

基于改进 RetinaNet 的电力设备红外目标精细化检测模型

苏海锋, 赵岩, 武泽君, 程博, 吕林飞

(华北电力大学 电力工程系, 河北 保定 071000)

摘要: 电力设备在运行过程中会产生大量红外图像, 当红外图像中的电力设备存在排列密集、具有倾斜角度、大长宽比的情况时, 基于水平矩形框的目标检测网络只能给出目标概略位置, 易发生目标检测区域重叠, 引入冗余背景信息, 使得检测结果不够精细。针对此问题, 提出在 RetinaNet 目标检测网络中引入旋转矩形框机制, 并在网络输入端引入 Mosaic 数据增强技术; 将原特征提取网络中 ReLU 函数替换为梯度流更平滑的 Mish 激活函数; 在原模型 FPN 模块后追加 PAN 模块进一步融合图像特征。最后利用现场采集的电力设备红外图像制作数据集, 将改进后的模型与 Faster R-CNN、YOLOv3、原 RetinaNet 三种基于水平矩形框定位的目标检测网络进行对比评估, 实验表明改进后的模型可以更为精细地检测出密集场景下带有倾角的电力设备红外目标, 在多类别电力设备检测准确率对比上高于以上 3 种模型。

关键词: 红外图像; RetinaNet; 电力设备; 卷积神经网络; 目标检测;

中图分类号: TM85 文献标志码: A 文章编号: 1001-8891(2021)11-1104-08

Refined Infrared Object Detection Model for Power Equipment Based on Improved RetinaNet

SU Haifeng, ZHAO Yan, WU Zejun, CHENG Bo, LYU Linfei

(Department of Electrical Engineering, North China Electric Power University, Baoding 071000, China)

Abstract: A large number of infrared images are generated during the operation of power equipment. When the power equipment in the infrared image is densely arranged, incline-angled, and has a large aspect ratio, the target detection network based on a horizontal rectangular frame can only provide the approximate position of the target, which is prone to overlap with the target detection area and introduce redundant background information, giving detection results that are not sufficiently accurate. To solve this problem, we propose to introduce a rotating rectangular box mechanism into the retina net target detection network and mosaic data enhancement technology at the network input, replacing the ReLU function in the original backbone network with a smoother mish activation function of gradient flow; the Pan module is added after the FPN module of the original model to further fuse image features. Finally, the data set is made by using the power-equipment infrared images collected on-site. The improved model is compared and evaluated with three target detection networks based on horizontal rectangular frame positioning: fast R-CNN, YOLOv3, and original RetinaNet. The experiments show that the improved model can detect the infrared targets of power equipment with inclination in dense scenes more accurately, and the detection accuracy of multi-category power equipment is higher than that of the above three models.

Key words: infrared image, RetinaNet, power equipment, convolutional neural network, object detection;

0 引言

电气设备的可靠运行直接关系到电力系统的安

全与稳定, 其在故障早期往往会出现温度异常的现象。红外热成像技术作为一种非接触式的测温技术, 可以全天候地获取电力设备温度分布信息并形成热

收稿日期: 2021-09-08; 修订日期: 2021-10-16.

作者简介: 苏海锋 (1977-), 男, 河北石家庄人, 讲师, 博士, 主要研究方向为智能配电网研究。E-mail: hfsups@163.com.

通信作者: 赵岩 (1996-), 男, 河北保定人, 硕士研究生, 主要研究方向为电气设备智能检测。E-mail: 1191755813@qq.com.

基金项目: 国家重点研发计划 (2017BJ0080)。

谱图进行显示^[1],及时有效地发现设备过热缺陷。目前该技术已广泛地应用到了变电站电气设备运行状态检测当中,但其产生的大量红外图像仍然要靠人工进行故障判别与分析^[2],效率低下、主观性强,难以提升设备状态检测过程的实时性与可靠性。因此利用快速、准确地自动检测算法对海量的红外图像数据进行分析成为当下研究的重点^[3]。而对红外图像中出现的多种电气设备进行精确的定位与识别,是后续进行设备故障诊断与精准报警的基础^[4]。

近年来深度学习的发展使得越来越多的学者将基于深度学习的目标检测网络引入到电气设备的红外图像检测中^[5-7]。目前主流的目标检测深度学习算法可分为两大类:①Two-stage 检测算法,此类方法分两步进行,首先利用 RPN (region proposal network) 等结构生成候选区域,之后对候选区域进行目标分类与位置调整。代表作有 Faster R-CNN^[8];②One-stage 检测算法,这类方法无需单独生成候选区域,只通过全卷积的结构即可得出目标的分类与边框回归参数。典型方法有 YOLOv3 (you only look once)^[9]、SSD (single shot multibox detector)^[10]等。One-stage 类算法较 Two-stage 类算法速度提升明显,但准确率低于 Two-stage 类算法。文献[11]利用 Faster R-CNN 对红外图像进行处理,实现了变压器、套管、断路器等7种电气设备的定位与识别,取得了较高的识别精度。文献[12]基于改进 YOLOv3 算法对5种高压开关设备红外图像进行识别,并达到了对异常发热点进行检测的目的。

目前用于电气设备红外图像目标检测的深度学习算法主要利用水平矩形框对物体进行定位与识别,在电力设备密集、大长宽比、多角度的场景下水平矩形框无法保证足够的定位精细度,甚至会框选到临近设备,造成识别错位。当红外图像中的电气设备目标存在一定旋转角度,尤其是接近 45°倾角时,水平矩形框易引入与目标无关的大量背景冗余信息^[13],不利于后续基于设备区域颜色值进行的状态评估操作。针对水平框定位不精细的问题,本文提出基于改进 RetinaNet^[14]的旋转目标精细化检测方法,利用带有角度的旋转矩形框对倾斜密集的电气设备红外目标进行精细化的定位,以减少目标框选过程中的干扰物影响。同时为进一步提升模型在红外图像中的检测准确率,在原 RetinaNet 网络中引入 Mosaic^[15]数据增强算法、Mish^[16]激活函数和路径聚合网络(Path Aggregation Network, PAN)^[17]进行优化。最后在隔离开关、断路器、绝缘子串、电压互感器、电流互感器5类纹理特征理相似,大长宽比,易因拍摄等问题产生倾斜角度

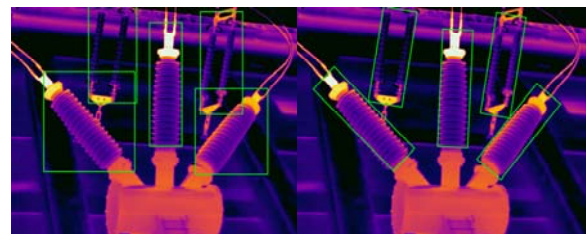
的设备中进行对比实验,证明本文所提模型在提升定位精细化程度上的有效性以及检测准确率上的优势。

1 RetinaNet 目标检测模型

RetinaNet 属于 One-stage 方法,该网络利用 Focal Loss 损失函数降低了图像中大量简单背景信息所占权重,突出图像目标区域正样本信息,解决了 One-stage 目标检测算法中存在的正负样本比例失衡的问题,从而使网络专注于难分样本的训练。Focal Loss 的利用使得 RetinaNet 网络在 COCO 数据集上的检测精度首次高于 Two-stage 算法^[14],且保有 One-stage 算法的速度优势。RetinaNet 目标检测网络主要包括特征提取网络、特征金字塔结构 (Feature Pyramid Networks, FPN)^[18]、预测网络3部分。特征提取部分采用深度残差网络 ResNet101 来进行图像特征的初步提取,残差结构的使用避免了因网络层次过深引起的梯度消失与梯度爆炸问题;FPN 模块对特征提取网络输出的特征图进行进一步的特征融合,从单一分辨率的图片构建出特征信息丰富、多尺度的特征金字塔,最终输出包含5个不同尺度目标区域信息的特征图集合;预测网络分为目标分类子网络与边框位置回归子网络两条支路,对5个特征图进行全卷积^[19]操作,其中目标分类子网负责预测目标的类别置信度分布,位置回归子网用于预测目标边框的位置信息。

2 改进 RetinaNet 的红外目标检测模型

为达到对红外图像中电气设备精细化检测的目的,本文主要对 RetinaNet 进行4个方面的改进:①在原网络中引入旋转框机制,实现在大长宽比、倾斜密集场景下的电气设备红外目标精细化定位。旋转框与水平框定位效果对比如图1所示;②在网络的图像



(a) 水平框 (b) 旋转框
(a) Horizontal box (b) Rotation box

图1 水平框与旋转框定位效果对比

Fig.1 Comparison of positioning effect between horizontal box and rotation box

输入端增加 Mosaic 数据增强算法,进一步丰富红外图像背景信息,增加样本的多样性并提升模型的泛化能力;③在特征提取网络引入 Mish 激活函数,用于

增强梯度变化的平滑性, 避免原 ReLU^[20] 激活函数导致负值梯度消失的问题; ④最后采用 FPN+PAN 结构

进行多尺度特征融合, 充分利用多层特征图的不同信息。改进后的网络结构如图 2 所示。

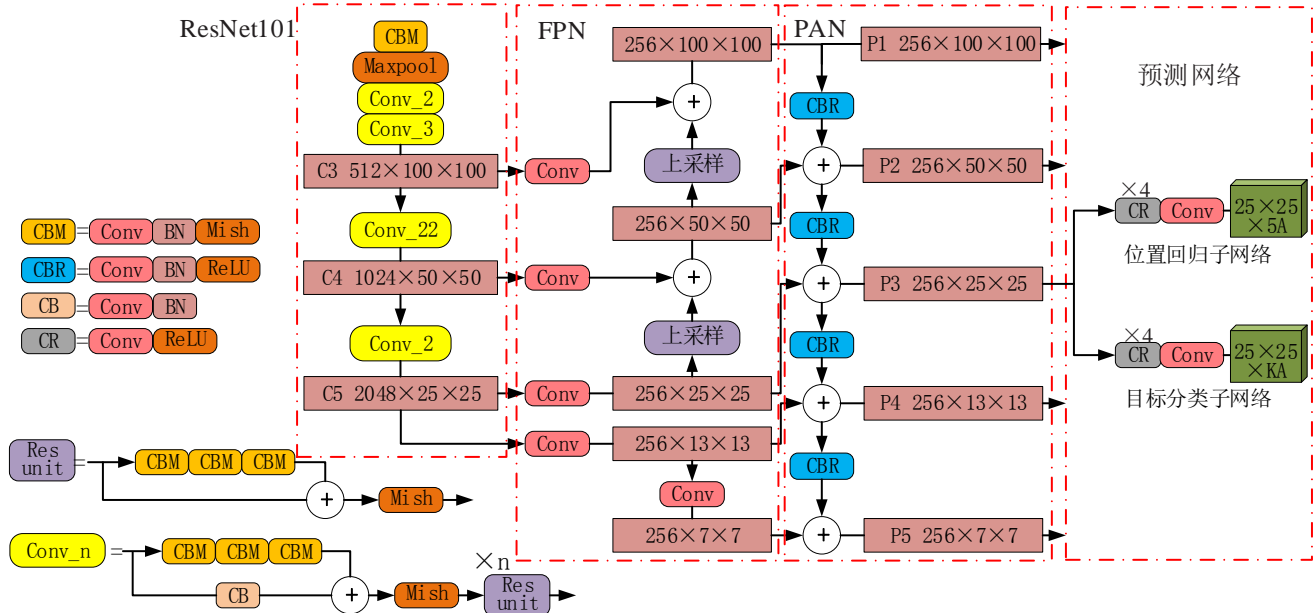
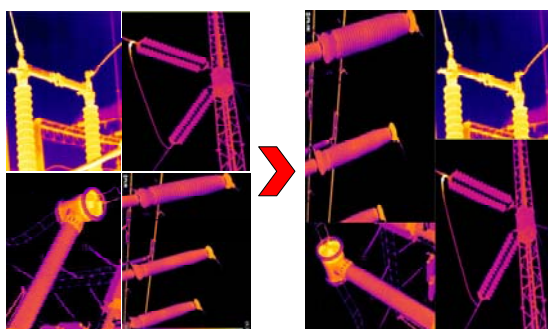


图 2 改进 RetinaNet 网络结构

Fig.2 The network architecture of improved RetinaNet mode

2.1 Mosaic 数据增强

改进后的 RetinaNet 旋转目标检测网络采用 Mosaic 数据增强方法处理输入端电气设备红外图片, Mosaic 数据增强方法每次读取 4 张红外图片, 首先对每张图片进行随机缩放、裁剪、色域变换、水平翻转中的一种或多种操作, 之后将处理过的 4 张图片进行随机排布拼接成一张图片, 如图 3 所示。将拼接好的



(a) 原始红外图像 (b) Mosaic 增强后图像
(a) Original infrared image (b) Mosaic enhanced image

图 3 Mosaic 数据增强

Fig.3 Mosaic data augmentation

红外图片以及标签数据送入神经网络进行进一步的特征提取。Mosaic 数据增强技术通过对多张图片的拼接操作丰富了待检测电气设备目标的背景信息, 使得改进 RetinaNet 网络对复杂背景的红外图像检测适应性更强, 对图片的缩放操作进一步丰富了图像内部小目标电气设备样本的数目, 增强了模型对不同尺度红

外目标的鲁棒性。该方法一次处理 4 张图片的数据, 因而 mini-batch 大小不需要很大, 降低了训练门槛, 单 GPU 即可达到较好的训练效果^[15]。

2.2 Mish 激活函数

原 ResNet101 网络中使用 ReLU 激活函数, 其表达式如下:

$$\text{ReLU}(x) = \max(0, x) \quad (1)$$

相比于 sigmoid 和 tanh 激活函数, ReLU 激活函数存在正区间不饱和的优势, 当 $x > 0$ 时激活值输出始终大于零, 解决了正区间 x 值过大进入饱和区造成的梯度消失的问题。但当 $x \leq 0$ 时激活值恒为 0, 对应导数为 0, 存在着负区间梯度消失的缺点, 在 $x < 0$ 时梯度无法继续更新导致神经元死亡。为解决 ReLU 函数存在的问题, 本文将 ReLU 函数替换为 Mish 激活函数, 具体表达式如下:

$$\text{Mish}(x) = x \tanh(\ln(1 + e^x)) \quad (2)$$

相较于 ReLU 函数, Mish 函数在负值的时候并非完全截断, 如图 4 所示, 而是允许一定的负值梯度存在, 从而稳定了网络的梯度流, 其平滑的特性能有效地允许电气设备图像信息深入地渗透到神经网络中。Mish 函数保证每一点的平滑性, 使得训练过程中梯度下降效果优于 ReLU 函数。随着特征提取网络层数的加深, 利用 Mish 激活函数替代 ReLU 激活函数取得了比原网络更高的测试准确率与训练稳定性^[16]。

2.3 PAN 模块

本文在原 RetinaNet 的 FPN 模块后面新增 PAN 模块, 电气设备红外图像输入 ResNet101 特征提取网络后, 得到不同深度的特征图, 不同深度的特征图所含特征信息有所不同, 深层特征图携带有更强的电气设备语义特征信息和较弱的目标位置信息, 而浅层的特征图含有丰富的电气设备目标位置特征与较弱的语义特征。FPN 模块采用自顶向下进行上采样的方式将深层特征图与浅层特征图进行叠加融合, 以达到向浅层特征图传递深层语义特征的目的, 但浅层丰富的目标位置信息并未传递到深层特征图。为了进一步融合电气设备红外图像深层语义特征与浅层目标位置特征, 在 FPN 模块后添加 PAN 模块, PAN 结构对 FPN 输出的特征图进行自底向上的下采样操作, 通过卷积模块调整特征图尺寸与通道数, 之后逐级与深层特征图叠加融合, 把浅层的丰富的目标位置特征传递到深层的特征图。FPN 和 PAN 结构的组合充分利用了红外图像不同尺度特征图的信息, 避免了 One-stage 方法利用单一尺度特征图无法同时有效地表征各个尺度的物体的缺点, 提高了模型对电气设备红外图像的特征利用能力。

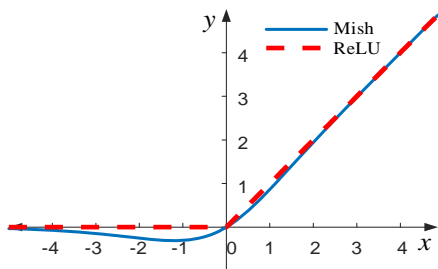


图4 两种激活函数

Fig.4 Two activation functions

2.4 预测网络与损失函数

RetinaNet 目标检测算法利用水平矩形框对物体进行定位识别, 因此在电气设备红外图像目标检测中的应用存在一定的局限性。当红外图像中的电气设备排列密集时, 一个矩形框可能包含多个目标, 矩形框之间的重叠区域较大; 在电气设备存在倾斜角度时, 水平矩形框会引入较多背景噪声信息, 影响后续图像处理工作。针对以上局限性, 本文将 RetinaNet 网络的水平矩形框调整为带有旋转角度的矩形框, 并对预测子网络与损失函数做出调整。

2.4.1 旋转矩形框机制

本文定义的旋转矩形框如图 5 所示, 矩形框中与 x 轴正方向成锐角的边定义为 h , 另一条边定义为 w , 角度 θ 为边 h 与 x 轴正方向所成的锐角, 取值区间为 $[-\pi/2, 0)$ 。旋转矩形框以 (x, y, w, h, θ) 位置向量的形式给

出, x, y 表示矩形框中心点坐标, w, h 代表矩形框的宽度和高度。

输入 RetinaNet 网络的图片大小为 800×800 , 在经过特征提取、特征图融合后输出 7×7 、 13×13 、 25×25 、 50×50 、 100×100 五个不同尺度的特征图来预测不同尺度的物体, 5 个特征图分别记为 $P_1 \sim P_5$ 。特征图在原图的映射点称为锚点, 以每个锚点为中心形成不同大小的先验框 (anchor)。先验框的面积尺度在 $P_1 \sim P_5$ 上分别为 $\{512^2, 256^2, 128^2, 64^2, 32^2\}$, 原 RetinaNet 模型对每个尺度的先验框添加 $\{2^0, 2^{1/3}, 2^{2/3}\}$ 3 种不同尺度的面积比例因子与 $\{1/2, 1/1, 2/1\}$ 3 种不同的长宽比例因子, 每个锚点生成 9 个不同的先验框。由于电气设备多具有大长宽比的特点, 本文将原来 3 种长宽比拓展为 $\{1:1, 1:2, 2:1, 1:3, 3:1, 5:1, 1:5\}$ 7 种长宽比, 增加了更大的长宽比例因子以适应电气设备红外目标的定位框选, 并对每个先验框引入 $\{-\pi/2, -5\pi/2, -\pi/3, -\pi/4, -\pi/6, -\pi/12\}$ 6 种角度变换因子, 实现对带有倾斜角度的红外电气设备精细化定位。每个锚点由原模型的 9 个水平矩形先验框增加为 126 个带有角度的先验框, 改进后的先验框策略如图 6 所示, 先验框的位置坐标向量由 $(x_a, y_a, w_a, h_a, \theta_a)$ 表示。

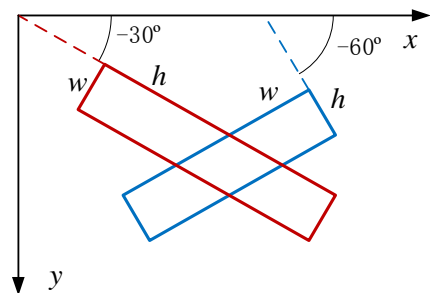


图5 旋转矩形框示意图

Fig.5 Schematic of the rotating rectangular box

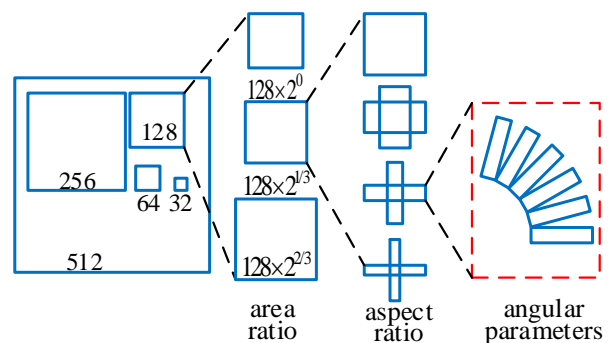


图6 本文使用的先验框策略

Fig.6 Anchor strategy in our method

2.4.2 预测网络

特征提取部分输出的 $P_1 \sim P_5$ 五个不同尺度特征

图分别进入对应的预测网络,如图 2 右侧所示。预测网络分为目标分类子网络与位置回归子网络两条支路,两支路均进行 4 个连续的卷积操作来提取特征,每次卷积操作后使用 ReLU 作为激活函数,最后通过一个卷积层调整特征图通道数。回归子网络负责预测每个锚点对应预测框的位置参数,原 RetinaNet 利用水平框对物体进行框选,只需要预测 (x', y', w', h') 4 个位置参数来表示预测框中心点位置与宽高,每个锚点周围生成 9 个预测框,故回归子网络输出 $W \times H \times 4A$ 的张量,其中 A 的值为 9, $W \times H$ 等于输入特征图对应锚点数量。改进 RetinaNet 采用旋转矩形框,在得到预测框中心点位置与宽高的基础上还需要增加额外的倾斜角度变量,需要预测 (x', y', w', h', θ) 5 个旋转框位置预测量,由于每个锚点生成 21 个预测框,每个框含有 6 个旋转角度因子,此时每个锚点生成 126 个大小、角度不同的框,故调整回归子网络最后一层卷积的通道数由 $4A$ 变为 $5A$,其中 A 为 126。分类子网络负责估计每个预测框的种类置信度分布,每个锚点对应预测框数量由 9 变为 126,故调整分类子网络最后一层卷积的通道数由 $K \times 9$ 变为 $K \times 126$,输出 $W \times H \times 126K$ 的张量, K 为待识别的电气设备种类数。本文总共识别隔离开关、断路器、绝缘子串、电压互感器、电流互感器 5 类电气设备,因此 K 值取 5。

2.4.3 损失函数

RetinaNet 损失函数分为分类损失与回归损失两部分。由于加入了角度预测量,改进 RetinaNet 回归损失函数在原损失函数上加入了角度这一回归变量。真实标注框 (Ground Truth, GT) 与先验框的相对坐标以及预测框与先验框的相对坐标可表示为:

$$\begin{cases} t_x = (x - x_a) / w_a \\ t_y = (y - y_a) / h_a \\ t_w = \lg(w / w_a) \\ t_h = \lg(h / h_a) \\ t_\theta = \theta - \theta_a \end{cases} \quad (3)$$

$$\begin{cases} t'_x = (x' - x_a) / w_a \\ t'_y = (y' - y_a) / h_a \\ t'_w = \lg(w' / w_a) \\ t'_h = \lg(h' / h_a) \\ t'_\theta = \theta' - \theta_a \end{cases} \quad (4)$$

式中: $(x_a, y_a, w_a, h_a, \theta_a)$ 、 $(x', y', w', h', \theta')$ 、 (x, y, w, h, θ) 分别表示先验框、预测框、真实标签框的位置参数及倾斜角; $(t_x, t_y, t_w, t_h, t_\theta)$ 表示标签框与先验框的相对坐标; $(t'_x, t'_y, t'_w, t'_h, t'_\theta)$ 为预测框与先验框的相对坐标。

边框位置回归损失函数通过预测框与真实标签框的差异计算损失值的大小,本文使用 smooth L1 函数,定义如下:

$$SL_1(t_i - t'_i) = \begin{cases} 0.5(t_i - t'_i)^2 & |t_i - t'_i| < 1 \\ |t_i - t'_i| - 0.5 & \text{其他} \end{cases} \quad (5)$$

式中: t_i 表示标签框与先验框的相对坐标,取值范围为 $(t_x, t_y, t_w, t_h, t_\theta)$, t'_i 表示预测框与先验框的相对坐标,取值范围为 $(t'_x, t'_y, t'_w, t'_h, t'_\theta)$ 。

总损失函数 L 为回归子网络损失与分类子网络损失加权,如下式:

$$L = \frac{\lambda_1}{N} \sum_{n=1}^N t'_n \sum_{i \in \{x, y, w, h, \theta\}} SL_1(t'_i, t_{ni}) + \frac{\lambda_2}{N} \sum_{n=1}^N L_{cls}(p_n, t_n) \quad (6)$$

式中: N 表示先验框数量; t'_n 有两个取值,当先验框为前景时 t'_n 取 1,当先验框为背景时取 0。 t_{ni} 表示第 n 个预测框相对于先验框的坐标, t_{ni} 表示第 n 个先验框相对于真实标签框的坐标。 p_n 表示分类子网络预测的第 n 个先验框的多类别置信度分布经过 Sigmoid 函数计算后的值, t_n 为第 n 个先验框对应真实目标的类别标签。 L_{cls} 表示类别判定损失,使用原 RetinaNet 的 Focal Loss 损失函数计算;参数 λ_1 、 λ_2 默认取 1。

3 实验结果与分析

实验环境为 Ubuntu16.04 系统,利用 Pytorch 框架搭建网络结构。计算机硬件配置为 GeForce GTX1080Ti 11-GB GPU, CPU 为 i7-10700, 16 GB 内存。利用 GPU 进行训练和测试过程的加速。

3.1 实验数据

本文采集我国西北某电力公司多个变电站的运检红外图像,主要包括隔离开关、断路器、绝缘子串、电压互感器 (potential transformer, PT)、电流互感器 (current transformer, CT) 5 类常见电气设备,部分红外图像中的电气设备存在排列密集、倾斜角度较大的问题。对原图片进行增强和扩充预处理,包括随机缩放填充、随机平移翻转、随机调整色调亮度等操作,最终得到 3000 张红外图像,统一调整为 800×800 大小,其中包括 897 个断路器、4219 个绝缘子串、1710 个隔离开关,831 个电压互感器、597 个电流互感器。利用 roLabelImg 工具对所有红外图像中的电气设备进行旋转矩形框标注制成数据集。最终对 3000 张红外图像随机选取,2400 张用于模型的训练,600 张用于模型测试。

3.2 训练过程

由于目前样本数量有限,本文采用迁移学习的方法训练新模型。迁移学习可将一个领域内学习到的知识迁移到另一个相似领域,充分利用了原始数据集和目标数据集之间潜在相同特征或者结构,可以明显加速网络收敛^[21]。在网络训练时,载入利用 PASCAL VOC 2007 数据集训练得到的权重,将网络参数迁移过来进行红外图像数据集的训练,训练时初始学习率设置为 $5 \times e^{-4}$,动量为 0.9,权重衰减为 0.0001,采用 Adam 优化器, batch size 大小为 4,共进行 30000 次迭代。图 7 为前 20000 训练过程中损失值随迭代次数增加而变化的曲线。训练开始时损失值随训练迭代次数的增加而迅速下降,收敛速度较快。当迭代次数超过 5000 时损失值下降明显变慢,迭代 10000 次以后损失值几乎不再下降,维持在 0.6 左右震荡,此时模型的训练达到收敛。

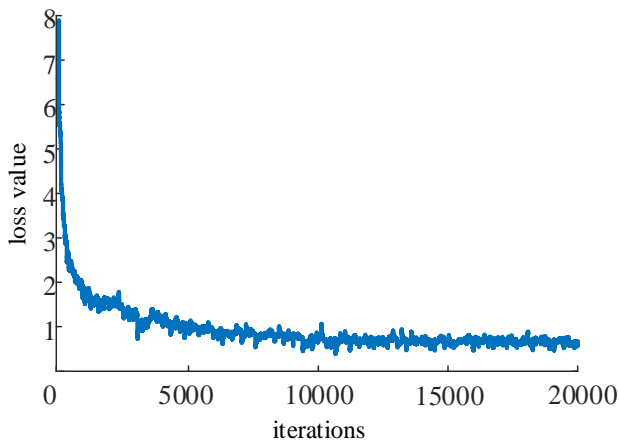


图7 损失曲线
Fig.7 Loss curve

3.3 性能比较与结果分析

为验证本文改进 RetinaNet 网络的检测精度,将所提算法与当前主流的目标检测器进行对比,包括 Faster R-CNN、YOLOv3 以及原 RetinaNet 网络。以上网络均在相同的电气设备红外数据集进行测试评估,实验条件以及训练参数完全相同。评估指标包括平均准确率(Average Precision, AP)、平均准确率均值(mean AP, mAP)。其中 AP 值是评判某一类电气设备目标检测效果的指标, mAP 是所有 5 类别电气设备的平均准确率均值。AP 与 mAP 定义如下:

$$\begin{cases} P = \frac{TP}{TP + FP} \\ R = \frac{TP}{TP + FN} \end{cases} \quad (7)$$

$$AP = \int_0^1 P(R) dR \quad (8)$$

$$mAP = \frac{\sum_{n=1}^4 AP(n)}{4} \quad (9)$$

式中: TP 表示正确分类的正样本个数; FP 表示错误标记成正样本的负样本个数; FN 表示将正样本错误判为负样本的个数; P、R 分别为查全率与查准率。表 1 给出了 4 种不同网络模型在 IoU 阈值为 0.5 时各个不同种类电气设备的 AP 值以及 mAP 值对比结果。

表 1 不同检测模型对比测试结果

Table 1 Comparison of the test results of different detection models

Method	AP				mAP	
	Breaker	Insulator	Switch	PT	CT	
Faster R-CNN	94.47	89.21	87.23	96.45	95.44	92.56
YOLOv3	90.62	86.52	82.09	92.03	91.37	88.53
RetinaNet	94.96	90.05	88.57	96.03	96.19	93.16
Ours method	97.51	92.84	90.61	98.69	97.86	95.50

由表 1 可知,在相同的 IoU 阈值下,本文所提模型在隔离开关、断路器、绝缘子、电压互感器、电流互感器 5 类设备的 AP 值均高于其余 3 种目标检测网络,改进后模型在测试集上的 mAP 值高于原 RetinaNet 网络 2.34%,对比结果证明了本模型在检测准确率上的优势以及本文所提改进措施对原模型准确率的提升。此外,在测试集的 513 张存在倾斜角度、排列密集的电气设备图像中,由于改进 RetinaNet 网络采用了带有角度的旋转框进行目标定位,取得了比其余 3 种目标检测网络更加精细的定位效果。在具有较大长宽比且存在倾斜角度的电气设备红外目标图像检测中,带有旋转角度的矩形框相较于水平矩形框引入了更少的冗余背景温度信息,降低了后续热故障区域判别及提取算法的数据处理量,减小了因背景异常温度干扰而产生的误报警几率,使得基于电气设备红外图像的热故障诊断算法可以有更快的处理速度以及更高的故障判别准确率。基于旋转矩形框的目标检测区域之间相互重叠的部分更少,这也为后续的电气设备红外图像分离操作提供便利。部分电气设备红外图像的定位效果对比如图 8 所示,(a)、(d)图中绝缘子与电流互感器存在较大的倾斜角度,利用水平矩形框框选引入了大量与目标无关的背景图像,而利用本文方法则能够精细地识别出其中的红外设备。(b)、(c)图由于设备存在倾斜角度且排列密集,导致矩形框之间发生了重叠,难以进行后续图像分离操作,本文基于旋转矩形框进行目标定位,避免了电气设备排列密集时水平框重复框选、相互重叠的缺点,实现了更精细的目标检测效果。

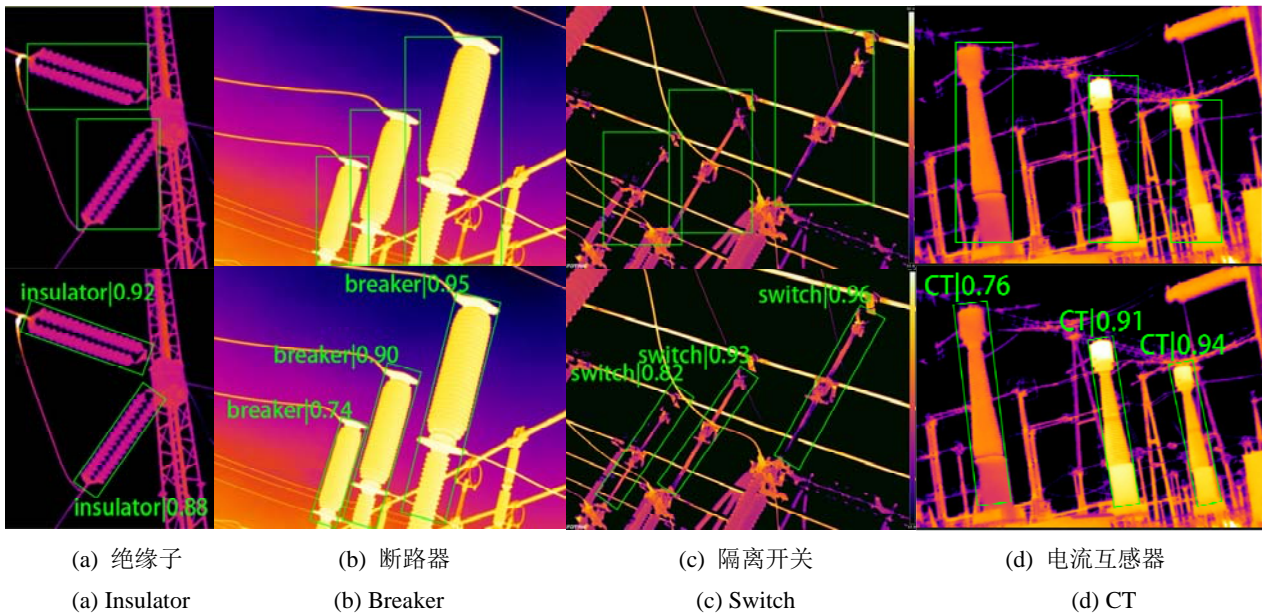


图 8 模型检测结果

Fig.8 Model test results

4 结论

本文提出一种基于改进 RetinaNet 的电气设备红外图像检测模型,将原网络的水平矩形框定位改为带有角度的旋转矩形框进行目标的定位。并且在网络的输入端、特征提取部分、特征融合部分均作出相应改进。改进后的模型在隔离开关、断路器、绝缘子串、电压互感器、电流互感器 5 类常见电气设备构成的数据集中进行了对比试验,结果表明:

1) 本文所提模型在 5 类常见电气设备上每一类的检测精度均达到 90% 以上,取得了较优的检测结果。

2) 在 IoU 阈值为 0.5 时,本文所提方法平均准确率均值高达 95.50%,高于当前主流的 Faster R-CNN、YOLOv3 目标检测算法,且优于原 Retinate 网络。

3) 本文采用旋转矩形框定位机制,在红外图像中的电气设备存在排列密集、有倾斜角度的情况下,减小了冗余干扰信息的框选,取得了更精细的目标检测效果,为后续电力设备的带电状态评估与诊断奠定了基础。

参考文献:

[1] 谭宇璇,樊绍胜.基于图像增强与深度学习的变电设备红外热像识别方法[J/OL].中国电机工程学报,[2021-07-30].<http://kns.cnki.net/kcms/detail/11.2107.tm.20210601.1000.002.html>.
TAN Yuxuan, FAN Shaosheng. Infrared thermal image recognition of substation equipment based on image enhancement and deep learn-

ing[J/OL]. *Proceedings of the CSEE*, [2021-07-30]. <http://kns.cnki.net/kcms/detail/11.2107.tm.20210601.1000.002.html>.
[2] 冯振新,周东国,江翼,等.基于改进 MSER 算法的电力设备红外故障区域提取方法[J].电力系统保护与控制,2019,47(5):123-128.
FENG Zhenxin, ZHOU Dongguo, JIANG Yi, et al. Fault region extraction using improved MSER algorithm with application to the electrical system[J]. *Power System Protection and Control*, 2019, 47(5): 123-128.
[3] Jadin M S, Taib S. Recent progress in diagnosing the reliability of electrical equipment by using infrared thermography[J]. *Infrared Physics & Technology*, 2012, 55(4): 236-245.
[4] 曾军,王东杰,范伟,等.基于红外热成像的电气设备组件识别研究[J].红外技术,2021,43(7):679-687.
ZENG Jun, WANG Dongjie, FAN Wei, et al. Research on electrical equipment component recognition based on infrared thermal imaging[J]. *Infrared Technology*, 2021, 43(7): 679-687.
[5] 朱惠玲,牛哲文,黄克灿,等.基于单阶段目标检测算法的变电设备红外图像目标识别及定位[J].电力自动化设备,2021,41(8):217-224.
ZHU Huiling, NIU Zhewen, HUANG Kecan, et al. Infrared image target recognition and location of substation equipment based on single-stage target detection algorithm[J]. *Power Automation Equipment*, 2021, 41(8): 217-224.
[6] 吴克河,王敏鉴,李渊博.基于 Mask R-CNN 的电力设备红外图像分割技术研究[J].计算机与数字工程,2020,48(2):417-422.
WU Kehe, WANG Minjian, LI Yuanbo. Research on infrared image segmentation technology of power equipment based on mask R-CNN[J]. *Computer & Digital Engineering*, 2020, 48(2): 417-422.
[7] 刘云鹏,裴少通,武建华,等.基于深度学习的输变电设备异常发热点红外图片目标检测方法[J].南方电网技术,2019,13(2):27-33.

- LIU Yunpeng, PEI Shaotong, WU Jianhua, et al. Deep learning based target detection method for abnormal hot spots infrared images of transmission and transformation equipment[J]. *Southern Power System Technology*, 2019, **13**(2): 27-33.
- [8] REN Shaoqing, HE Kaiming, 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, 2015: 91-99.
- [9] Redmon J, Farhadi A. YOLOv3: An incremental improvement[J/OL]. [2018-04-08]. <https://arxiv.org/abs/1804.02767>.
- [10] LIU W, Anguelov D, Erhan D, et al. SSD: single shot multibox detector[C]// *Proceedings of the European Conference on Computer Vision. Amsterdam*, 2016: 21-37
- [11] 李文璞, 谢可, 廖道, 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 Wenpu, XIE Ke, LIAO Xiao, 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.
- [12] 王永平, 张红民, 彭闯, 等. 基于 YOLO v3 的高压开关设备异常发热点目标检测方法[J]. *红外技术*, 2020, **42**(10): 983-987.
- WANG Yongping, ZHANG Hongmin, PENG Chuang, et al. The Target detection method for abnormal heating point of high-voltage switchgear based on YOLO v3[J]. *Infrared Technology*, 2020, **42**(10): 983-987.
- [13] 梁杰, 李磊, 周红丽. 基于改进 SSD 的舰船目标精细化检测方法[J]. *导航定位与授时*, 2019, **6**(5): 43-51.
- LIANG Jie, LI Lei, ZHOU Hongli. A ship target refinement detection method based on improved SSD[J]. *Navigation Positioning & Timing*, 2019, **6**(5): 43-51.
- [14] 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: 2999-3007.
- [15] Bochkovskiy A, WANG C Y, LIAO H Y M. YOLOv4: optimal speed and accuracy of object detection[J/OL]. *Computer Vision and Pattern Recognition*, 2020, <https://arxiv.org/abs/2004.10934>.
- [16] Misra D. Mish: a self regularized non-monotonic neural activation function[J/OL]. *Computer Science*, 2019, <https://arxiv.org/abs/1908.08681>.
- [17] LIU Shu, QI Lu, QIN Haifang, et al. Path aggregation network for instance segmentation[C]// *IEEE Conference on Computer Vision and Pattern Recognition*, 2018: 8759-8768.
- [18] LIN T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]// *IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 2117-2125.
- [19] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2014, **39**(4): 640-651.
- [20] NAIR V, HINTON G E. Rectified linear units improve restricted boltzmann machines[C]// *Proceedings of the 27th International Conference on Machine Learning(ICML-10)*, 2010: 807-814.
- [21] WEN Long, GAO Liang, LI Xinyu. A new deep transfer learning based on sparse auto-encoder for fault diagnosis[J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2019, **49**(1): 136-144.