and 4.7% improvements in mAP@50 and mAP@50:95, respectively. Compared with classical algorithms, the proposed method demonstrates higher detection accuracy in scenarios involving complex backgrounds, occluded targets, and blurred images.

Keywords: transmission tower; YOLO11; object detection; masked generative distillation

0　引言

电力系统是能源资源的基础平台，也是人类生产生活不可或缺的重要保障。输电塔作为电力传输系统的重要组成部分，是连接大型能源基地和用电地区的关键纽带。我国输电塔分布范围广，所属区域的地形地貌条件复杂，不可避免地受到地震等自然灾害的威胁[1]。输电塔一旦遭受破坏，将严重影响电力系统的正常运行，不仅会造成巨大的经济损失[2]，还可能削弱生命线工程中其他系统的运行能力，对震后应急救援和灾民生活造成严重影响。因此，准确判断输电塔的结构状态及其周边环境的演变具有重要意义，而输电塔识别是完成上述任务的基础。

早期的输电塔检测主要依赖人工设计的特征提取器和分类器，虽然取得了一定成效，但在不同光照、背景和角度条件下，其检测精度难以保证。随着数字图像处理技术的发展，基于卷积神经网络（Convolutional Neural Network, CNN）的目标检测算法逐渐应用于输电塔检测。目标检测算法可分为两类：双阶段目标检测和单阶段目标检测[3]。双阶段目标检测首先生成候选区域，再对候选区域进行边界框回归和分类，代表算法包括区域卷积神经网络（Region-based CNN, R-CNN）[4]、更快的基于区域的卷积神经网络（Faster Region-based CNN, Faster R-CNN）[5]等。郑鑫等[6]基于可变形网络和迁移学习改进了Faster R-CNN，以较高精度实现了输电塔识别。Wang等[7]基于Faster R-CNN和风格迁移图像合成技术，在小样本条件下同时识别了输电铁塔和绝缘子。曹志勇等[8]引入多尺度特征融合技术，以50层残差网络（Residual Network 50 layers, ResNet50）代替视觉几何组16层网络（Visual Geometry Group 16-layer Network, VGG-16）改进了Faster R-CNN，在电力塔基识别中取得了较高精度。

单阶段目标检测直接在图像上预测目标类别和边界框，代表算法是YOLO系列[9]。宋成根[10]通过改进YOLOv3模块，构建了基于金字塔池化模块、可变形卷积和注意力机制的特征提取模块，以较高精度识别了输电铁塔。方明等[11]结合样本数据增广方法，使用YOLOv4算法提升了输电杆塔自动检测精度。孙乐杨等[12]改进了YOLOv5算法，针对小目标增加大尺度检测层，在识别输电塔的同时计算其地理坐标。苏晓等[13]结合加权框融合推理机制，集成YOLOv5s和YOLOv5x算法，改善了小目标输电杆塔识别精度低的问题，并在不同天气、不同光照及复杂背景条件下取得了较好的检测结果。雷磊等[14]改进了YOLOv7模型，引入卷积块注意力模块和基于动态非单调聚焦机制的边界框损失函数，提高了输电铁塔塔基的识别精度。

国内外学者围绕输电塔检测已开展大量研究，针对数据集样本数量不足、输电塔目标尺度小、背景复杂等问题给出了切实有效的改进方案，但仍存在以下难题：受建筑、道路等背景干扰及成像模糊影响，易出现误检、漏检现象，检测框回归精度较低；输电塔被阴影、建筑或树木遮挡时，同样存在漏检问题，检测框回归精度不足；目标检测算法在检测精度和检测速度之间的平衡也需要进一步提升。针对上述问题，本文对YOLO11s算法进行改进，提出结合知识蒸馏技术的Linear Attention-YOLO11（简称LA-YOLO11）算法，主要工作如下：

（1）在YOLO11s算法骨干网络中，使用受

Mamba启发的线性注意力(Mamba-Inspired Linear Attention, MILA)模块替代原来的跨阶段部分与金字塔挤压注意力(Cross Stage Partial with Pyramid Squeeze Attention, C2PSA)模块,增强模型对全局上下文信息的感知能力,进一步提升模型对被遮挡及成像模糊的输电塔特征信息的提取能力。

(2) 在YOLO11s算法检测头部分,使用基于最小点距离的交并比(Minimum Point Distance-based Intersection over Union, MPDIoU)损失函数替换原来的完全交并比(Complete IoU, CIoU)损失函数。该函数对不同尺度目标使用相同的度量方式,对边界框的形状差异与位置差异更加敏感,有利于不同尺度目标检测,增强了模型在复杂场景中的泛化能力。

(3) 利用掩码生成蒸馏(Masked Generative Distillation, MGD)技术,以LA-YOLO11m模型为教师模型、LA-YOLO11s模型为学生模型进行知识蒸馏,在不增加模型参数量与计算量的前提下,进一步提高LA-YOLO11s模型的检测精度。

1 研究方法

1.1 结合知识蒸馏技术的LA-YOLO11算法

YOLO11算法是一种单阶段目标检测模型,延续了经典的Backbone-Neck-Head检测框架,在检测精度和检测速度上均有显著提升。主干网络通过层级卷积结构进行特征提取,构建多尺度特征金字塔,并引入C3K2模块和C2PSA模块。C3K2模块通过一系列较小内核的卷积优化网络中的信息流,降低了计算成本;C2PSA模块则在提取的特征上应用空间注意力,使模型关注感兴趣区域,更适于目标的精确提取。颈部网络通过双向跨尺度连接对主干网络输出的多级特征进行连接,提高特征金字塔的表达能力。头部网络负责模型的最终决策,同步优化分类置信度和边界框回归精度,并采用解耦设计,大幅减少了模型参数量和计算量,提高了检测速度。其损失函数采用二元交叉熵损失、分布式焦点损失和完全交并比损失的组合,并沿用动态的任务对齐分配方式。

YOLO11算法延续了YOLO系列模型的复合缩放策略,通过调整深度系数、宽度系数及最大通道数,提供YOLO11n、YOLO11s、YOLO11m、YOLO11l和YOLO11x五种模型,以满足不同场景的应用需求。各模型的复合缩放系数列于表1。其中,深度系数主要通过控制模型层数,直接影响特征的抽象层次和网络感受野,从而控制特征表示的复杂程度;宽度系数通过调整各层通道数,直接影响模型的参数数量和特征容量,从而控制模型特征表达能力;最大通道数则用于限制模型各层最大通道数,避免单层结构过度膨胀。

表1 YOLO11模型复合缩放系数

Table 1 Compound scaling coefficients for YOLO11 models

模型	深度系数	宽度系数	最大通道数
YOLO11n	0.50	0.25	1 024
YOLO11s	0.50	0.50	1 024
YOLO11m	0.50	1.00	512
YOLO11l	1.00	1.00	512
YOLO11x	1.00	1.50	512

YOLO11算法在多个数据集上均表现出良好性能[15],但在检测复杂背景、被遮挡或成像模糊的输电塔时,精度较低。本文对YOLO11算法进行改进,提出LA-YOLO11算法,其网络结构如图1所示。

LA-YOLO11算法的具体改进包括:首先,在YOLO11模型的骨干网络中使用MILA模块替代原有的C2PSA模块,以促进相邻窗口间的信息交换,同时提高计算效率;其次,使用MPDIoU损失函数替换原头部网络中的CIoU损失函数,将位置参数和形状参数分离优化;最后,利用掩码生成蒸馏(MGD)技术,以检测精度高、参数量大的LA-YOLO11m模型作为教师模型,以检测精度较低、参数量较小的LA-YOLO11s模型作为学生模型进行训练,在不增加模型参数量和计算量的前提下进一步提升检测精度。

1.2 线性注意力模块MILA

在机器学习中,注意力机制能够增强模型对特征信息的感知能力。本文针对输电塔识别任务的实际需求,在骨干网络中引入MILA模块替代原来的C2PSA模块,通过门控机制抑制被遮挡区域对特征提取的干扰,使注意力权重聚焦于可见区域。同时,通过特征移位操作增强模型对上下文信息的感知能力,弥补了局部注意力的局限性。

MILA模块是Han等[16]通过比较Mamba模型和线性注意力Transformer模型的相似之处,将遗忘门及模块设计集成到线性注意力中而提出的。其结构如图1所示,处理流程如下:

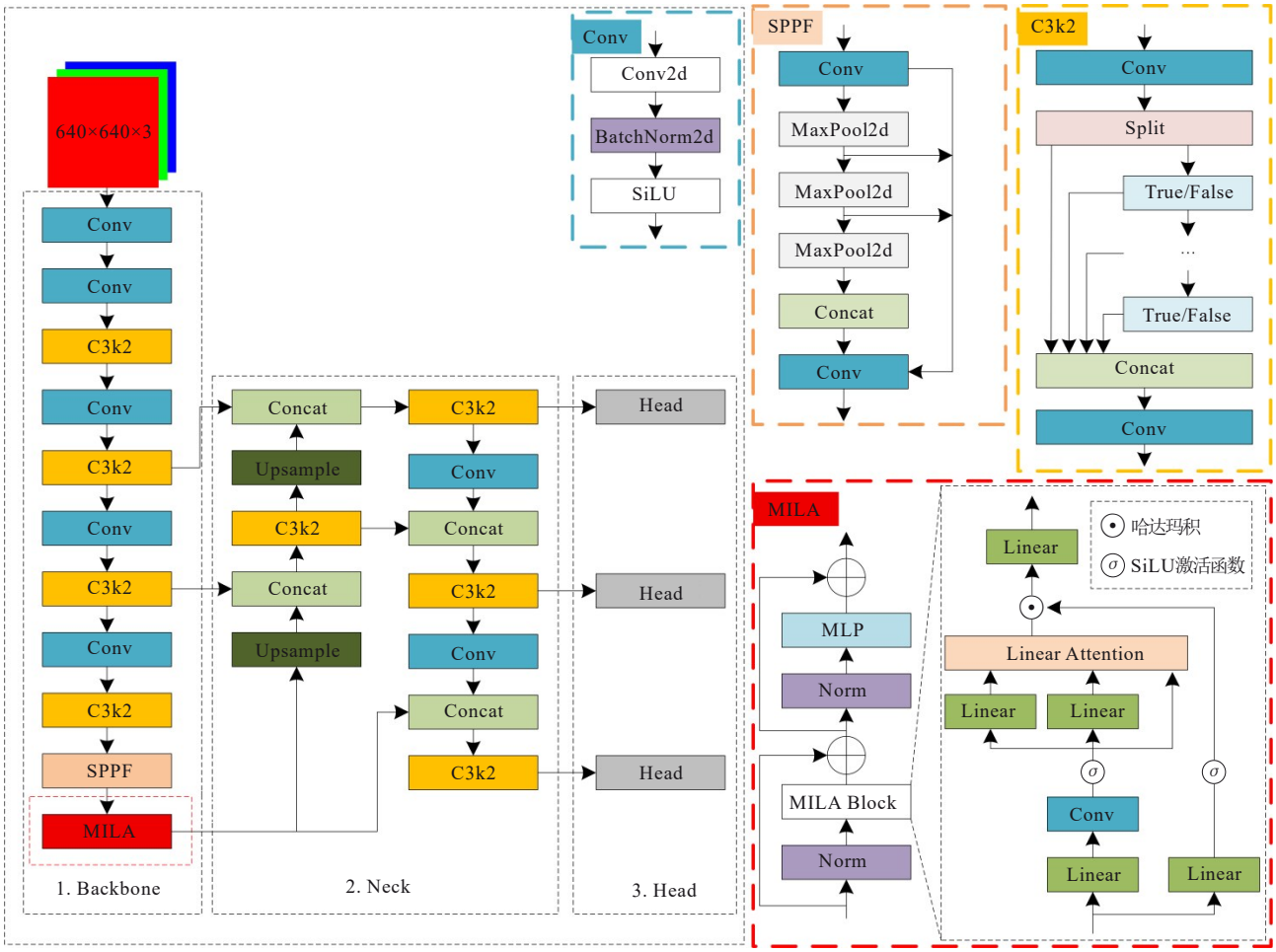


图 1　LA-YOLO11 算法网络结构

Fig. 1　Network architecture of the LA-YOLO11 algorithm

$$\text{Attention}=\text{LinearAttention}\left\{\text{Conv}\left[L(x)\right]\right\} \tag{1}$$

$$\text{Output}=L\left[\text{Attention}\odot L(x)\right] \tag{2}$$

式中：x 为输入特征；L 表示线性变换；Conv 表示卷积层；$\odot$ 表示哈达玛积。LinearAttention 操作如下：

$$\boldsymbol{Q}=f(x\boldsymbol{W}_Q),\boldsymbol{K}=f(x\boldsymbol{W}_K),\boldsymbol{V}=x\boldsymbol{W}_V \tag{3}$$

$$\boldsymbol{y}_i=\sum_{j=1}^{N}\frac{\boldsymbol{Q}_i\boldsymbol{K}_j^{\mathrm{T}}}{\sum_{j=1}^{N}\boldsymbol{Q}_i\boldsymbol{K}_j^{\mathrm{T}}}\boldsymbol{V}_j \tag{4}$$

式中：f 表示核函数；$\boldsymbol{W}_Q$、$\boldsymbol{W}_K$、$\boldsymbol{W}_V$ 均表示投影矩阵；$\boldsymbol{Q}$、$\boldsymbol{K}$、$\boldsymbol{V}$ 分别表示查询矩阵、关键矩阵、数值矩阵。

1.3　MPDIoU 损失函数

在机器学习中，损失函数是衡量预测框和真实框差异程度的指标。合适的损失函数不仅可以加快模型的收敛速度，还能提高目标检测的准确度。在本文的应用场景中，输电塔处于复杂背景中，且存在被遮挡和成像模糊的情况，原有的 CIoU 损失函数未充分考虑目标的尺度相关性，且难以有效区分相邻实例的空间位置差异。因此，本文引入 MPDIoU 损失函数[17]，替代检测头部分的 CIoU 损失函数。

MPDIoU 损失函数通过最小化预测框与真实框之间的顶点距离优化回归过程，适用于复杂场景中的目标检测。MPDIoU 计算公式如下：

$$\text{MPDIoU}=\text{IoU}-\frac{d_1^2}{w^2+h^2}-\frac{d_2^2}{w^2+h^2} \tag{5}$$

式中：w 和 h 分别为输入图像的宽度和高度；d_1^2 和 d_2^2 计算方法如下：

$$d_1^2=(x_1^{\text{pred}}-x_1^{\text{gt}})^2+(y_1^{\text{pred}}-y_1^{\text{gt}})^2 \tag{6}$$

$$d_2^2=(x_2^{\text{pred}}-x_2^{\text{gt}})^2+(y_2^{\text{pred}}-y_2^{\text{gt}})^2 \tag{7}$$

式中：$(x_1^{\text{pred}},\ y_1^{\text{pred}})$ 为预测框左上角坐标；$(x_1^{\text{gt}},\ y_1^{\text{gt}})$ 为标注框左上角坐标；$(x_2^{\text{pred}},\ y_2^{\text{pred}})$ 为预测框右下角坐标；$(x_2^{\text{gt}},\ y_2^{\text{gt}})$ 为标注框右下角坐标。

MPDIoU 损失函数的定义式为：

$$L_{\text{MPDIoU}}=1-\text{MPDIoU} \tag{8}$$

1.4 MGD技术

与LA-YOLO11s模型相比，LA-YOLO11m模型通过增加各层通道数增强了特征表达能力，但模型计算量的增加也导致推理速度下降。针对这一性能与效率的权衡问题，知识蒸馏技术提供了一种切实有效的解决方案[18]。本文使用MGD技术，将参数量多、精度较高的LA-YOLO11m模型作为教师模型，将参数量少、精度较低的LA-YOLO11s模型作为学生模型进行训练，在不增加学生模型参数量的条件下进一步提升其检测精度。

MGD是一种基于特征空间变换的知识蒸馏方法，其核心思想是通过引入随机掩膜机制，利用教师模型指导学生模型进行特征恢复，从而提高学生模型的特征表达能力[19]。首先，设置随机掩码覆盖学生模型的第l层特征，方法如下：

$$M_{i,j}^{l}=\begin{cases}0, & R_{i,j}^{l}<\lambda\\ 1, & \text{其他}\end{cases} \tag{9}$$

式中：$R_{i,j}^{l}$是区间(0, 1)范围内的随机数；i、j分别为特征图的横坐标与纵坐标；λ为表示掩码比例的超参数。

然后，使用该掩码覆盖学生模型的特征层，并根据剩余像素尝试生成教师模型的特征层：

$$\zeta\left[f_{\text{align}}(S^{l})\cdot M^{l}\right]\rightarrow T^{l} \tag{10}$$

$$\zeta(F)=W_{l2}\left\{\text{ReLU}\left[W_{l1}(F)\right]\right\} \tag{11}$$

式中：ζ表示投影层，该层包含卷积层W_{l1}、W_{l2}和激活函数层ReLU；T^{l}、S^{l}分别表示教师模型和学生模型的第l层特征；f_{align}为使学生模型特征和教师模型特征保持一致的适配器层。

MGD技术的蒸馏损失函数表达式为：

$$L_{\text{dis}}(S,T)=\sum_{l=1}^{L}\sum_{k=1}^{C}\sum_{i=1}^{H}\sum_{j=1}^{W}\left\{T_{k,i,j}^{l}-\zeta\left[f_{\text{align}}(S_{k,i,j}^{l})\cdot M_{i,j}^{l}\right]\right\}^{2} \tag{12}$$

式中：L是蒸馏层数的总和；C、H、W分别表示特征图的通道数、高度与宽度。

在模型训练过程中，采用的总损失函数如下：

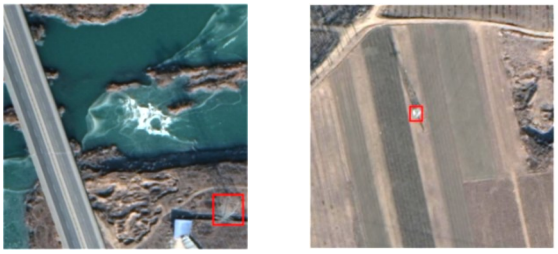
$$L_{\text{all}}=L_{\text{original}}+\alpha L_{\text{dis}} \tag{13}$$

式中：L_{original}为模型原来的损失函数；α为平衡两个损失函数的超参数。

2 数据集

选取2018—2023年间中国河北省唐山市Google 19级卫星遥感影像(分辨率0.2～0.3 m)，基于ArcGIS平台对输电塔进行标注，并将遥感影像与标注矢量同步裁剪为640×640像素的图块，制作成YOLO格式数据集。图像文件格式为JPG，标签文件格式为TXT，包含类别标识，标注框中心点横、纵坐标，标注框宽、框高。

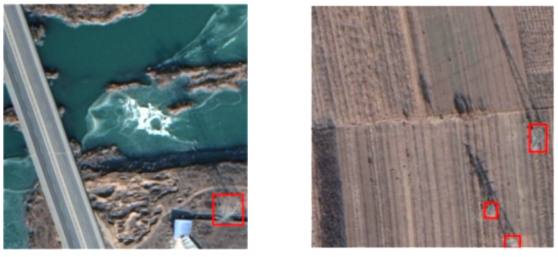
数据集共包含图像2 278张，其中含有一定数量的困难样本(源于微小目标、极端光照、部分遮挡)，其纹理和颜色信息存在不同程度的丢失。简单样本与困难样本的对比如图2所示。数据集中每张图像至少包含1个输电塔目标，且处于不同光照、角度及背景条件下。复杂背景包括森林、草地、裸土、道路、农田、建筑、河流等，如图3所示。



(a) 正常目标与微小目标对比(尺寸)



(b) 正常光照与极端光照对比(平均亮度)



(c) 无遮挡目标与有遮挡目标对比

图2 简单样本与困难样本对比

Fig. 2 Comparison between simple and difficult samples

在模型训练过程中，对数据集样本进行随机增强(如图像平移、图像旋转、图像翻转、图像缩放、色彩抖动、透视变换、Mosaic增强等)，以扩充数据集，提高模型在复杂场景中的泛化能力。

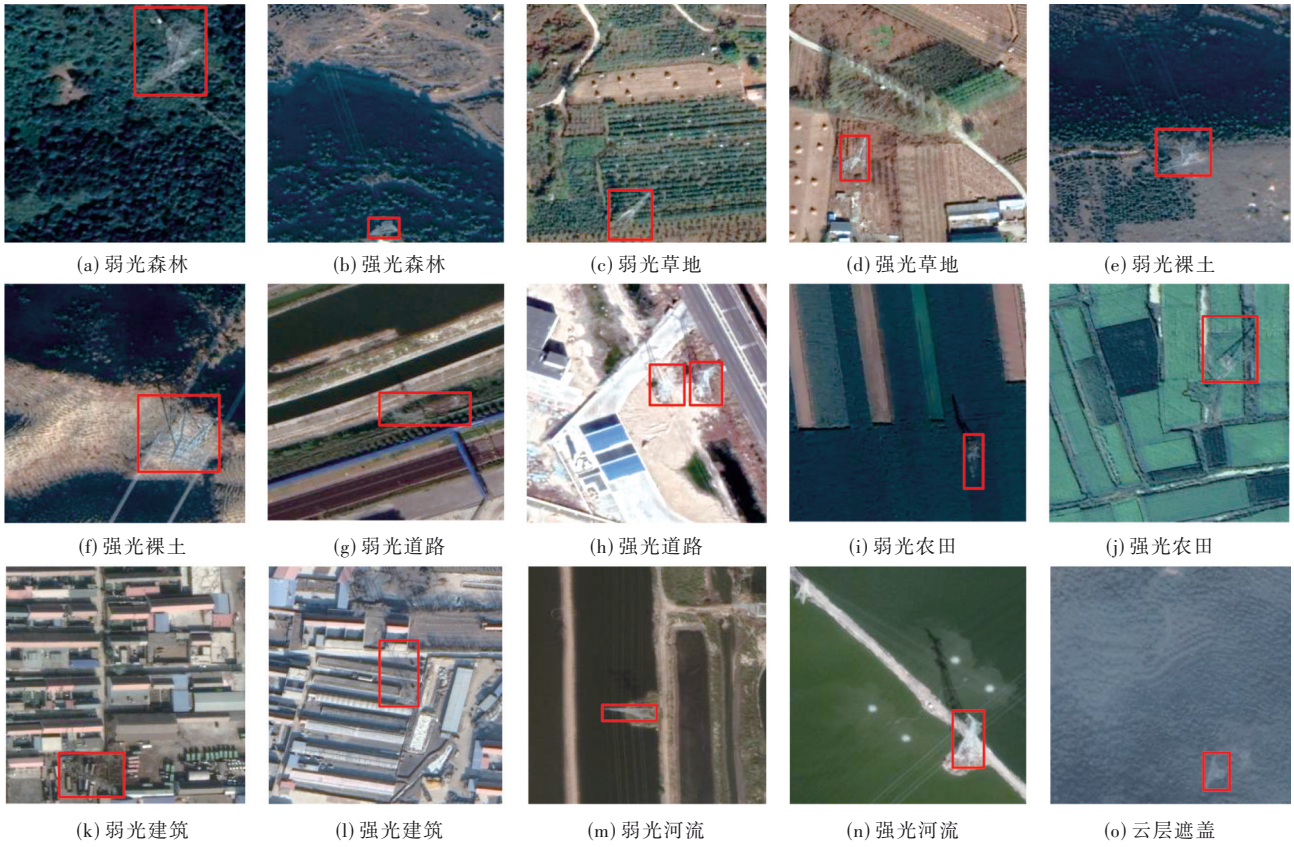


图 3　数据集背景

Fig. 3　Dataset backgrounds

3　实验结果与分析

3.1　实验环境与参数设置

实验平台采用的操作系统为 Windows Server 2019 Standard，Python 版本为 3.11，深度学习框架为 Pytorch 2.5，CUDA 版本为 12.2，cuDNN 版本为 9.1.0。硬件配置方面，CPU 型号为 Intel(R) Xeon (R) Gold 6342，GPU 型号为 NVIDIA A30。训练过程中，Batch Size 设置为 16，初始学习率为 0.01，优化器使用随机梯度下降。为保证实验的可重复性，随机数生成器种子设置为 42，并启用确定性模式进行训练。

3.2　评价指标

本文采用精确度(P)、召回率(R)和平均精度均值(mean Average Precision，mAP)作为模型性能的评价指标。精确度是指所有被模型预测为正类的样本中实际为正类的比例，定义如式(14)所示。召回率是指所有实际为正类的样本中被模型预测为正类的比例，定义如式(15)所示。

$$P=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}} \tag{14}$$

$$R=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \tag{15}$$

式中：TP 表示真正例；FP 表示假正例；FN 表示假反例。

平均精度(Average Precision，AP)是模型对预测结果准确性和可靠性的综合评估指标，反映了模型在不同召回率下精确度的平均值。以召回率为横坐标、精确度为纵坐标绘制 P-R 曲线，对曲线积分得到 AP。平均精度均值(mAP)是指所有类别 AP 的平均值，计算公式如下：

$$\mathrm{mAP}=\frac{1}{N}\sum_{i=1}^{N}\int_{0}^{1}P_i(R)\mathrm{d}R \tag{16}$$

mAP@50 是交并比(IoU)为 0.5 时的平均精度均值，mAP@50:95 是 IoU 从 0.5 到 0.95 的平均精度均值。mAP@50 和 mAP@50:95 值越高，表明模型的检测精度越高。

其他性能评价指标包括模型参数量(Params)、模型计算量(FLOPs)和模型大小(Model Size)。模型参数量越少、计算量越小、模型越小，说明模型的

检测速度越快[20]。

3.3 实验结果分析

按照训练集∶验证集∶测试集=8∶1∶1的比例对自建输电塔数据集进行划分，基于该数据集使用改进后的模型进行训练、验证和测试。

3.3.1 MILA有效性验证

为验证MILA模块替换YOLO11s骨干网络中C2PSA模块在自建输电塔数据集上的有效性，分别使用上下文锚点注意力(Context Anchor Attention，CAA)模块[21]、高效多尺度注意力(Efficient Multi-scale Attention，EMA)模块[22]及MILA模块[16]替代YOLO11s模型中的相同位置，进行实验对比，结果列于表2。

表2 注意力机制实验结果对比

Table 2 Comparison of attention mechanism experimental results

模型	mAP@50/%
YOLO11s(C2PSA)	92.5
YOLO11s(CAA)	92.9
YOLO11s(EMA)	93.1
YOLO11s(MILA)	93.3

由表2可知，不同注意力机制对模型的影响程度不同。使用MILA模块替换YOLO11s模型中的C2PSA模块后，mAP@50值提高了0.8%，在对比实验中表现最佳，说明MILA模块能够更好地帮助模型关注输电塔的特征信息，提高目标检测精度。

3.3.2 MPDIoU损失函数有效性验证

为验证MPDIoU损失函数替换YOLO11s模型检测头中CIoU损失函数在自建输电塔数据集上的有效性，分别使用统一交并比(Unified IoU，UIoU)损失函数[23]、更聚焦的交并比(More Focused IoU，Focaler-IoU)损失函数[24]，以及MPDIoU损失函数[17]替代YOLO11s模型中的相同位置，进行实验对比，结果列于表3。

表3 损失函数实验结果对比

Table 3 Comparison of loss function experimental results

模型	mAP@50/%
YOLO11s(CIoU)	92.5
YOLO11s(UIoU)	92.8
YOLO11s(Focaler-IoU)	92.5
YOLO11s(MPDIoU)	93.4

由表3可知，不同损失函数对模型的影响程度不同。使用MPDIoU损失函数替换YOLO11s模型中的CIoU损失函数后，mAP@50值提高了0.9%，在对比实验中表现最佳，说明MPDIoU损失函数能更好地帮助模型评价预测框和真实框的差异程度，提高目标检测精度。

3.3.3 消融实验

在YOLO11s模型的骨干网络中引入MILA模块，在检测头部分引入MPDIoU损失函数，构建LA-YOLO11s模型。为验证所提改进策略对目标检测精度的提升效果，以YOLO11s模型为基准，基于自建输电塔数据集开展消融实验，结果列于表4。

表4 消融实验结果对比

Table 4 Comparison of ablation experimental results

模型	mAP@50/%	mAP@50∶95/%	Params/M	FLOPs/G	Model Size/MB
YOLO11s	92.5	58.9	9.413	21.5	19.2
YOLO11s(MILA)	93.3	59.5	9.349	21.5	19.1
YOLO11s(MPDIoU)	93.4	60.9	9.413	21.5	19.2
YOLO11s(MILA，MPDIoU)	93.3	61.9	9.349	21.5	19.1

由表4可知，在YOLO11s模型中引入MILA模块后，模型参数量略有下降，计算量不变，mAP@50和mAP@50∶95值分别提升0.8%和0.6%。在YOLO11s模型中引入MPDIoU损失函数后，模型参数量及计算量均无变化，mAP@50和mAP@50∶95值分别提升0.9%和2%。上述结果表明，两种改进策略对YOLO11s基准模型均有效。LA-YOLO11s模型综合这两种改进策略，在模型参数量下降的同时，其mAP@50值提升0.8%，mAP@50∶95值提升3%，进一步验证了所提改进策略在提升目标检测精度方面的有效性。

3.3.4 MGD实验结果对比

选取模型参数量大、计算量高的LA-YOLO11m模型作为教师模型，参数量小、计算量低的LA-YOLO11s模型作为学生模型，利用MGD策略进行训练，蒸馏后的模型记为LA-YOLO11s+distill。

LA-YOLO11s 和 LA-YOLO11m 模型具有相同的网络结构，且参数量相差不大，有利于提升学生模型的泛化能力。蒸馏前后模型的各项指标列于表5。

表 5　蒸馏前后模型指标对比

Table 5　Comparison of model metrics before and after distillation

模型	mAP@50/%	mAP@50:95/%	Params/M	FLOPs/G	Model Size/MB
LA-YOLO11s	93.3	61.9	9.349	21.5	19.1
LA-YOLO11m	92.7	62.5	19.966	68.1	40.4
LA-YOLO11s+distill	94.3	63.6	9.349	21.5	19.1

由表5可知，教师模型的参数量和计算量远大于学生模型，在测试集上的mAP@50:95值比学生模型高0.6%。采用MGD策略后，蒸馏模型在不改变参数量和计算量的情况下，mAP@50值提升了1%，mAP@50:95值提升了1.7%，说明了MGD策略在输电塔检测任务中的有效性。

3.3.5　经典算法对比实验

为验证本文算法在输电塔识别任务中的优越性，在自建输电塔数据集上将其与多种经典算法进行对比，包括双阶段目标检测算法 Faster RCNN[5]、单阶段目标检测算法——单激发多框探测器（Single Shot MultiBox Detector, SSD）[25]、YOLOv5s[26]、YOLOv8s[26]、YOLO11s[26]，以及实时检测算法 RT-DETR-L[27]，结果列于表6。

从表6可以看出，基准模型 YOLO11s 相比其他经典算法，计算量最小，参数量和模型大小仅略高于 YOLOv5s 算法，且检测精度（mAP@50值）较高。改进后的模型 LA-YOLO11s+distill，其检测精度最高：mAP@50为94.3%，mAP@50:95为63.6%；计算量为21.5 G，低于其他对比算法；参数量为9.349 M，模型大小为19.1 MB，仅略高于 YOLOv5s 算法。综上，本文算法在输电塔识别任务中检测精度最高，且计算量最低，推理速度较快。

表 6　经典算法实验结果对比

Table 6　Comparison of classical algorithm experimental results

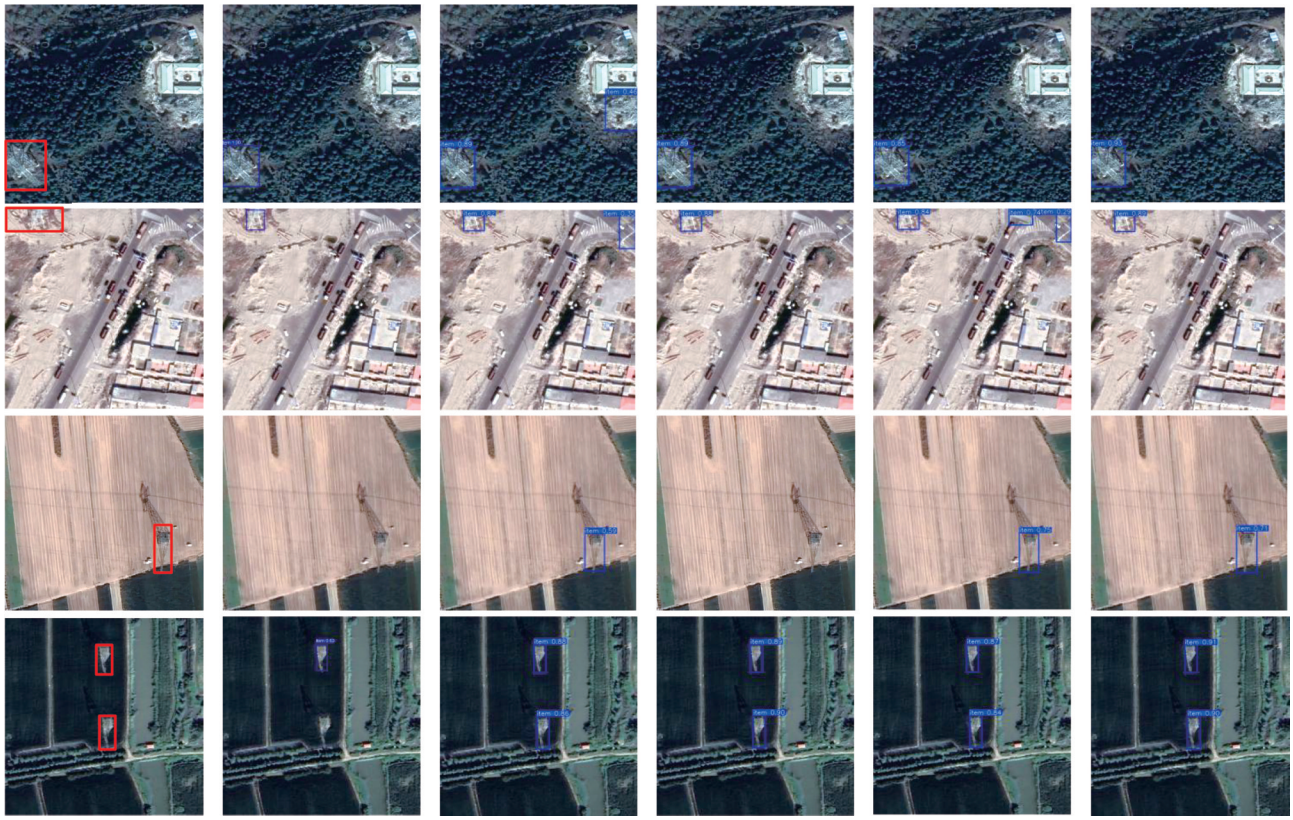
模型	mAP@50/%	mAP@50:95/%	Params/M	FLOPs/G	Model Size/MB
SSD	81.6	60.9	22.943	30.1	87.5
Faster R-CNN	83.5	37.3	28.275	253.7	107.9
YOLOv5s	92.1	60.4	9.111	24.0	18.6
YOLOv8s	93.4	61.1	11.125	28.6	22.6
YOLO11s	92.5	58.9	9.413	21.5	19.2
RT-DETR-L	92.3	60.3	31.985	108.0	66.2
LA-YOLO11s+distill	94.3	63.6	9.349	21.5	19.1

3.3.6　检测结果对比实验

为了更加直观地展示本文算法在输电塔识别任务中的优越性，使用该算法和其他计算量较低、推理速度较快的经典算法，对复杂背景（如森林、建筑、道路、裸地、农田）中的图片进行检测，结果对比如图4所示；对目标被遮挡的图片进行检测，结果对比如图5所示；对目标成像模糊的图片进行检测，结果对比如图6所示。

由图4可知，在以森林和建筑为背景的场景中，YOLOv5s 算法存在误检（将建筑旁边的道路误检为输电塔），其余算法检测结果较准确。在以道路为背景的场景中，YOLOv5s 和 YOLO11s 算法均存在误检（将斑马线和路灯误检为输电塔），其余算法检测结果较准确。在以裸地为背景的场景中，SSD 和 YOLOv8s 算法存在漏检现象，其余算法检测结果较准确。在以农田为背景的场景中，SSD 算法存在漏检现象，其余算法检测结果较准确。根据各算法在不同背景下的检测结果对比，本文算法更适合复杂背景中的输电塔检测。

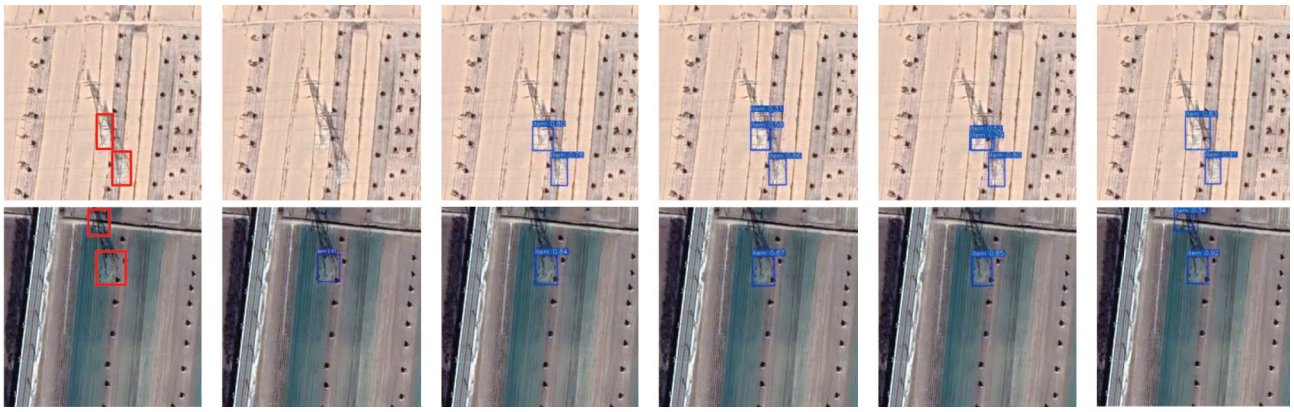
根据图5中第一张图片的检测结果，在目标被遮挡的场景下，SSD 算法存在漏检，YOLOv8s 和 YOLO11s 算法对输电塔边界框预测不够准确，其余算法检测结果较准确。使用各算法对图5第二张图片进行检测，仅本文算法能够检测到两个目标输



(a) 原图像 (b) SSD (c) YOLOv5s (d) YOLOv8s (e) YOLO11s (f) 本文算法

图 4 检测结果对比(复杂背景)

Fig. 4 Comparison of detection results (complex backgrounds)



(a) 原图像 (b) SSD (c) YOLOv5s (d) YOLOv8s (e) YOLO11s (f) 本文算法

图 5 检测结果对比(目标被遮挡)

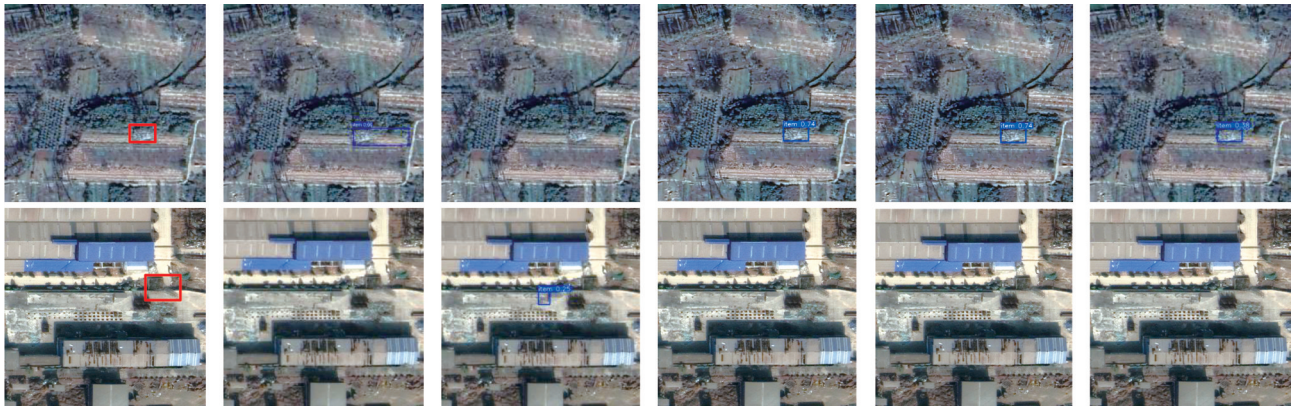
Fig. 5 Comparison of detection results (occluded targets)

电塔。上述对比表明,本文算法更适合被遮挡场景中的目标检测。

由图 6 可知,使用各算法对输电塔成像模糊的图片进行预测:第一张图片中,SSD 算法对输电塔的边界框预测不够准确,YOLOv5s 算法无法正确识别输电塔,其余算法检测结果较准确;第二张图片,各算法均无法准确识别模糊目标,且 YOLOv5s 算法在背景区域存在误检。以上结果表明,除某些极端模糊情况外,本文算法在目标成像模糊场景中能够给出较为准确的识别结果。

4 结论

本研究针对目前输电塔识别研究中存在的主要问题,以提升复杂场景(如不同背景、目标被遮



(a) 原图像　(b) SSD　(c) YOLOv5s　(d) YOLOv8s　(e) YOLO11s　(f) 本文算法

图 6　检测结果对比（模糊图像）

Fig. 6　Comparison of detection results (blurred images)

挡、成像模糊）下输电塔识别的准确性为目标，改进了 YOLO11s 目标检测算法，通过引入 MILA 模块和 MPDIoU 损失函数，并结合掩码生成蒸馏技术，提出了 LA-YOLO11s 模型。

首先，基于高分辨率卫星遥感影像建立了输电塔数据集。该数据集图片涵盖不同背景、光照和角度条件，且包含一定数量的目标被遮挡、目标尺寸小、目标成像模糊等困难样本。基于该数据集进行模型训练、验证和测试，通过理论分析和实验对比，验证了 MILA 模块和 MPDIoU 损失函数对目标检测的有效性。结合掩码生成蒸馏技术，在不增加模型参数量和计算量的前提下，进一步提升了模型的检测精度。与 YOLO11s 基线模型相比，改进后模型的 mAP@50 提高了 1.8%，mAP@50：95 提高了 4.7%。

将本文算法与其他经典算法在不同背景、目标被遮挡、成像模糊等区域进行输电塔检测。对比实验结果表明：本文算法在不同背景及目标被遮挡的场景中具有较高的识别精度，表现出一定的泛化能力；在目标成像模糊的场景中，本文算法在一定程度上能够得到较准确的检测结果，但在极端情况下仍存在漏检现象。后续研究将致力于进一步提升模型的目标检测性能，并将本文模型应用到输电塔检测任务中，拓展其应用价值和实际意义。

参考文献（References）

[1] 缪惠全，高思远，刘如山，等. 生命线工程的可靠性研究进展分析（Ⅱ）：国内电力系统的抗震可靠性[J]. 北京工业大学学报，2024，50(3)：362-384.
Miao Huiquan，Gao Siyuan，Liu Rushan，et al. Research progress on reliability of lifeline engineering (Ⅱ)：domestic research on seismic reliability of electricity system[J]. Journal of Beijing University of Technology，2024，50(3)：362-384.

[2] 张中近. 电力设施地震经济损失快速评估[D]. 哈尔滨：中国地震局工程力学研究所，2017：1.
Zhang Zhongjin. Rapid evaluation of electric power facility economic loss caused by earthquake[D]. Harbin：Institute of Engineering Mechanics，China Earthquake Administration，2017：1.

[3] 郭庆梅，刘宁波，王中训，等. 基于深度学习的目标检测算法综述[J]. 探测与控制学报，2023，45(6)：10-20，26.
Guo Qingmei，Liu Ningbo，Wang Zhongxun，et al. Review of deep learning based object detection algorithms[J]. Journal of Detection & Control，2023，45(6)：10-20，26.

[4] Girshick R，Donahue J，Darrell T，et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]// Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus：IEEE，2014：580-587.

[5] Ren S Q，He K M，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.

[6] 郑鑫，潘斌，张健. 可变形网络与迁移学习相结合的电力塔遥感影像目标检测法[J]. 测绘学报，2020，49(8)：1042-1050.
Zheng Xin，Pan Bin，Zhang Jian. Power tower detection in remote sensing imagery based on deformable network and transfer learning[J]. Acta Geodaetica et Cartographica Sinica，2020，49(8)：1042-1050.

[7] Wang T，Wei R Z，Wang L，et al. Detection of transmission towers and insulators in remote sensing images with deep learning[C]// Proceedings of the 2021 China Automation Congress (CAC). Beijing：IEEE，2022：3298-3303.

[8] 曹志勇，丰佳，毛文利，等. 基于特征融合 Faster R-CNN 的电力塔基目标检测[J]. 浙江电力，2021，40(11)：72-77.
Cao Zhiyong，Feng Jia，Mao Wenli，et al. Object detection of power tower bases based on faster R-CNN with feature fusion[J]. Zhejiang Electric Power，2021，40(11)：72-77.

[9] 徐彦威，李军，董元方，等. YOLO系列目标检测算法综述[J]. 计算机科学与探索，2024，18(9)：2221-2238.
Xu Yanwei，Li Jun，Dong Yuanfang，et al. Survey of development of YOLO object detection algorithms[J]. Journal of Frontiers of Computer Science and Technology，2024，18(9)：2221-2238.

[10] 宋成根. 基于高分辨率航空遥感影像的输电铁塔智能检测方法研究[D]. 阜新：辽宁工程技术大学，2021：21.
Song Chenggen. Research on intelligent detection method of transmission tower based on high-resolution aerial remote sensing image[D]. Fuxin：Liaoning Technical University，2021：21.

[11] 方明，孙晓敏，黄然，等. 面向高分辨率卫星遥感的电力杆塔自动检测[J]. 航天返回与遥感，2021，42(5)：118-126.
Fang Ming，Sun Xiaomin，Huang Ran，et al. Research on automatic detection algorithm of power tower using high resolution remote sensing satellite image[J]. Spacecraft Recovery & Remote Sensing，2021，42(5)：118-126.

[12] 孙乐杨，凌振宝，王永志. 基于改进的YOLOv5识别遥感影像中输电塔的方法[J]. 实验技术与管理，2022，39(4)：19-24.
Sun Leyang，Ling Zhenbao，Wang Yongzhi. Transmission tower identification method based on improved YOLOv5 in remote sensing image[J]. Experimental Technology and Management，2022，39(4)：19-24.

[13] 苏晓，张明晖，陈峻宇，等. 基于集成YOLOv5算法的输电线路杆塔目标检测[J]. 农村电气化，2023(5)：33-39.
Su Xiao，Zhang Minghui，Chen Junyu，et al. Target detection of transmission line towers based on integrated YOLOv5 algorithm[J]. Rural Electrification，2023(5)：33-39.

[14] 雷磊，魏小龙，梁俊，等. 基于改进YOLOv7的输电铁塔塔基检测算法[J]. 陕西师范大学学报(自然科学版)，2024，52(3)：85-95.
Lei Lei，Wei Xiaolong，Liang Jun，et al. A novel algorithm based on the improved YOLOv7 for detecting transmission tower base[J]. Journal of Shaanxi Normal University (Natural Science Edition)，2024，52(3)：85-95.

[15] 王宁，智敏. 深度学习下的单阶段通用目标检测算法研究综述[J]. 计算机科学与探索，2025，19(5)：1115-1140.
Wang Ning，Zhi Min. Review of one-stage universal object detection algorithms in deep learning[J]. Journal of Frontiers of Computer Science and Technology，2025，19(5)：1115-1140.

[16] Han D，Wang Ziyi，Xia Zhuofan，et al. Demystify mamba in vision：a linear attention perspective[J]. Advances in Neural Information Processing Systems，2024，37：127181-127203.

[17] Ma S L，Xu Y. MPDIoU：a loss for efficient and accurate bounding box regression[EB/OL]. (2023-07-14) [2025-07-02]. https://arxiv.org/abs/2307.07662.

[18] 黄震华，杨顺志，林威，等. 知识蒸馏研究综述[J]. 计算机学报，2022，45(3)：624-653.
Huang Zhenhua，Yang Shunzhi，Lin Wei，et al. Knowledge distillation：a survey[J]. Chinese Journal of Computers，2022，45(3)：624-653.

[19] Yang Zhendong，Li Zhe，Shao Mingqi，et al. Masked generative distillation[C]// Proceedings of the 2022 European Conference on Computer Vision. Cham：Springer Nature Switzerland，2022：53-69.

[20] 李珺，丁彬彬，史维娟，等. HM-YOLO：融合多尺度特征的轻量化航拍图像检测算法[J]. 计算机工程与应用，2026，62(1)：87-100.
Li Jun，Ding Binbin，Shi Weijuan，et al. HM-YOLO：lightweight aerial image detection algorithm fusing multi-scale features[J]. Computer Engineering and Applications，2026，62(1)：87-100.

[21] Cai X H，Lai Q X，Wang Y W，et al. Poly kernel inception network for remote sensing detection[C]//Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle，WA：IEEE，2024：27706-27716.

[22] Ouyang D L，He S，Zhang G Z，et al. Efficient multi-scale attention module with cross-spatial learning[C]// Proceedings of the 2023 International Conference on Acoustics，Speech and Signal Processing (ICASSP). Rhodes Island，Greece：IEEE，2023：1-5.

[23] Luo X J，Cai Z H，Shao B，et al. Unified-IoU：for high-quality object detection[EB/OL]. (2024-08-13) [2025-07-02]. https://arxiv.org/abs/2408.06636.

[24] Zhang H，Zhang S J. Focaler-IoU：more focused intersection over union loss[EB/OL]. (2024-01-19) [2025-07-02]. https://arxiv.org/abs/2401.10525.

[25] Liu W，Anguelov D，Erhan D，et al. SSD：single shot MultiBox detector[C]// Proceedings of the 2016 European Conference on Computer Vision. Cham：Springer，2016：21-37.

[26] 岑伟迪，江佳玲，黄勃，等. YOLO十年演进：从实时检测先锋到多任务智能前沿[J/OL]. 武汉大学学报(理学版)，2025-07-02. https://doi.org/10.14188/j.1671-8836.2025.0116.
Cen Weidi，Jiang Jialing，Huang Bo，et al. Ten years of YOLO evolution：from real-time detection pioneer to multi-tasking intelligence frontier[J/OL]. Journal of Wuhan University (Natural Science Edition)，2025-07-02. https://doi.org/10.14188/j.1671-8836.2025.0116.

[27] Zhao Y A，Lv W Y，Xu S L，et al. DETRs beat YOLOs on real-time object detection[C]// Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle，WA，USA：IEEE，2024：16965-16974.

(本文编辑：赵乘程)