DOI: 10.19650/j.cnki.cjsi.J2311576

# 基于改进 $R^3$ det 的无人机电力杆塔倾斜程度检测*

胡 霞[1],仲林林[2]
(1. 东南大学-蒙纳士大学苏州联合研究生院(东南大学) 苏州 215123; 2. 东南大学电气工程学院 南京 210096)

**摘 要**:无人机巡检图像中的电力杆塔具有多姿态、大长宽比等特点,难以利用特有的先验知识来准确定位和判别不同倾斜程度的杆塔。本文提出了一种改进的 $R^3$det 网络模型(Multi-Head-KF-$R^3$det),可提高电力杆塔倾斜程度检测精度。首先,在原始 $R^3$det 中引入倾斜程度分支,实现了电力杆塔类别和倾斜程度的判别以及电力杆塔的准确定位。然后,将基于卡尔曼滤波的旋转交并比损失项引入回归损失函数中,在不增加额外超参的情况下,进一步提升了模型整体检测精度以及倾斜程度检测召回率。最后,基于 Ghost 轻量化网络设计原理对改进后的模型进行合理压缩,为模型在嵌入式设备中的部署奠定基础。实验结果表明,Multi-Head-KF-$R^3$det 在多尺度和多姿态的电力杆塔数据集上检测精度和召回率分别可达 94.5% 和 94.9%。

**关键词**: 电力杆塔;无人机巡检;旋转目标检测;倾斜程度检测
**中图分类号**: TP391.4 TH701 **文献标识码**: A **国家标准学科分类代码**: 470.4051

# Incline detection of power towers from UAV images based on the improved $R^3$det

Hu Xia[1], Zhong Linlin[2]
(*1. Southeast University-Monash University Joint Graduate School (Suzhou), Southeast University, Suzhou 215123, China*;
*2. School of Electrical Engineering, Southeast University, Nanjing 210096, China*)

**Abstract**: Power towers in unmanned aerial vehicle (UAV) inspection images have the characteristics of multi-attitude and large-aspect ratios, which are difficult to accurately locate and distinguish towers with different degrees of inclination according to unique prior knowledge. To improve the incline detection accuracy, this article proposes a method for processing UAV images based on the improved refined rotation RetinaNet (Multi-Head-KF-$R^3$det). Firstly, the incline detection head is added to the original $R^3$det model to achieve the classification of power towers and their inclination degrees, as well as the accurate location. Then, the Kalman-filter intersection over union loss is introduced into the regression loss function to effectively improve the overall detection accuracy and recall rates of incline detection without additional hyperparameters. Finally, the improved model is reasonably compressed based on the design principles of ghost lightweight network, which lays a foundation for the deployment of the model in embedded devices. The experimental results show that the mAP and recall rates of Multi-Head-KF-$R^3$det on multi-scale and multi-attitude power tower datasets can reach 94.5% and 94.9%, respectively.

**Keywords**: power tower; UAV inspection; rotation object detection; incline detection

## 0 引 言

电力杆塔作为支撑架空输电线路的主要设备,在电能从发电厂经由变电站传输到终端用户的过程中起着重要作用[1]。随着我国能源电力发展进程的不断加快,分布日趋广泛的电力杆塔在自然灾害频发的地区容易出现杆塔倾斜甚至倒断等异常情况,影响电能的正常传输,若不及时维护将带来更多安全隐患[2]。

当前,传统人工巡检方式不仅检测效率低下,且其检测精度易受到巡检人员专业水平、工作状态等多方面主观因素影响[3]。为此,电力运维部门引入无人机巡检技

收稿日期:2023-06-20 Received Date: 2023-06-20
*基金项目:国家自然科学基金(92066106)、江苏省科协青年科技人才托举工程(2021031)、东南大学“至善青年学者”支持计划(中央高校基本科研业务费)(2242022R40022)项目资助

术,通过无人机搭载各类传感装置实现输电线路及其部件检测[4]。同时,借助深度学习强大的特征提取优势来准确提取无人机巡检图像中的目标特征,为下游任务的进一步执行打下坚实基础[5-6]。文献[7]提出了一种改进的 EfficientDet 用于均压环、绝缘子等电网元件检测,通过在特征网络中引入双向特征金字塔结构和坐标注意力机制,其检测精度可提升 8.6%。文献[8-10]针对电力设备图像背景噪声复杂、设备轮廓模糊等问题,在 YOLOv5、Faster R-CNN 等目标检测模型基础上,采用改进骨干残差网络、优化区域提议网络等方法增强网络特征提取能力,同时利用马赛克数据增强、改进损失函数等方法来进一步提升电力设备缺陷检测的精度。上述研究对象主要集中于体积相对较小的输电线路部件或大部件的局部缺陷检测,而对于电力巡检图像中具有多角度、多尺度、多姿态和大长宽等特点的异常电力杆塔,需要有针对性地改进模型来实现不同倾斜程度的判别。文献[11]将杆塔类别分为两类正常和一类异常样本,并借助 YOLOv3 模型实现了电力杆塔的异常检测,但倾斜杆塔的水平包围框中包含了较多的背景噪声,经过多次卷积操作后将掩盖杆塔部分特征信息,影响判别精度。文献[12]针对巡检拍摄的杆塔图像背景复杂且正负样本极不均衡问题,提出了一种基于压缩激活改进的快速异常检测对抗网络(squeeze-and-excitation improved fast unsupervised anomaly detection with generative adversarial network, SE-f-AnoGAN),有效提升了异常电力杆塔的检测精度,但该方法在判别前需要预先定位并裁剪杆塔目标,这种"两步走"的方法将增加检测时间,可能无法满足实际应用需求。

基于上述分析可知,常规目标检测及其改进模型已无法满足电力杆塔倾斜程度高精度检测的需求,且仅考虑杆塔状态判别这一单一任务的模型也已不符合目前端到端模型的设计理念[13]。2017 年,Liu 等[14]提出的旋转目标检测器为提升具有旋转方向性目标的检测精度提供了新的解决思路,并在接下来的几年里被广泛应用于各种领域。文献[15]在基于区域的快速卷积神经网络(faster region-based convolutional neural network, Faster R-CNN)中引入带旋转角度的锚框,实现了方向自适应的遥感目标检测。文献[16]提出了一种由多尺度特征融合模块、特征精炼模块和头部双分支等部件构成的旋转目标检测网络(refined rotation RetinaNet, $R^3$det),进一步提升了遥感目标检测精度。文献[17]在回归损失函数中引入基于像素旋转交并比(pixels-intersection over union, PIoU)损失项,在一定程度上降低了旋转目标的定位误差。上述研究主要用于遥感目标检测任务,而对于电力杆塔倾斜程度检测这一特定任务,需要从任务功能实现、损失函数设计、模型构建等方面来进一步分析并进行有针对性地改进。

因此,本文提出了一种基于轻量级多头分支卡尔曼滤波交并比的 $R^3$det(multi-head Kalman-filter intersection over union refined rotation RetinaNet, Multi-Head-KF-$R^3$det)网络模型,用于无人机电力杆塔倾斜程度检测。该方法主要包括:1)在原始 $R^3$det 中添加倾斜程度分支,实现了电力杆塔倾斜程度和杆塔类别的双重分类以及电力杆塔目标的准确定位。2)在回归损失函数中引入基于卡尔曼滤波的旋转 IoU 损失项[18],借助高斯分布的优势提升了带有倾斜角度杆塔的定位精度,进而有助于多任务中倾斜程度的判别。3)在改进的 $R^3$det 基础上,设计了轻量化网络结构,基于 Ghost[19] 轻量化网络设计原理来严格控制模型参数量,可实现在精度略微下降的情况下大大减少模型参数量。最后,本文通过实验证明了改进模型的有效性提高了检测精度。

## 1 基于 $R^3$det 的旋转自适应定位检测

### 1.1 旋转自适应定位检测

电力杆塔倾斜程度检测是目前电力巡检目标检测领域的一大难点。与普通输电线路部件识别相比,电力杆塔图像背景复杂,杆塔具有多角度、多尺度、多姿态和大长宽比等特点,难以利用特有的先验知识和上下文信息来准确定位和判别不同倾斜程度的杆塔[20]。图 1(1)、(3)分别为因受不同程度自然灾害以及不同地形的影响,电力杆塔具有任意倾斜角度,导致在某些拍摄视角下其分布较为紧密。若直接用水平框标注,不仅会引入较多混淆目标的背景噪声,影响电力杆塔倾斜程度分类判别,还可能在通过非极大抑制(non-maximum suppression, NMS)去除冗余框时,由于两框之间的交并比超过阈值,出现相互抑制的情况,造成漏检或错检[21-22]。此外,相比于水平框,旋转框之间存在多种相对位置关系,如图 1(4)、(5)所示,难以采用框架中已有的算子直接计算预测框和真实框之间的旋转交并比,而现有近似计算方法表现性能不佳[23],导致电力杆塔的定位精度较低,进而影响多任务中其他任务(如倾斜程度判别)的执行。因此,如何将引入旋转锚框后的目标检测模型有效应用于电力巡检图像,实现电力杆塔等目标物体的高精度检测,是目前电力巡检研究中亟待解决的难点。

旋转框定位表示法可分为五参数标定法和八参数标定法两种[24],本文采用五参数标定法中的 OpenCV 表示法作为旋转框的参数预设,如图 1(6)所示。用五元组 $(x,y,w,h,\theta)$ 表示目标框的位置信息,其中,$x$,$y$ 为目标框的中心点坐标,$w$ 为与 $x$ 轴成锐角的其中一边,作为目标框的宽度,另一边 $h$ 则作为目标框的高度,$\theta$ 为 $w$ 与 $x$ 轴之间的夹角,其角度范围为$[-90°,0°]$。

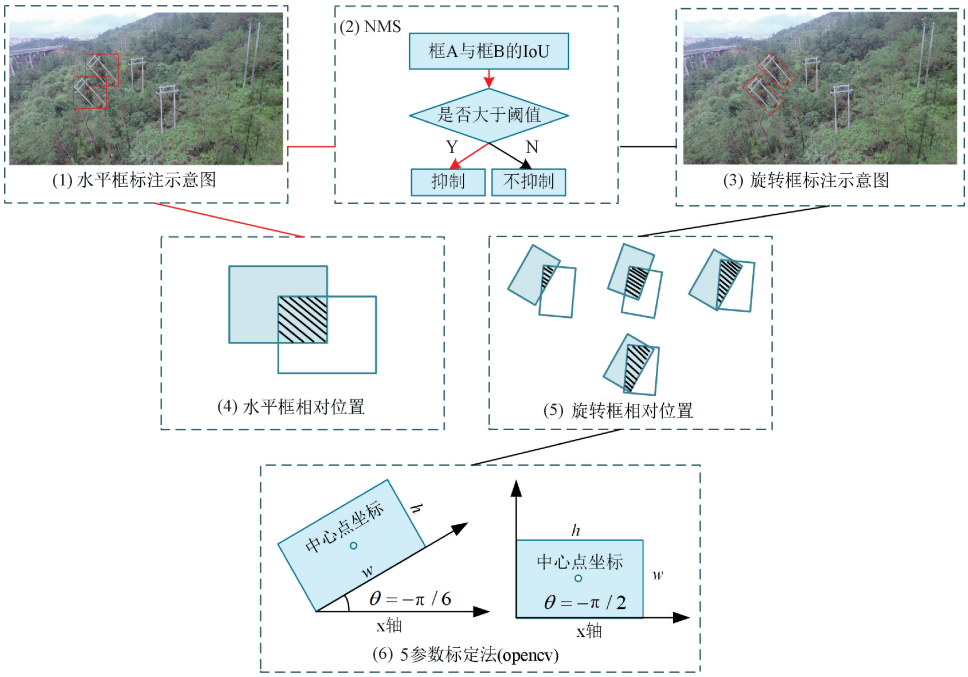



图 1　旋转框定位

Fig. 1　Positioning of rotation boxes

## 1.2　R³det 模型

R³det 是一种基于 RetinaNet[25] 的单阶段快速旋转目标检测器,主要应用于遥感影像、倾斜文本等多方向物体的检测,其模型结构如图 2 所示。区别于 RetinaNet 模型,该模型在头部双分支后加入了特征精炼模块,通过双路卷积、坐标点双线性插值法对特征图进行重构,使模型能够准确修正发生位置偏移的预测框,并结合逐级预测,降低了下游回归任务的难度。此外,为了解决旋转偏移角和中心位置相等而长宽比不等的情况,采用 smooth L1 方法计算可能得到相同损失结果的问题(如图 3(a)所示),在回归损失函数中引入了旋转 IoU 偏差项。R³det 损失函数 $L$ 可表示为:

$$L=\frac{\lambda_1}{N}\sum_{n=1}^{N} obj_n \frac{L_{\mathrm{reg}}(\boldsymbol{v}'_n,\boldsymbol{v}_n)}{|L_{\mathrm{reg}}(\boldsymbol{v}'_n,\boldsymbol{v}_n)|}|f(SkewIoU)|+\frac{\lambda_2}{N}\sum_{n=1}^{N}L_{\mathrm{cls}}(\boldsymbol{p}_n,\boldsymbol{t}_n) \tag{1}$$

$$L_{\mathrm{reg}}(\boldsymbol{v}',\boldsymbol{v})=L_{\mathrm{smooth-L1}}(\boldsymbol{v}'_\theta,\boldsymbol{v}_\theta)-\mathrm{IoU}(\boldsymbol{v}'_{\{x,y,w,h\}},\boldsymbol{v}_{\{x,y,w,h\}}) \tag{2}$$

式中:前一项为回归损失,后一项为类别损失;$obj_n$ 为 1 表示前景;$L_{reg}(\boldsymbol{v}'_n,\boldsymbol{v}_n)/|L_{reg}(\boldsymbol{v}'_n,\boldsymbol{v}_n)|$ 为回归损失的单位向量,用于控制梯度反向传播方向;$\boldsymbol{v}'$ 为预测框和锚框之间的偏移量;$\boldsymbol{v}$ 为真实框和锚框之间的偏移量;*SkewIoU* 为预测框和真实框之间的旋转交并比;$f(\cdot)$ 表示与 *SkewIoU* 相关的损失函数,用于控制反向传播时梯度的大小。$\boldsymbol{p}_n$ 为经过 *sigmoid* 函数后的概率分布向量,$\boldsymbol{t}_n$ 为目标类别的独热向量,$N$ 为旋转锚框的数量,$\lambda_1$ 和 $\lambda_2$ 为超参数,用于调整分类损失和定位损失的相对权重。当预测框和真实框之间的角度偏差较大时,对应的 SkewIoU 较小,经过 $f$ 函数反转后,可以加大定位损失项的惩罚力度,从而学到更精确的定位信息。$L_{reg}(\boldsymbol{v}',\boldsymbol{v})$ 具体可展开为式(2),其中前一项为角度偏差的 smooth L1 损失。为避免由于角度周期性问题使得先验锚框在回归预测框的过程中舍近求远,导致模型迭代训练过程中回归损失值发生突变(如图 3(b)所示),R³det 进一步引入了后一项 IoU 坐标损失。

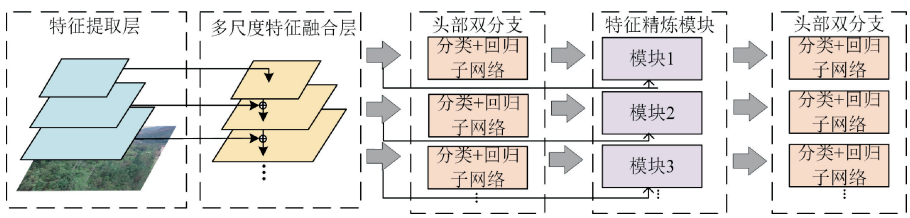



图 2　R³det 模型结构

Fig. 2　Model architecture of R³det

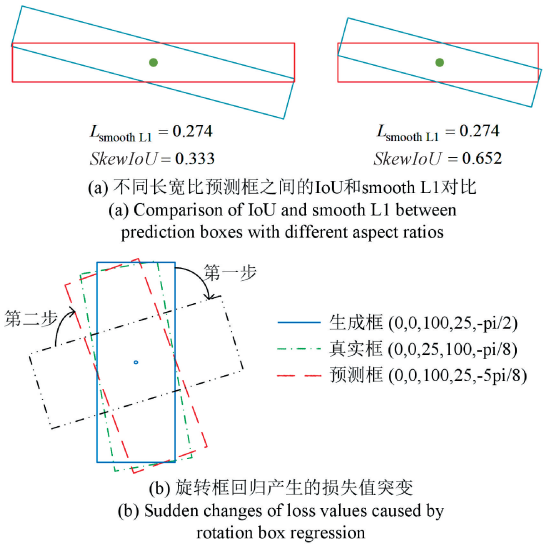



图 3 旋转框回归时可能出现的问题

Fig. 3 Possible issues during the rotation box regression

## 2 基于轻量级 Multi-Head-KF-R$^3$det 的无人机电力杆塔倾斜程度检测

### 2.1 Multi-Head-KF-R$^3$det 网络结构

本文在 R$^3$det 基础上，提出了用于电力杆塔倾斜程度检测的 Multi-Head-KF-R$^3$det。其网络结构如图 4 所示，该模型由特征提取层、多尺度特征融合层、特征精炼模块和多头任务分支网络构成。其中，特征提取网络中引入了残差结构，用于解决深层网络退化问题并提升网络的特征表达能力。多尺度特征融合层采用特征金字塔结构，在残差结构基础上，先自底向上映射并保留不同尺度的特征，然后自顶向下逐层传递深层语义信息，最后通过横向连接将浅层细节特征和深层语义特征有机结合起来，这在一定程度上可提升不同尺度杆塔的检测精度。多头分支网络由杆塔类别分支、定位分支和倾斜程度分支构成，利用参数共享机制共享特征提取网络的权重参

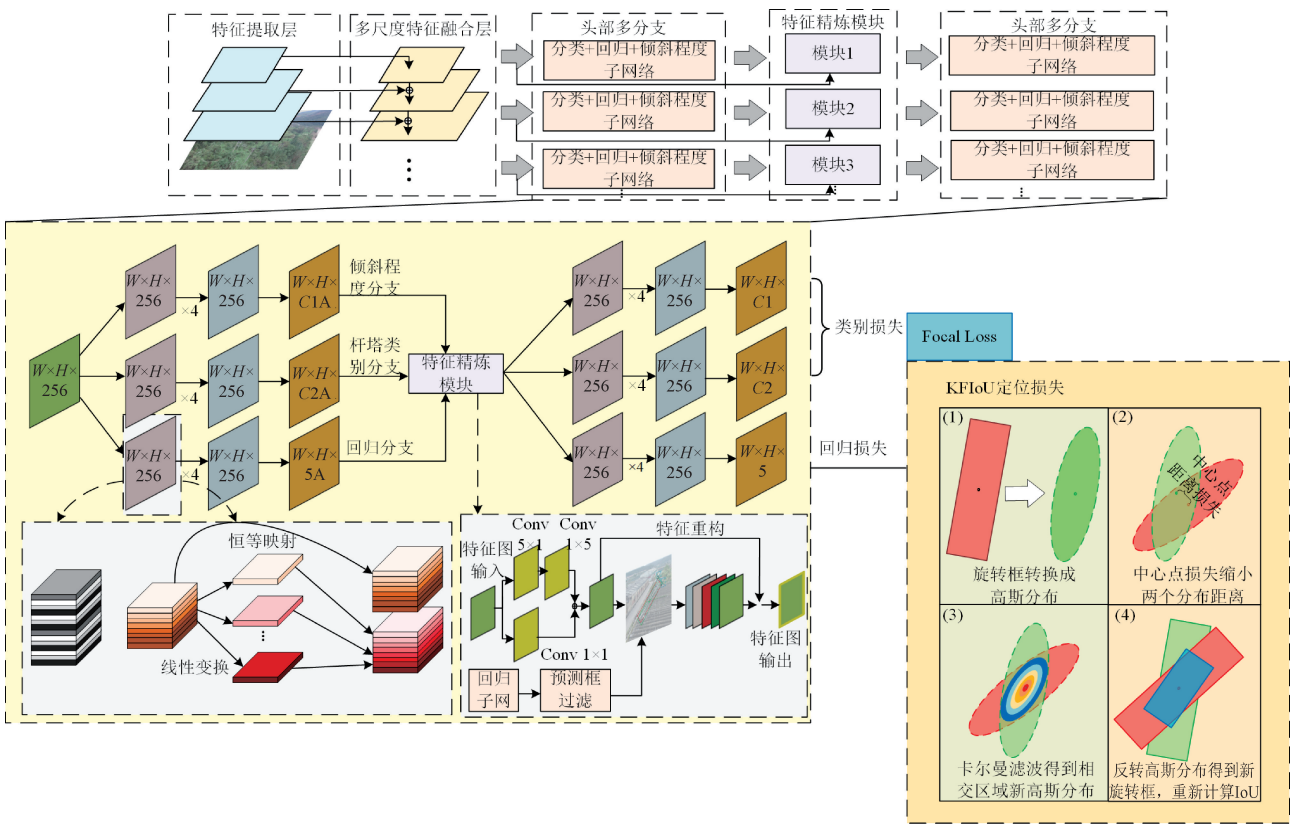



图 4 Multi-Head-KF-R$^3$det 模型结构

Fig. 4 Model architecture of Multi-Head-KF-R$^3$det

数，在反向传播时根据各任务预先设置的权重来有针对性地促进共享网络的特征融合。通过这种多任务学习方式，使某一任务可以为其他相关但又有区别的任务提供有价值的梯度噪声，当共享层聚合梯度时合理提高网络训练难度，将有助于进一步提升模型的泛化能力。在回归损失函数设计方面，借助高斯分布并引入卡尔曼滤波

思想来近似计算目标框和预测框之间的旋转交并比,解决了因现有旋转交并比计算精度低而使得目标定位不准确的问题,进而减轻了多任务中对倾斜程度判别带来的影响。此外,针对 Multi-Head-KF-$R^3$det 模型参数量多而不利于部署在无人机等嵌入式设备中的问题,本文将模型各部件依次替换为 Ghost 轻量化网络,利用少量卷积核以及线性映射方式提取关键特征,并以最低代价挖掘前景特征的幻影来补充冗余特征,弥补了因卷积数量的减少对模型特征提取带来的影响。

## 2.2　基于卡尔曼滤波交并比的损失函数

式(1)中的回归损失项虽然可以近似估计真实框和预测框之间的相对位置关系,但其梯度方向仍由坐标定位偏差主导,并没有准确引入旋转 IoU 的梯度方向。此外,旋转框之间的相对位置关系多样,采用现有近似方法计算精度较低,将影响多任务中其他任务(如倾斜程度判别)的表现性能。基于上述分析,本节以高斯分布为理论依据,设计了一个基于卡尔曼滤波思想的旋转 IoU 计算方法,并将其引入回归损失函数中指导模型训练。

图 5 显示了基于卡尔曼滤波的旋转 IoU 计算方法。具体步骤如下:

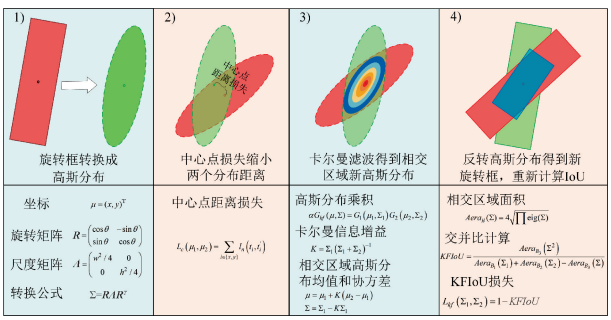



图 5　基于 KFIoU 的交并比计算方法

Fig. 5　Intersection over union algorithm based on KFIoU

1)首先,将旋转框转换为一个与之对应的二维高斯分布,使该旋转框成为该分布的最小外接矩形框,转换过程如下:

(1)构建二元高斯分布模型:

$$f(\boldsymbol{X})=\frac{1}{(2\pi)^{d/2}\left|\boldsymbol{\Sigma}\right|^{1/2}}\times \exp\left[-\frac{1}{2}(\boldsymbol{X}-\boldsymbol{\mu})^{\top}\boldsymbol{\Sigma}^{-1}(\boldsymbol{X}-\boldsymbol{\mu})\right],\ \boldsymbol{X}=(x_1,x_2,\cdots,x_n) \tag{3}$$

式中:$d$ 表示维度,取值为 2,$\boldsymbol{\mu}=(\mu_1,\mu_2,\cdots,\mu_n)$ 为各维度变量的均值,$\boldsymbol{\Sigma}=\begin{bmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{13} & \sigma_{14}\end{bmatrix}$ 为协方差矩阵,用于描述各维度变量之间的关系,$\sigma_{ij}$ 表示维度 $i$ 和维度 $j$ 之间的方差。

(2)对标准正态二维高斯分布进行线性变换,使变换后的二维高斯分布内切于旋转框,变换公式如下:

$$\boldsymbol{R}=\begin{bmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{bmatrix} \tag{4}$$

$$\boldsymbol{S}=\begin{bmatrix}S_w & 0\\ 0 & S_h\end{bmatrix} \tag{5}$$

$$\boldsymbol{A}=\boldsymbol{RS} \tag{6}$$

式中:$\boldsymbol{R}$ 为旋转矩阵,用于控制高斯分布的旋转角度,$\boldsymbol{S}$ 为尺度矩阵,用于控制二维高斯分布在两个维度上的缩放值,$\boldsymbol{A}$ 为线性变换矩阵,$\theta$ 为旋转角度,$S_w$ 和 $S_h$ 分别为 $w/2$ 和 $h/2$,$w$ 和 $h$ 分别为旋转框的长和宽。

(3)在步骤(2)基础上计算二维高斯分布的参数 $\boldsymbol{\mu}$ 和 $\boldsymbol{\Sigma}$:

$$\boldsymbol{\mu}=(x,y)^{\top} \tag{7}$$

$$\boldsymbol{\Sigma}=\boldsymbol{AA}^{\top}=\boldsymbol{RSS}^{\top}\boldsymbol{R}^{\top}=\boldsymbol{R\Lambda R}^{\top} \tag{8}$$

式中:$\boldsymbol{\Lambda}$ 为包含 $\boldsymbol{\Sigma}$ 特征值的对角矩阵,$(x,y)$ 表示旋转框的中心点坐标;

2)参照卡尔曼滤波矫正方法,将旋转目标检测中的预测框作为估计值,真实框作为观测值,重叠区域作为估计值和观测值之间的最优估计,利用预测框和真实框对应的两个高斯分布 $\mathcal{G}_1$,$\mathcal{G}_2$ 的乘积来计算新重叠区域的高斯分布 $\mathcal{G}_{kf}$:

$$\alpha\,\mathcal{G}_{kf}(\boldsymbol{\mu},\boldsymbol{\Sigma})=\mathcal{G}_1(\boldsymbol{\mu}_1,\boldsymbol{\Sigma}_1)\,\mathcal{G}_2(\boldsymbol{\mu}_2,\boldsymbol{\Sigma}_2) \tag{9}$$

式中:$\boldsymbol{\mu}_i$ 和 $\boldsymbol{\Sigma}_i$ 分别为高斯分布的均值和协方差矩阵,$\alpha$ 为常量系数,$\boldsymbol{\mu}=\boldsymbol{\mu}_1+\boldsymbol{K}(\boldsymbol{\mu}_2-\boldsymbol{\mu}_1)$,$\boldsymbol{\Sigma}=\boldsymbol{\Sigma}_1-\boldsymbol{K\Sigma}_1$,$\boldsymbol{K}=\boldsymbol{\Sigma}_1(\boldsymbol{\Sigma}_1+\boldsymbol{\Sigma}_2)^{-1}$ 为卡尔曼信息增益。可以发现,两个高斯分布的乘积结果为一个被系数 α 控制的新高斯分布,且该分布的协方差矩阵只和预测框和真实框对应的高斯分布有关。结合式(11)计算高斯分布对应的最小外接矩形框面积,由该式可知,若预测框和真实框对应的两个高斯分布的协方差矩阵保持不变,则相交区域的高斯分布反转为旋转框后,对应的旋转框面积并不会随着两个分布之间的距离$(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)$的改变而改变。为了解决上述问题,将系数 α 拆解为:

$$\alpha=\mathcal{G}_{\alpha}(\boldsymbol{\mu}_2,\boldsymbol{\Sigma}_1+\boldsymbol{\Sigma}_2)=\frac{1}{\sqrt{det(2\pi(\boldsymbol{\Sigma}_1+\boldsymbol{\Sigma}_2))}}\times \mathrm{e}^{-\frac{1}{2}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)^{\top}(\boldsymbol{\Sigma}_1+\boldsymbol{\Sigma}_2)^{-1}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)} \tag{10}$$

式中:引入了两个高斯分布中心点之间的距离 $(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)$。进一步,利用中心点损失函数 $L_c$(见式(15))的下降,减小$(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)$的值。当该值接近于0时,则α约等于1个常数,随后可根据重叠区域 $\mathcal{G}_{kf}(\boldsymbol{\mu},\boldsymbol{\Sigma})$ 中的协方差矩阵重新计算相交区域面积。

3)根据二维高斯分布的协方差矩阵,求解矩阵特征值,特征值累乘结果即为高斯分布对应最小外接旋转框

的面积：

$$\nu_{\mathcal{B}}(\boldsymbol{\Sigma}) = 2^n\sqrt{\prod \mathrm{eig}(\boldsymbol{\Sigma})} = 2^n \cdot |\boldsymbol{\Sigma}^{\frac{1}{2}}| = 2^n \cdot |\boldsymbol{\Sigma}|^{\frac{1}{2}} \tag{11}$$

式中：$n$ 为 $\boldsymbol{\Sigma}$ 的维度，$eig(\cdot)$ 为 $\boldsymbol{\Sigma}$ 的特征值。基于卡尔曼滤波交并比（KFIoU）最终可表示为：

$$KFIoU = \frac{\nu_{\mathcal{B}_3}(\boldsymbol{\Sigma})}{\nu_{\mathcal{B}_1}(\boldsymbol{\Sigma}_1) + \nu_{\mathcal{B}_2}(\boldsymbol{\Sigma}_2) - \nu_{\mathcal{B}_3}(\boldsymbol{\Sigma})} \tag{12}$$

式中：$\beta_1$，$\beta_2$，$\beta_3$ 分别为图 5 步骤 4）中的 3 个不同旋转框，在维度为 $n$ 的情况下，KFIoU 的值域为 $0 \leqslant KFIoU \leqslant 1/(2^{n/2+1}-1)$。在二维旋转目标检测中，KFIoU 的值域为[0，1/3]。

Multi-Head-KF-R$^3$det 损失函数设计如下：

$$L_{\text{total}} = \lambda_1 \sum_{n=1}^{N_{\text{pos}}} L_{\text{reg}}(\mathcal{G}(b_n), \mathcal{G}(gt_n)) + \frac{\lambda_2}{N}\sum_{n=1}^{N} L_{\text{cls}}(p_n, t_n) \tag{13}$$

$$L_{\text{reg}} = L_{\text{c}} + L_{\text{kf}} \tag{14}$$

$$L_{\text{kf}}(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2) = \mathrm{e}^{1-KFIoU} - 1 \tag{15}$$

$$L_{\text{c}}(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\Sigma}_1) = \ln((\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)^{\top}\boldsymbol{\Sigma}_1^{-1}(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1) + 1) \tag{16}$$

式（13）为整体损失，前一项为回归损失，由中心点损失 $L_{\text{c}}$ 和卡尔曼滤波损失 $L_{\text{kf}}$ 构成，其中中心点损失参考文献[26]中基于中心优化机制的 KLD 损失的第一项。$N$ 和 $N_{\text{pos}}$ 分别为所有锚框和正样本锚框的数量，$\mathcal{G}(b_n)$，$\mathcal{G}(gt_n)$ 分别为第 $n$ 个预测框和第 $n$ 个真实框的高斯分布函数；后一项为类别损失，采用焦点损失（focal loss）计算得到，$t_n$ 为第 $n$ 个目标的标签，$p_n$ 为经过 sigmoid 函数后的第 $n$ 个预测框的概率分布，$\lambda_1$ 和 $\lambda_2$ 为超参数，取值范围为[0.01，1]。

### 2.3 基于 Ghost 模块的轻量化网络

虽然 Multi-Head-KF-R$^3$det 相比于通用目标检测模型在杆塔定位、倾斜程度判别等多个分支的检测精度上有了一定程度的提升，但复杂的网络结构难以部署在资源有限的嵌入式设备中。其主要原因是原始 R$^3$det 中的特征精炼模块增加了网络堆叠层数，使网络参数量和计算量增加，模型推理速度下降。此外，不同层级网络之间需要根据预设的 IoU 阈值来进一步筛选符合前景条件的锚框，这将进一步增加额外的计算量。因此，本节采用 Ghost 轻量化模块替换 R$^3$det 各部件中的卷积层，如图 4 左下角所示。其原理为先用少量卷积提取关键特征，得到部分特征图：

$$\boldsymbol{Y}' = \boldsymbol{X} \times \boldsymbol{f}' \tag{17}$$

式中：$\boldsymbol{f}' \in R^{c\times k\times k\times m}$ 表示通道数为 $c$，大小为 $k\times k$，数量为 $m$ 的卷积核，其中，$m \leqslant n$，为了方便解释，这里省略了偏差项。$\boldsymbol{Y}' \in R^{h'\times w'\times m}$ 表示通道数为 $m$，大小为 $h'\times w'$ 的输出特征图。然后对 $\boldsymbol{Y}'$ 中的每张特征图进行 $s$ 次线性变换得到幻影特征图：

$$\boldsymbol{y}_{ij} = \boldsymbol{\Phi}_{i,j}(\boldsymbol{y}'_i), \forall i = 1,\cdots,m, j = 1,\cdots,s \tag{18}$$

式中：$\boldsymbol{y}'$ 为 $\boldsymbol{Y}'$ 中第 $i$ 个通道的特征图，$\boldsymbol{\Phi}_{i,j}$ 为第 $i$ 个特征图对应的第 $j$ 个线性运算，$\boldsymbol{y}_{ij}$ 为第 $i$ 个特征图经过第 $j$ 个线性运算得到的幻影特征图，最终特征图数量为 $m\times s = n$，该值与采用 $n$ 个卷积核输出特征图数量一致。

不考虑偏差项的情况下，普通卷积的参数量为 $n\times c\times k\times k$（卷积核数量为 $n$），Ghost 模块的第一步卷积参数量为 $n\times c\times k\times k = n/s\times c\times k\times k$，表示 $m$ 个通道数为 $c$，大小为 $k\times k$ 的卷积核对应的参数量，第二步产生冗余特征图所需的参数量为 $m\times(s-1)\times d\times d = n/s\times(s-1)\times d\times d$，表示 $m$ 个输入特征图经过 $(s-1)$ 个线性变换后得到大小为 $n\times w\times h$ 的输出特征图，每个线性变换对应的卷积核大小为 $d\times d$。因此，普通卷积与 Ghost 模块卷积的参数压缩比可表示为：

$$r_c = \frac{n\times c\times k\times k}{\frac{n}{s}\times c\times k\times k + (s-1)\times\frac{n}{s}\times d\times d} \approx \frac{s\times c}{s+c-1} \approx s \tag{19}$$

式中：$k$ 和 $d$ 均表示卷积核的大小，一般设置为 3×3 或 5×5，线性变换数量 $s$ 远小于卷积通道数 $c$。经过近似化简后，得到两者压缩比约等于 $s$。同理，两者的理论计算加速比 $r_s$ 也约等于 $s$。

### 2.4 算法流程

Multi-Head-KF-R$^3$det 的训练和检测流程如下。

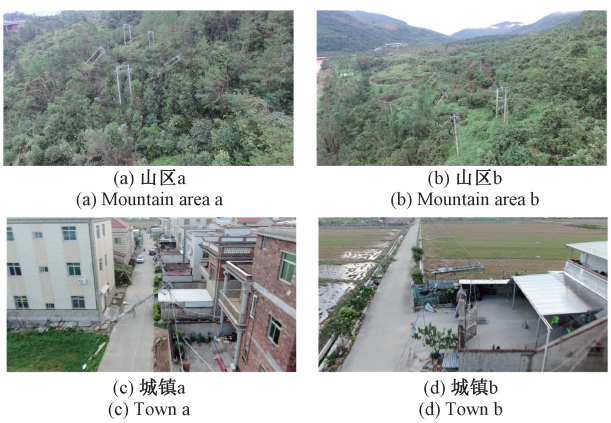
1）模型训练：首先，对无人机电力杆塔图像进行预处理操作，包括图像裁剪和数据增强两部分（见 3.1 节）。接着，将预处理后的图像依次输入到特征提取网络、多尺度特征融合网络和多头分支网络，初步得到所有杆塔预测锚框。然后，借助特征精炼网络实现锚框筛选以及预测锚框与真实框之间的坐标特征对齐。将对齐后的特征图进一步输入到多头分支网络，根据其预测结果分别计算杆塔类别、倾斜程度和杆塔定位损失值，并通过反向传播更新网络参数。重复上述操作，直至达到指定迭代次数，保存模型参数，结束训练。

2）杆塔检测：首先，输入无人机电力杆塔原图，并用滑动窗口裁剪图像。然后，将预处理后的杆塔图像依次经过训练好的特征提取网络、多尺度特征融合网络、多头分支网络（初步预测）、特征精炼网络和多头分支网络（最终预测），得到杆塔类别、倾斜程度以及位置信息。

## 3　实验结果与分析

### 3.1　实验数据集

为验证 Multi-Head-KF-R³det 相比于其他目标检测模型的优越性和有效性,本文基于无人机电力杆塔倾斜程度检测开展了实验研究。实验采用多旋翼无人机在巡线模式下拍摄的电力杆塔图片,数据集共包含了 2 610 张图片。其中,山区图片 2 041 张,图像分辨率为 1 291 pixels×726 pixels;城镇图片 569 张,图像分辨率为 3 840 pixels×2 160 pixels,如图 6 所示。杆塔累计数量为 6 380 个,其中正常杆塔数量为 5 376 个,轻度异常杆塔数量为 248 个,重度异常杆塔数量为 756 个。由于原始无人机巡检图像分辨率较大,若直接将原图输入网络,则用于存储中间结果的内存将呈爆炸式增长,甚至超出硬件内存上限,影响模型训练。因此,本文采用大小为 600 pixels×600 pixels,步长为 450 pixels 的滑动窗口对原图进行滑动切割,然后将得到的裁剪块大小调整至 800 pixels×800 pixels。该方法在减小图片尺寸的同时,还能最大限度避免杆塔截断。经裁剪后,图片共计 8 721 张,其中 6 647 张图片用于训练,2 047 张图片用于测试。此外,以 0.5 的概率随机水平翻转裁剪块用于数据增强。



(a) 山区a　(b) 山区b
(a) Mountain area a　(b) Mountain area b
(c) 城镇a　(d) 城镇b
(c) Town a　(d) Town b

图 6　山区与城镇电力杆塔

Fig. 6　Mountain and urban power towers

### 3.2　不同模型评估结果

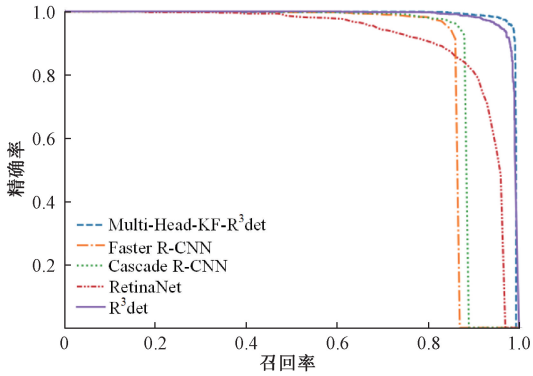
Multi-Head-KF-R³det 模型总共训练 40 个轮次(epoch),每个 epoch 中共包含了 3 556 次不同图片的迭代,初始学习率为 0.004,当训练到 30、35 个 epoch 时,学习率分别缩小为原来的 1/10,特征精炼次数设置为 1。

基于第 2 节的改进方法,将常规目标检测和旋转目标检测模型作为对照组,初步测试了 R³det 改进前后的性能。如图 7 所示,分别为不同模型在正常、轻度异常和重度异常这 3 个类别中的 PR 曲线。如图 7(a)、(b)、(c)所示,实线、虚线均位于另外 3 条曲线的右上方且曲线下方面积更大,说明 R³det 相比于常规目标检测模型,在电力杆塔倾斜程度检测上的性能更优。特别在轻度异常检测中,如图 7(b)所示,常规目标检测模型对应的双点虚线、单点虚线、点线 3 条曲线均有向左下方收缩的趋势,而基于 R³det 的旋转目标检测模型对应的曲线仍保持在最右上角的位置。当引入基于卡尔曼滤波的回归损失项后,PR 曲线会略微向右上方膨胀,表明卡尔曼滤波损失这一改进有助于模型性能的提升。由图 7(d)、(e)、(f)可知,其他旋转目标检测模型对应的 PR 曲线均位于 R³det 的左下方,且双点虚线、单点虚线、点线 3 条曲线中的折点较多,说明其他模型可能产生了较多无用的检测框,导致模型精度较低。

如表 1 所示,展示了用 mAP 指标定量评估常规目标检测模型、旋转目标检测模型以及 R³det 模型改进前后的结果。mAP 值可以避免单一评价指标(如精确率或召回率)对结果的偏向性,能较为全面地评价模型的分类和回归能力。其中,$AP_n$,$AP_{ab1}$,$AP_{ab2}$ 分别为正常、轻度异常、重度异常的 AP 值。由表可知,原始 R³det 的整体 mAP 值相比于常规目标检测和旋转目标检测模型,均有不同程度的提升,分别提高了 9.2%、6.6%、5.2%、14.4%、12.6%、10.8%,这表明 R³det 网络结构的设计以及引入旋转目标框的优越性。当引入多头分支后,虽然轻度倾斜类别的 AP 值略微下降,但在重度倾斜类别检测中,其性能提高了 3.2%,整体性能提升了 0.8%,验证了通过多头分支引入多任务信息有助于模型泛化能力的提升。在此基础上,在回归损失项中引入基于卡尔曼滤波的旋转框交并比计算方法,使模型整体精度可达到 94.5%,其中,正常、轻度倾斜、重度倾斜的 AP 值分别为 90.8%、98.7%、94.1%,验证了基于卡尔曼滤波交并比损失函数的有效性。

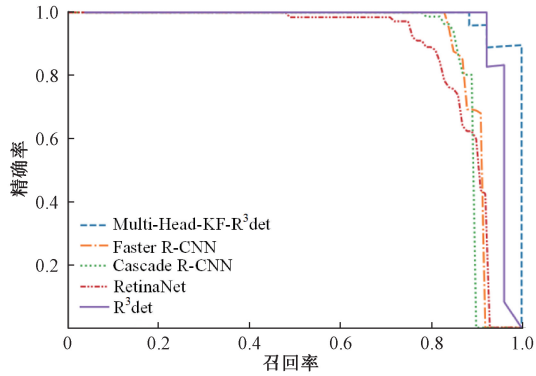
表 2 则进一步展示了 R³det 改进前后的召回率,其平衡点的计算方法为:计算 PR 曲线与 $y=x$ 的交点,作为精确率与召回率达到相对最佳的平衡点。表中,$Recall_n$,$Recall_{ab1}$,$Recall_{ab2}$ 分别为正常、轻度异常、重度异常的召回率。由表可知,R³det 改进前后的召回率均接近于 95%,说明这些模型可以检测出绝大多数杆塔,其漏检率较低。当引入多头分支和卡尔曼滤波交并比损失后,模型召回率有微小提升,分别提升了 0.1% 和 0.5%。

为了更直观地检验 Multi-Head-KF-R³det 在杆塔定位、杆塔类别以及倾斜程度方面的检测能力,本节分别选取了部分山区和城镇图片进行可视化检测,结果如图 8 所示。在山区场景中,对于不同倾斜程度、倒断、倒塌等多种姿态的异常杆塔,Multi-Head-KF-R³det 不仅可以准确判别各个杆塔的倾斜程度和杆塔类型,还能定位检



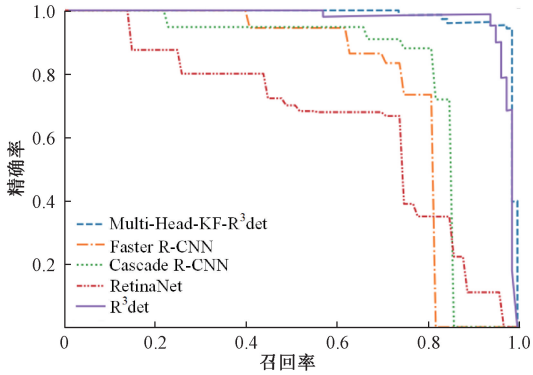

(a) 改进前后的$R^3$det与常规目标检测模型的PR曲线（正常）

(a) PR curve of $R^3$det and conventional object detection model before and after improvement (normal)



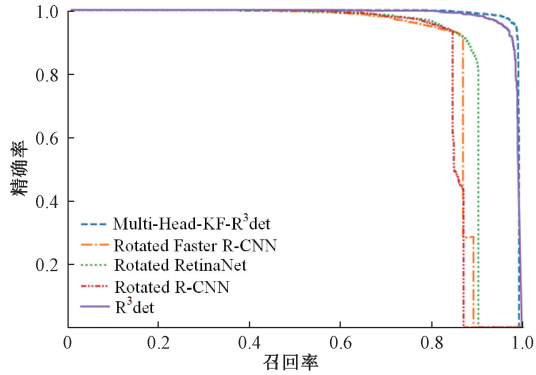

(b) 改进前后的$R^3$det与常规目标检测模型的PR曲线（轻度异常）

(b) PR curve of $R^3$det and conventional object detection model before and after improvement (mildly abnormal)



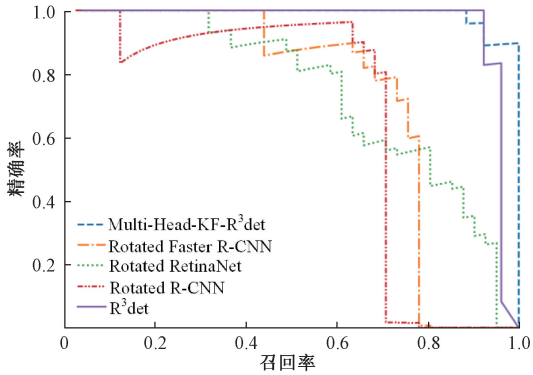

(c) 改进前后的$R^3$det与常规目标检测模型的PR曲线（重度异常）

(c) PR curve of $R^3$det and conventional object detection model before and after improvement (severely abnormal)



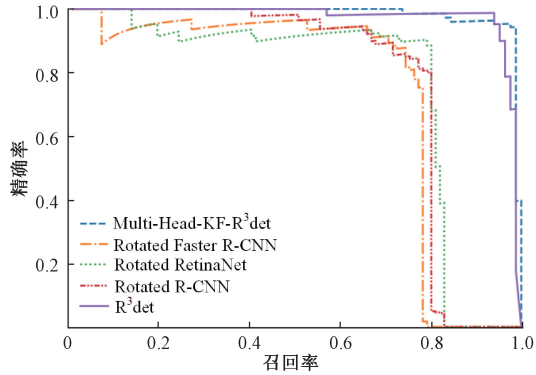

(d) 改进前后的$R^3$det与旋转目标检测模型的PR曲线（正常）

(d) PR curve of $R^3$det and rotation object detection model before and after improvement (normal)




(e) 改进前后的$R^3$det与旋转目标检测模型的PR曲线（轻度异常）

(e) PR curve of $R^3$det and rotation object detection model before and after improvement (mildly abnormal)




(f) 改进前后的$R^3$det与旋转目标检测模型的PR曲线（重度异常）

(f) PR curve of $R^3$det and rotation object detection model before and after improvement (severely abnormal)

图 7 不同模型的 PR 曲线

Fig. 7 PR curves of different models

测出部分小目标杆塔，如图 8(a)所示，且检测框的旋转方向与杆塔倾斜朝向保持一致，说明 Multi-Head-KF-$R^3$det 具备了当前通用目标检测框架所缺少的多功能检测能力，同时也显示出该模型对小目标检测定位的优

**表 1　不同模型 mAP 值对比**

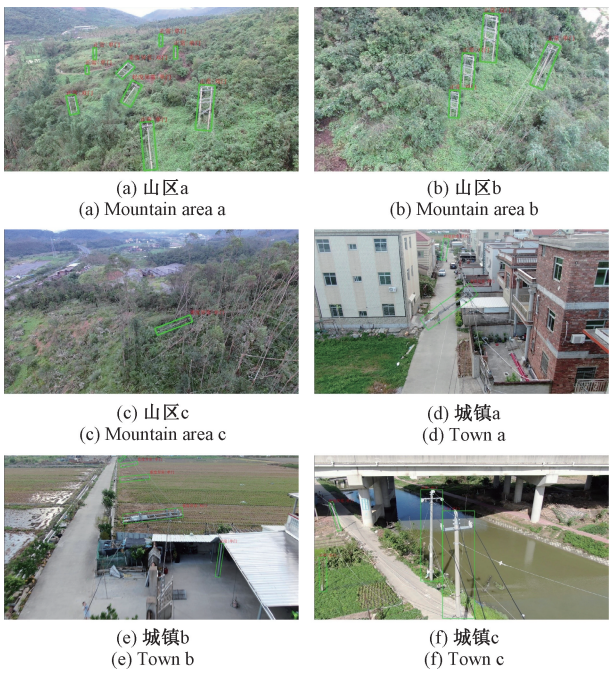
**Table 1　Comparation of mAP values among different models**　%

| 模型类型 | 模型 | mAP | $AP_n$ | $AP_{ab1}$ | $A_{pab2}$ |
|---|---|---|---|---|---|
| 常规目标检测模型 | RetinaNet | 81.5 | 91.7 | 87.5 | 65.3 |
| | Faster R-CNN | 84.1 | 86.0 | 89.7 | 76.6 |
| | Cascade R-CNN | 85.5 | 88.2 | 88.2 | 80.2 |
| 旋转目标检测模型 | Rotated R-CNN | 76.3 | 81.2 | 69.3 | 78.4 |
| | Rotated Faster R-CNN | 78.1 | 81.0 | 76.0 | 77.3 |
| | Rotated RetinaNet | 79.9 | 88.9 | 73.7 | 77.0 |
| 改进模型 | 原始 R³det | 90.7 | 90.7 | 90.9 | 90.5 |
| | Multi-Head-R³det | 91.5 | 90.7 | 90.2 | 93.7 |
| | Multi-Head-KF-R³det | **94.5** | **90.8** | **98.7** | **94.1** |

**表 2　R³det 改进前后召回率对比**

**Table 2　Comparation of recall rates before and after R³det improvement**　%

| 模型 | Recall | $Recall_n$ | $Recall_{ab1}$ | $Recall_{ab2}$ |
|---|---|---|---|---|
| 原始 R³det | 94.3 | 95.6 | 92.3 | 95.2 |
| Multi-Head-R³det | 94.4 | 95.8 | 92.3 | 95.2 |
| Multi-Head-KF-R³det | **94.9** | **97.2** | **92.3** | **95.2** |



(a) 山区a (a) Mountain area a　(b) 山区b (b) Mountain area b

(c) 山区c (c) Mountain area c　(d) 城镇a (d) Town a

(e) 城镇b (e) Town b　(f) 城镇c (f) Town c

图 8　不同场景下的可视化检测结果

Fig. 8　Visual detection results in different scenarios

越性。在城镇场景中,可以基本检测出图片中的所有杆塔,并能准确地判别杆塔类别和倾斜程度,但也出现了部分错检。例如如图 8(e)所示,将钢铁材质的柱子误认为单门正常杆塔,说明城镇场景中包含的内容信息更加丰富多样,对杆塔检测造成了一定的干扰,导致与杆塔具有相似材质和形态的其他目标物被错检。

最后,为了得到一个轻量化模型,压缩 Multi-Head-KF-R³det 中各部分网络,计算得到不同压缩网络的参数量、浮点运算量(floating point operations, Flops)和 mAP 值,如表 3 所示。对比模型压缩前结果可知,当压缩比为 2 时,压缩不同部件后的 mAP 值均有提升,骨干网络(backbone)、头部网络(head)、特征精炼头部网络(refine-head)以及三者组合的压缩网络分别提升了 1.1%、1.2%、1.6%、0.8%,说明基于 Ghost 模块的轻量化网络

**表 3　模型压缩前后参数量和运算量结果**

**Table 3　The parameter and computational results before and after model compression**

| GhostNet 添加位置 | 压缩比 | 参数量/M | FLOPs/GMacs | mAP/% |
|---|---|---|---|---|
| 压缩前 | – | 54.55 | 269.53 | 94.5 |
| Backbone | 2 | 43.00 | 244.55 | 95.6 |
| | 3 | 39.19 | 236.45 | 94.1 |
| | 4 | 37.24 | 232.19 | 89.4 |
| Head | 2 | 51.02 | 222.46 | 95.7 |
| | 3 | 49.86 | 207.03 | 93.5 |
| | 4 | 49.26 | 198.94 | 96.5 |
| Refine-Head | 2 | 47.49 | 222.46 | 96.1 |
| | 3 | 45.18 | 207.03 | 94.2 |
| | 4 | 43.97 | 198.94 | 95.6 |
| Backbone+Head+Refine-Head | 2 | 32.41 | 150.40 | 95.3 |
| | 3 | 25.13 | 111.44 | 89.8 |
| | 4 | 21.37 | 91.00 | 92.1 |
| | 5 | **19.26** | **79.84** | 89.1 |

设计通过适度减少卷积个数并引入非线性变换来增加特征多样性,有助于模型精度的提升,但随着压缩比的增加,卷积数量的减少,mAP 值波动下降。其中,backbone 的压缩对模型精度影响最大,当压缩比为 4 时,其 mAP 值相比压缩前降低了 5.1%,说明 backbone 对图像整体特征提取有更重要的作用。由表 3 最后一行可知,当 backbone、head、refine-head 组合网络压缩为原来的 1/5 时,mAP 值为 89.1%,参数量可减少为 19.26 M,FLOPs 最终降为 79.84 GMacs。在实际使用中,可根据具体模型以及应用场景来设置压缩比。

## 4 结 论

本文针对无人机巡检图像中的电力杆塔具有多角度、多尺度、多姿态和大长宽比等特点,提出了一种基于 Multi-Head-KF-$R^3$det 的电力杆塔倾斜程度检测新模型,本模型的改进及性能提升包括:

1)通过在原始 $R^3$det 中添加倾斜程度分支,实现了电力杆塔类别和倾斜程度的判别以及电力杆塔目标的准确定位。同时,借助多任务学习中相关任务特征共享、信息互通的优势,使模型整体 mAP 值相比改进前提升了 0.8%,重度倾斜杆塔 mAP 值提升了 3.2%。

2)通过构建旋转框的二维高斯分布,将基于卡尔曼滤波的旋转交并比损失项引入回归损失函数中用于指导模型训练,可使电力杆塔整体样本的检测精度和召回率分别达到 94.5% 和 94.9%。

3)通过将 Multi-Head-KF-$R^3$det 各部件的卷积依次替换为 Ghost 轻量化网络,可使模型在压缩为原来 1/5 的情况下,其检测精度仍高于常规目标检测模型和旋转目标检测模型,为 Multi-Head-KF-$R^3$det 在嵌入式设备中的部署奠定基础。

本算法模型适用于山区和城镇场景下的电力杆塔倾斜程度检测,未来可将注意力机制、杆塔上下文信息特征融合等算法加入 Multi-Head-$R^3$det 模型中,用于聚焦杆塔检测和减少背景噪声干扰,从而进一步提高模型检测精度和鲁棒性。

## 参考文献


[1] LI Z, ZH Y, WU H, et al. Design and application of a UAV autonomous inspection system for high-voltage power transmission lines[J]. Remote Sensing, 2023, 15(3): 865-888.

[2] XU C, LI Q, ZHOU Q, et al. Power line-guided automatic electric transmission line inspection system[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 3512118.

[3] 吴华, 梁方正, 刘草, 等. 输电杆塔弱纹理部件的可迁移式检测[J]. 仪器仪表学报, 2021, 42(6): 172-178.
WU H, LIANG F ZH, LIU C, et al. Transferable detection for low texture components of transmission tower[J]. Chinese Journal of Scientific Instrument, 2021, 42(6): 172-178.

[4] 王宏超, 邵云峰, 马中静, 等. 基于深度学习增强的 LSD 杆塔倾斜度检测[J]. 电子测量与仪器学报, 2021, 35(9): 204-213.
WANG H CH, SHAO Y F, MA ZH J, et al. Tower tilt detection based on LSD enhancement by deep learning[J]. Journal of Electronic Measurement and Instrumentation, 2021, 35(9): 204-213.

[5] 刘志颖, 缪希仁, 陈静, 等. 电力架空线路巡检可见光图像智能处理研究综述[J]. 电网技术, 2020, 44(3): 1057-1069.
LIU ZH Y, MU X R, CHEN J, et al. Review of visible image intelligent processing for transmission line inspection[J]. Power Grid Technology, 2020, 44(3): 1057-1069.

[6] 孙备, 党昭洋, 吴鹏, 等. 多尺度互交叉注意力改进的单无人机对地伪装目标检测定位方法[J]. 仪器仪表学报, 2023, 44(6): 54-65.
SUN B, DANG ZH Y, WU P, et al. Multi scale cross attention improved method of single unmanned aerial vehicle for ground camouflage target detection and localization[J]. Chinese Journal of Scientific Instrument, 2023, 44(6): 54-65.

[7] 宋立业, 刘帅, 王凯, 等. 基于改进 EfficientDet 的电网元件及缺陷识别方法[J]. 电工技术学报, 2022, 37(9): 2241-2251.
SONG L Y, LIU SH, WANG K, et al. Identification method of power grid components and defects based on improved efficientDet [J]. Transactions of China Electrotechnical Society, 2022, 37(9): 2241-2251.

[8] 郝帅, 杨磊, 马旭, 等. 基于注意力机制与跨尺度特征融合的 YOLOv5 输电线路故障检测[J]. 中国电机工程学报, 2023, 43(6): 2319-2331.

HAO SH, YANG L, MA X, et al. YOLOv5 Transmission line fault detection based on attention mechanism and cross-scale feature fusion [J]. Proceedings of the CSEE, 2023, 43(6): 2319-2331.

[9] 刘传洋, 吴一全. 基于深度学习的输电线路视觉检测方法研究进展[J]. 中国电机工程学报,2023,43(19): 7423-7446.
LIU CH Y, WU Y Q. Research progress of vision detection methods based on deep learning for transmission lines[J]. Proceedings of the CSEE, 2023, 43(19): 7423-7446.

[10] 李文璞, 谢可, 廖逍, 等. 基于 Faster RCNN 变电设备红外图像缺陷识别方法[J]. 南方电网技术, 2019, 13(12): 79-84.
LI W P, XIE K, LIAO X, et al. Intelligent diagnosis method of infrared image for transformer equipment based on improved faster RCNN[J]. Southern Power System Technology, 2019, 13(12): 79-84.

[11] 郭敬东, 陈彬, 王仁书, 等. 基于 YOLO 的无人机电力线路杆塔巡检图像实时检测[J]. 中国电力, 2019, 52(7): 17-23.
GUO J D, CHEN B, WANG R SH, et al. Real time detection of UAV power line tower inspection image based on Yolo [J]. China Power, 2019, 52(7): 17-23.

[12] 仲林林, 胡霞, 刘柯妤, 等. 基于改进生成对抗网络的无人机电力杆塔巡检图像异常检测[J]. 电工技术学报, 2022, 37(9): 2230-2240,2262.
ZHONG L L, HU X, LIU K Y, et al. Power tower anomaly detection from unmanned aerial vehicles inspection images based on improved generative adversarial network [J]. Transactions of China Electrotechnical Society, 2022, 37(9): 2230-2240,2262.

[13] 蒲天骄,张中浩,谈元鹏,等. 电力人工智能技术理论基础与发展展望(二):自主学习与应用初探[J]. 中国电机工程学报,2023,43(10):3705-3718.
PU T J, ZHANG ZH H, TAN Y P, et al. Theoretical primer and directions of electric power artificial interlligence (II): Self-directed learning and preliminary application [J]. Proceedings of the CSEE, 2023, 43(10):3705-3718.

[14] LIU L, PAN Z, LEI B. Learning a rotation invariant detector with rotatable bounding box[J]. ArXiv Preprint, 2017, ArXiv:171109405.

[15] YANG X, YANG J, YAN J, et al. Scrdet: Towards more robust detection for small, cluttered and rotated objects[C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 8232-8241.

[16] YANG X, YAN J, FENG Z, et al. $R^3$det: Refined single-stage detector with feature refinement for rotating object[C]. Proceedings of the AAAI Conference on Artificial Intelligence, Vancouver, 2021: 3163-3171.

[17] CHEN Z, CHEN K, LIN W, et al. Piou loss: Towards accurate oriented object detection in complex environments[C]. Computer Vision-ECCV 2020: 16th European Conference, 2020: 195-211.

[18] YANG X, ZHOU Y, ZHANG G, et al. The KFIoU loss for rotated object detection[J]. ArXiv Preprint, 2022, ArXiv: 220112558.

[19] HAN K, WANG Y, TIAN Q, et al. Ghostnet: More features from cheap operations[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 1580-1589.

[20] WANG R, ZHANG S, CHEN B, et al. Unmanned aerial vehicle (UAV) vision-based detection and location of power towers for transmission line maintenance[C]. The 16th IET International Conference on AC and DC Power Transmission, 2020: 1937-1941.

[21] GONE M, WANG D, ZHAO X, et al. A review of non-maximum suppression algorithms for deep learning target detection [C]. Seventh Symposium on Novel Photoelectronic Detection Technology and Applications, 2021: 821-828.

[22] TANG X, XIE X, HAO K, et al. A line-segment-based non-maximum suppression method for accurate object detection[J]. Knowledge-Based Systems, 2022, 251: 108885.

[23] YANG X, YAN J. On the arbitrary-oriented object detection: Classification based approaches revisited[J]. International Journal of Computer Vision, 2022, 130(5): 1340-1365.

[24] DAI Y, YU J, ZHANG D, et al. ROD Former: High-precision design for rotating object detection with transformers[J]. Sensors, 2022, 22(7): 2633.

[25] 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, 2017:

2980-2988.

[26] YANG X, YANG X, YANG J, et al. Learning high-precision bounding box for rotated object detection via kullback-leibler divergence [J]. Advances in Neural Information Processing Systems, 2021, 34: 18381-18394.

## 作者简介

**胡霞**,2020 年于南京工程学院获得学士学位,现为东南大学硕士研究生,主要研究方向为人工智能技术在电气领域中的应用。
E-mail: huxia@seu.edu.cn

**Hu Xia** received her B. Sc. degree from Nanjing Institute of Technology in 2020. She is currently a master student at Southeast University. Her main research interests include the application of artificial intelligence in the electrical engineering.

**仲林林**(通信作者),2012 年于西安交通大学获得学士学位,2017 年于西安交通大学和法国图卢兹第三大学获得博士学位,现为东南大学副研究员/博士生导师,主要研究方向为高电压技术、放电等离子体技术、人工智能技术。
E-mail: linlin@seu.edu.cn

**Zhong Linlin** (Corresponding author) received his B. Sc. degree from Xi′an Jiaotong University in 2012, and received his Ph. D. degree from Xi′an Jiaotong University and Université Toulouse III - Paul Sabatier in 2017. He is currently an associate professor and a Ph. D. advisor at Southeast University. His main research interests include the interdisciplinary study of high voltage engineering, plasma engineering, and artificial intelligence.