

一种基于雷达和相机数据融合网络的 输电线路鸟类多目标识别方法

范程涛¹ 高伟¹ 靳小喜²

(1. 福州大学电气工程与自动化学院, 福州 350108;
2. 福州电力设计院有限公司, 福州 350007)

摘要 本文提出一种基于雷达与相机数据融合的输电线路鸟类多目标识别网络(RVFNet)。该网络结合雷达射频(RF)数据与视觉图像,实现监控范围内鸟类目标的高精度识别。针对多模态数据语义差异的问题,计算雷达射频信号与图像位置信息的对应关系,确保特征表达的一致性。在网络结构方面,设计鸟类姿态卷积(BPC)网络,有效融合多模态信息,提升对小目标特征的提取与细节保留能力;同时,采用特征融合模块(FFM)整合不同模态特征,在显著提升特征交互能力的同时保持较低计算成本。实验表明,RVFNet在多种天气条件下对鸟类的平均识别准确率达到了80.18%,展现出良好的鲁棒性。

关键词: 探鸟驱鸟; 视觉图像; 雷达射频图像; 传感器融合; 深度卷积神经网络

A multi-target bird recognition method for transmission lines based on radar and camera data fusion

FAN Chengtao¹ GAO Wei¹ JIN Xiaoxi²

(1. College of Electrical Engineering and Automation, Fuzhou University, Fuzhou 350108;
2. Fuzhou Electric Power Design Institute Co., Ltd, Fuzhou 350007)

Abstract This paper proposes a multi-target recognition network for birds on power transmission lines, called RVFNet, based on the fusion of radar and camera data. The network achieves high-precision recognition of bird targets within the monitoring range by integrating radar radio frequency (RF) data with visual images. To address the semantic differences between multimodal data, the correspondence between radar RF signals and image positional information is calculated to ensure consistency in feature representation. Structurally, the network incorporates a bird posture convolutional network (BPC) to effectively fuse multimodal information, enhancing the extraction of small-target features and preserving fine details. Additionally, a feature fusion module (FFM) is introduced to integrate multimodal features, significantly improving feature interaction while maintaining low computational costs. Experimental results demonstrate that RVFNet achieves an average bird recognition accuracy of 80.18% under various weather conditions, highlighting its robustness.

Keywords: identify and repel birds; visual images; radar radio frequency images; sensor fusion; deep convolutional neural networks

0 引言

随着城市化进程不断推进,输电线路大规模建设与鸟类活动的冲突日益加剧,产生4种鸟害问题:鸟粪类灾害、鸟巢类灾害、鸟体短接类灾害和鸟啄

类灾害^[1-3]。这对输电线路的安全运行造成了严重影响。目前,人工巡检和传统非识别驱鸟器是防治鸟害的主要手段^[4-8]。然而,人工巡检成本高、滞后性大、风险高,难以满足预防和快速处置的需求;非识别驱鸟器虽短期有效且成本低,但长期使用会导

致鸟类产生耐受性并影响生态环境。

随着机器视觉技术日新月异,输电线路上鸟类目标识别逐渐成为业界研究热点。然而,基于监控摄像头和无人机图像的目标识别^[9-13],存在检测难度大、识别精度低等问题。文献[14]针对鸟类检测任务提出基于自监督预测卷积注意力块(self-supervised predictive convolution attention block, SSPCAB)、多分支块(multi-branch block, MBB)和双向特征金字塔网络(bidirectional feature pyramid network, BiFPN)的轻量级机场飞行鸟类检测算法,即SMB-YOLOv5(you only look once v5)模型,通过引入多分支模块并融合不同感受野的特征分支,显著增强了网络的表达能力,但模型性能易受天气条件影响。在视觉相机因低光照或烟雾遮挡等环境因素而性能下降时,雷达目标分类技术提供了有效替代方案。文献[15]提出雷达目标检测网络(radar object detection network, RODNet),该网络以射频(radio frequency, RF)图像为输入,学习并预测雷达视场内目标的存在概率,为雷达数据的处理和解释提供了新的视角。针对输电线路杆塔上鸟类目标的检测与识别,许多研究建议结合mm波雷达(简称雷达)和高清摄像头,这种传感器融合技术能够有效克服复杂条件下目标检测的难题。文献[16]提出一种基于雷达点云和视觉图像融合的鸟类检测与驱鸟方法,解决了基于视觉图像的检测技术易受天气条件和图像质量限制的问题,但在小目标检测方面仍存在一定局限性。文献[17]提出新型鸟害辨识网络(radar camera vision network, RCVNet),融合雷达射频图像与视觉图像,用于输电线路杆塔防鸟害任务,可预测鸟类的位置和类别,有效提升了检测的鲁棒性和准确性。

本文量化雷达射频图像与视觉图像的位置关系,弥补在各种天气和光照条件下图像数据中存在的特征消失和杂波现象,提出一种基于相机-雷达传感器融合和交叉监督的新型卷积神经网络,即基于雷达和相机数据融合的输电线路杆塔鸟类多目标识别方法(a multi-target recognition network for birds based on radar and visual image fusion, RVFNet),用于实现鸟类目标识别,为防治鸟害提供决策支持。针对鸟类等小目标的识别问题,引入鸟类姿态卷积(bird posture convolution, BPC)、特征融合模块(feature fusion module, FFM)等多个改进模块,通过优化特征融合策略改进特征提取网络结构,提升

对表层信息的利用率,并有效保留微小特征,从而增强模型在复杂场景中识别小目标的能力。通过实时监测输电线路区域内的鸟类活动,判定是否有鸟类停留,为采取对应驱鸟措施提供支持,预防鸟害。

1 相机-雷达数据融合

1.1 雷达射频数据的处理链

雷达作为自动驾驶辅助系统中的核心传感器,凭借其在恶劣环境下的稳定性与鲁棒性,能够与摄像头、激光雷达等传感器协同工作,有效克服传统感知系统的局限性。雷达能够精准测量目标与雷达之间的距离及其相对速度,功耗低,支持长时间连续运行,具备卓越的耐用性与稳定性,工作在mm波段(30~300GHz,波长1~10mm),具备穿透雾、烟、灰尘等障碍物的能力,并能提供广阔的感知视野。目前,雷达点云与雷达射频图像^[18]已成为雷达研究的两大主流方向。相比雷达点云,雷达射频图像在空间语义表达、数据处理灵活性及抗干扰能力方面展现出显著优势。雷达射频数据处理流程如图1所示,射频数据由雷达传感器采集的原始信号经过一系列数字与模拟信号处理后生成。具体来说,首先通过距离傅里叶变换提取目标反射的调频连续波信号,以获取目标的反射距离;然后,通过低通滤波器抑制高频噪声,并结合角度傅里叶变换获取目标的反射方位角,同时对多普勒方向进行恒虚警检测处理;最后,生成以横坐标表示方位、纵坐标表示距离的雷达射频图。该图可直观地表现为空间中的鸟瞰视图,清晰地描述目标的空间分布信息。

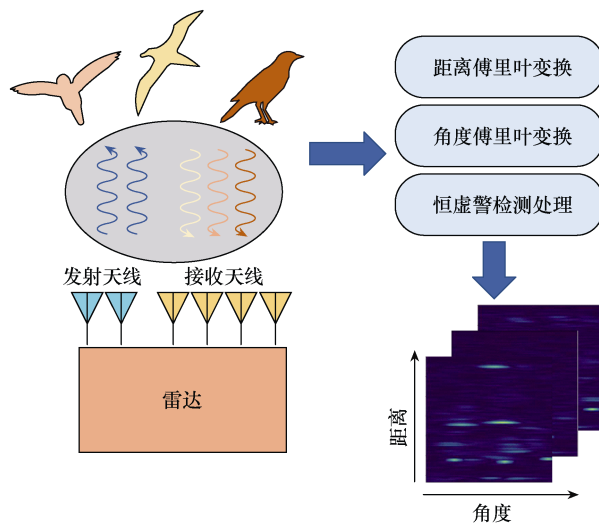
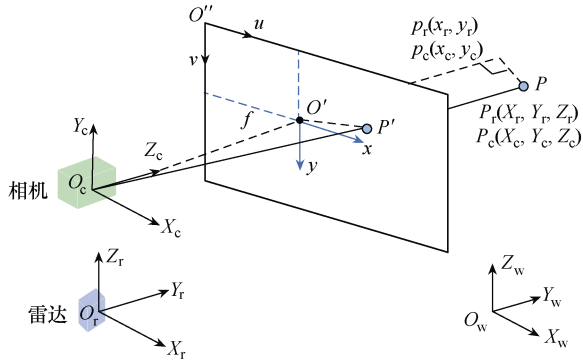


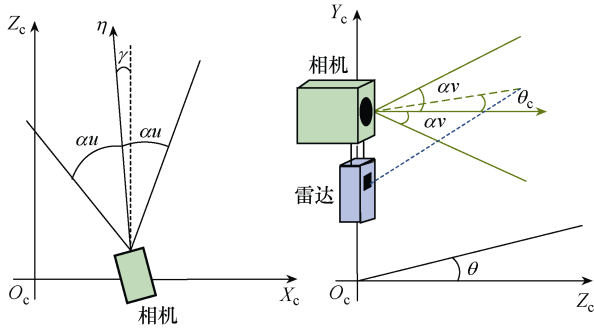
图1 雷达射频数据处理流程

1.2 射频图像和视觉图像匹配处理

为获取鸟类的雷达射频图像与视觉图像，将雷达与视觉摄像头垂直布置，其相对位置如图2(a)所示。数据融合涉及多个坐标系，包括像素坐标系 $O''-uv$ 、图像坐标系 $O'-xy$ 、相机坐标系 $O_c-X_cY_cZ_c$ 和世界坐标系 $O_w-X_wY_wZ_w$ 。由于世界坐标系仅起参考作用，为方便后续计算可将雷达坐标系 $O_r-X_rY_rZ_r$ 设定为世界坐标系。同时，外界因素（如大风或检测设备旋转）可能导致拍摄方向与主拍摄视角之间产生偏航角 γ ，如图2(b)所示。此外，相机通常的安装方式是使摄像头光轴方向与 $Y-Z$ 平面形成一个初始的俯仰角 θ_c ，但在实际场景下，由于安装位置可能存在坡度变化 $\Delta\theta$ ，因此总俯仰角 θ 包含摄像头初始俯仰角 θ_c 和坡度变化 $\Delta\theta$ 两个部分，如图2(c)所示。



(a) 相机、像素、图像、雷达坐标系的相对位置



(b) 相机的偏航角 (c) 雷达、相机对应的俯仰角

图2 坐标系与偏转角的关系

在自然场景中通常采用逆透视映射方法，该方法将图像坐标系中的点在垂直方向和水平方向上以相机位置、俯仰角、偏航角和相机角光圈为坐标，映射到世界坐标系。将原始图像与真实的世界坐标进行线性插值，得到无透视变形的鸟瞰图坐标。其中，二维图像坐标转化为三维世界坐标由式(1)和式(2)计算。

$$\begin{cases} \xi_r = \left(1 - \frac{2u}{w-1}\right) \tan\left(\arctan \frac{W}{2f}\right) \\ \xi_c = \left(1 - \frac{2v}{h-1}\right) \tan\left(\arctan \frac{H}{2f}\right) \end{cases} \quad (1)$$

$$\begin{cases} X'_c = C_y \frac{1 + \xi_r \tan \theta}{\tan \theta - \xi_r} + C_x \\ Z'_c = C_y \frac{\xi_c / \cos \theta}{\tan \theta - \xi_r} + C_z \end{cases} \quad (2)$$

式中： (u, v) 为像素坐标系中点的坐标； w, h 分别为图像的宽度和高度； f 为相机的焦距； H, W 分别为相机感光组件的高度和宽度； ξ_r, ξ_c 分别为图像行方向和列方向对应的透视畸变消除后的鸟瞰图坐标； X'_c 和 Z'_c 为从图像坐标系映射到世界坐标系的横坐标和纵坐标； (C_x, C_y, C_z) 为摄像机在真实世界中的位置坐标，选择合适的世界坐标系可以使 $C_x=C_z=0$ ， C_y 为相机距地面的高度。

由式(1)和式(2)可知，世界坐标系的 X 轴坐标只与像素坐标系中 u 轴的值有关，因此能将图像坐标系中的一行映射到世界坐标系中的同一行，从根本上解决“水平线弯曲”问题，同时解决由俯仰角所引起的透视效应。但实际上，由于风力和设备旋转等因素影响，其方向可能会因转动产生轻微的偏航角，需要补偿偏航角来消除这种旋转现象。为了使反透视映射结果中的主要拍摄方向与像素坐标系中的 v 轴平行，使用式(3)进行修正。

$$\begin{cases} X''_c = X'_c \cos \gamma + Z'_c \sin \gamma \\ Z''_c = X'_c (-\sin \gamma) + Z'_c \cos \gamma \end{cases} \quad (3)$$

式中， X''_c 和 Z''_c 为经过 γ 角补偿后从图像坐标系映射到世界坐标系的横坐标和纵坐标。

通过式(3)获取视觉图像相应位置坐标后，根据式(4)~式(6)计算对应射频图像坐标 (ρ_i^c, θ_i^c) 。

$$\begin{bmatrix} X_r \\ Y_r \\ Z_r \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{3 \times 3} & \mathbf{T}_{3 \times 1} \\ \mathbf{0}_{1 \times 3} & \mathbf{I}_{1 \times 1} \end{bmatrix}^{-1} \begin{bmatrix} X''_c \\ Y_c \\ Z''_c \\ 1 \end{bmatrix} \quad (4)$$

$$\theta_i^c = \arctan \frac{X_r - X_{or}}{Z_r - Z_{or}} \quad (5)$$

$$\rho_i^c = \sqrt{(X_r - X_{or})^2 + (Z_r - Z_{or})^2 + H_{cr}^2} \quad (6)$$

式中： (X_r, Y_r, Z_r) 为目标在雷达坐标系下的三维空间

坐标; X_{or} 、 Z_{or} 为雷达在雷达坐标系下的位置坐标; H_{cr} 为相机与雷达位置相对高度差; $\mathbf{0}_{1 \times 3}$ 为零矩阵; $\mathbf{I}_{1 \times 1}$ 为单位矩阵; $\mathbf{R}_{3 \times 3}$ 和 $\mathbf{T}_{3 \times 1}$ 分别为相机与雷达的旋转矩阵和偏移矩阵。

2 基于雷达和相机数据融合的输电线路杆塔鸟类多目标识别方法

2.1 RVFNet 的网络结构

RVFNet 的网络结构如图 3 所示, RVFNet 在融合雷达射频图像与视觉图像特征的基础上, 对 YOLO (you only look once) 框架进行优化, 形成一种适用于多模态输入的网络结构。RVFNet 的输入阶段有两个并行输入通道, 分别接收雷达射频图像和视觉图像, 其核心在于深度整合两种模态数据的目标信息, 通过特征筛选机制去除缺乏空间分布特性和物理意义的冗余特征点。在雷达射频图像处理通道中, 首先利用具有残差结构的浅层网络 ResNet13 进行初步特征提取, 然后采用卷积核大小分别为 1×1 、 3×3 和 5×5 的卷积 (Conv) 操作, 将卷积层相

加形成空间多尺度融合模块, 并结合激活函数生成雷达加权矩阵, 最后将该加权矩阵与经过 Focus 模块压缩尺寸及卷积处理后的视觉图像相乘, 生成特征向量输入主干网络。在视觉图像通道中, 为充分提取表层语义信息, 结合自注意力机制对视觉特征进行编码, 生成 3 种特征向量, 并分别输入网络的不同层次, 以捕捉多尺度视觉特征。在特征提取部分, RVFNet 引入鸟类姿态卷积, 设计灵感来源于鸟类在飞行与静止状态下翅膀的动态变化, 有效增强了模型在复杂场景中捕捉与表达目标特征的能力。RVFNet 的颈部结构基于特征金字塔设计理念, 能够接收 4 种不同尺寸的特征图, 并通过多次卷积、上采样和下采样操作, 将特征图生成 4 种不同尺度的输出。在头部结构中, 网络采用包含 Tiny-YOLOHead 在内的四解耦头进行多尺度的目标回归与分类。通过上述融合机制, RVFNet 实现了多尺度特征的高效捕捉, 有效减少了因目标尺寸变化导致的检测混乱与干扰, 同时优化了模型的反向传播过程, 大幅提升了学习效率与检测性能。

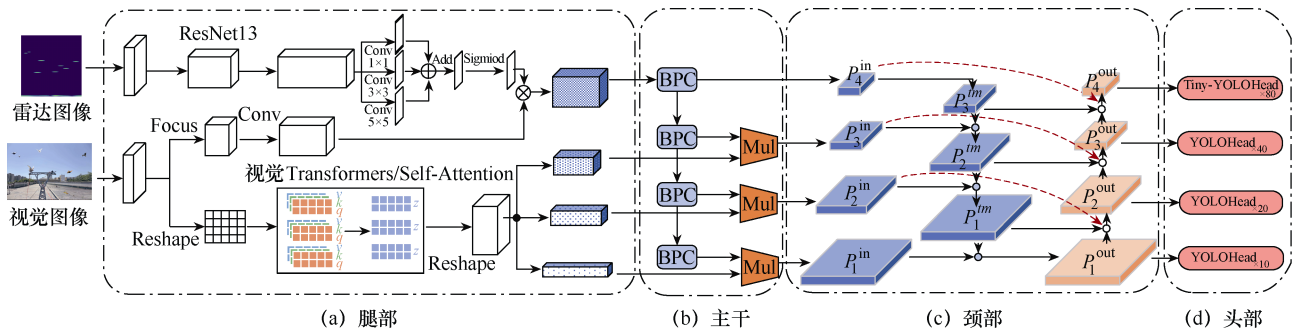


图 3 RVFNet 的网络结构

2.2 鸟类姿态卷积网络

雷达射频图像的像素大小受雷达天线与目标物体距离、物体运动速度及目标大小的影响。鸟类动作幅度大、飞行速度快, 容易导致图像中出现重影和模糊等现象, 还会引起射频图像中部分像素点动态变化。因此, 视觉图像与雷达射频图像融合特征中包含了大量动态特征。有效提取这些动态特征对准确定位鸟类具有重要意义。

设计一种基于扩张卷积 (dilated convolution, DC) 和可变形卷积 (deformable convolution, DConv) 的鸟类姿态卷积网络, BPC 网络结构如图 4 所示。图 4 中 (a) 扩张卷积相较于传统的二维卷积, 在不增加参数量的情况下, 显著扩大了卷积核的感受野, 同时保持输出特征映射的尺寸不变。可变形卷积通

过自适应调整卷积核的形状, 更好地捕捉不同尺度和形状的特征, 如图 4 中 (b) 所示。以大小为 (b, h, w, n) 的输入特征图及卷积核 3×3 为例, 其中 b 为批次大小, h 和 w 为输入特征图的高和宽, n 为输入通道数。扩张卷积将卷积核的感受野扩展为 5×5 , 并生成偏移特征图, 特征图大小为 $(b, h, w, 2n)$, 其中包含 x 和 y 方向上的偏移量。可变形卷积中的偏移量并非固定值, 通常为小数, 从而允许卷积核在输入特征图上以非网格化 (即非整数坐标) 的方式进行采样。通过双线性插值计算偏移后的像素值, 其计算公式如式 (7) 和式 (8) 所示。最后, 如图 4 中 (c) 所示, 通过交叉重建的方式将偏移特征图与原始特征图进行融合。这种融合方法能够充分结合两种信息丰富的加权特征, 并增强它们之间的信

息流，从而生成具有更高表现力和更加丰富的空间内容的特征。

$$y(p_0) = \sum_{p_k \in R} w_a(p_k) \cdot x(p_0 + p_k + \Delta p_k) \quad (7)$$

式中： p_0 为输出特征图上的一个特征点； R 为感受野坐标集合， $R = \{(-2, -2), (-2, -1), \dots, (1, 2), (2, 2)\}$ ； p_k 为枚举 R 中的位置； Δp_k 为正则网络的增广偏移量； k 为索引，用于遍历 R 中每一个位置， $k \in (0, N)$ ， $N = |R|$ ； w_a 为卷积核的权值； x 为输入特征映射； y 为输出特征图中的一个像素的值。

$$\begin{aligned} x(p) &= \sum_q G(p, q) \cdot x(q) \\ &= \sum_q g(p_x, q_x) \cdot g(p_y, q_y) \cdot x(q) \\ &= \sum_q \max(0, 1 - |q_x - p_x|) \cdot \max(0, 1 - |q_y - p_y|) \cdot x(q) \end{aligned} \quad (8)$$

式中： p 为任意位置， $p = p_0 + p_k + \Delta p_k$ ； q 为枚举特征映射 x 中的所有积分空间位置； $G(\cdot, \cdot)$ 为双线性插值核函数； $g(\cdot, \cdot)$ 为线性插值核函数。

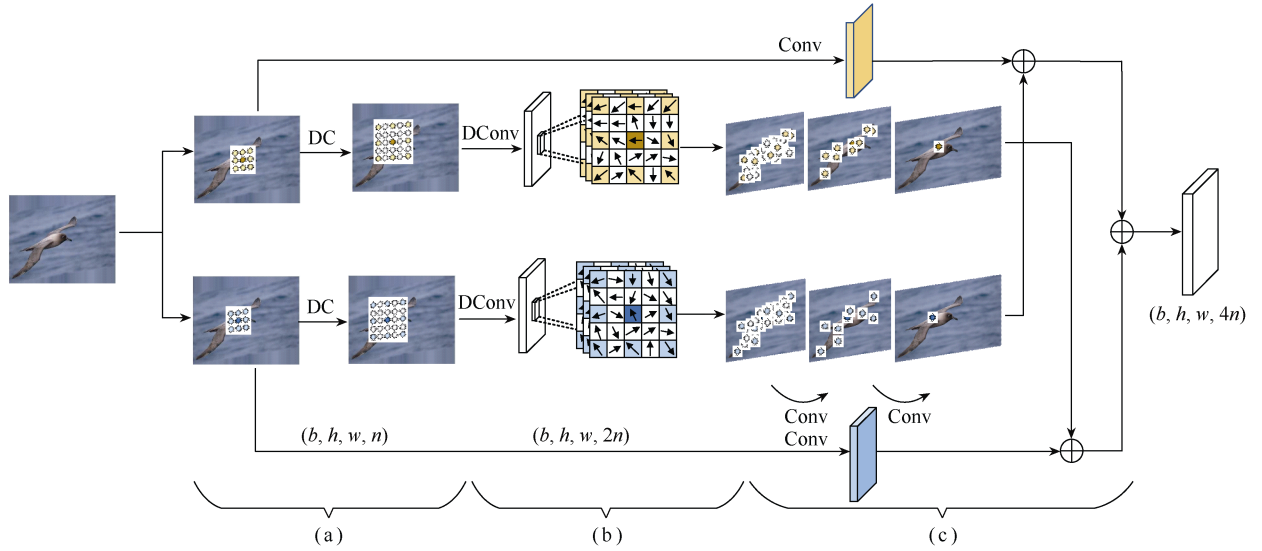


图4 BPC网络结构

2.3 特征融合模块

由于鸟类在高空飞行，其视野中的物体通常尺寸较小且视觉细节较少，这要求模型具备出色的小目标检测能力。同时，在各种突发情况下，鸟类的飞行高度可能发生显著变化，从而导致视野中的物体尺度也会出现较大波动。这要求模型具有较高的多尺度目标检测能力。

多尺度目标检测是一项具有挑战性的任务，因为需要权衡特征图的尺寸与语义特征。深层特征图包含更丰富的高级语义信息和更大的感受野，适用于大尺度目标的识别和分类；浅层特征图具有更高的分辨率和更多的位置信息，对小目标的识别与定位尤为关键。本文设计特征融合模块，FFM网络结构如图5所示。该模块通过在输入与输出节点之间添加跳跃连接，实现更多特征的有效融合，同时不会显著增加计算成本，如图5中（红色）虚线所示。此外，由于不同分辨率的输入特征对融合的贡献不同，本文为每个输入特征添加可学习的权重，并提

出一个高效的权重融合模块，其计算公式为

$$X_{\text{Output}} = \frac{\sum_i \omega_i X_{\text{Input}}^i}{\sum_i \omega_i + \varepsilon} \quad (9)$$

式中： X_{Output} 为输出特征； X_{Input}^i 为输入特征； ω_i 为可学习权重， $\omega_i \geq 0$ ，为了提高训练的稳定性，将权重归一化至 $[0, 1]$ ； ε 为一个修正值，用于提高计算稳定性。

2.4 鸟类驻留识别

针对鸟类识别结果的预测框，采用一种三帧差分广义交并比 $S_{3F-GIoU}$ 的计算方法^[16]，用以描述多个预测框之间的位置变化和重叠程度，并作为鸟害发生的判据。 $S_{3F-GIoU}$ 的计算公式为

$$\begin{aligned} S_{3F-GIoU} &= J(A, B, C) - \\ &\frac{1}{|D|} [|D| - (|A \cup B \cup C| - I(A, B, C) + |A \cap B \cap C|)] \end{aligned} \quad (10)$$

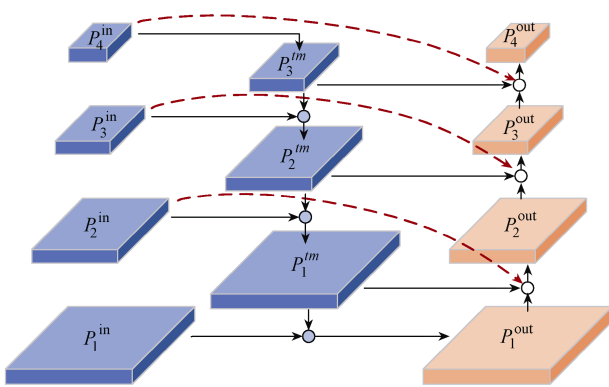


图 5 FFM 网络结构

$$I(A, B, C) = |A \cap B| + |B \cap C| + |C \cap A| \quad (11)$$

$$J(A, B, C) = \frac{I(A, B, C) - 2|A \cap B \cap C|}{|A \cup B \cup C| - I(A, B, C) + |A \cap B \cap C|} \quad (12)$$

式中： A 、 B 和 C 为一帧中的任意鸟类预测框； D 为能够同时包含它们的最小闭包框； $J(\cdot)$ 为改进的杰卡德系数； $I(\cdot)$ 为差分交集面积之和。

3 实验测试

3.1 数据集获取与分布

本文使用的数据集来自 CRB2022^[17]，用于训练 RVFNet 模型。CRB2022 是一个用于研究鸟类识别定位的雷达射频图像和视觉图像数据集，由雷达射频图像和视觉图像数据集 CRUW^[15] 和多鸟种的数据集 Caltech-UCSD Birds 200—2011^[19] 融合而来。该数据集包括视觉图像和雷达射频图像各 8 141 帧，囊括 29 431 个 YOLO 格式的注释框，视觉图像分辨率为 1 440×864 像素，雷达射频图像分辨率为 128×128 像素。CRB2022 数据集如图 6 所示，显示了 CRB2022 中鸟类在视觉图像和 RF 图像中的位置。

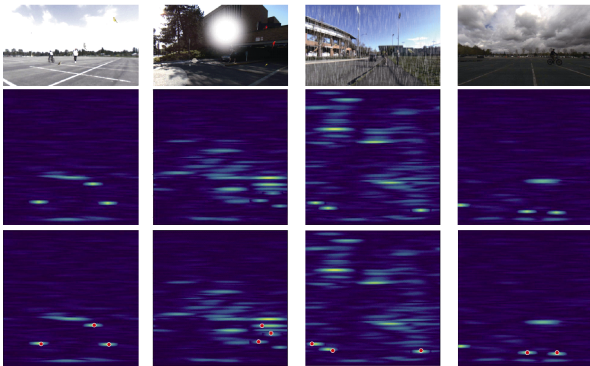


图 6 CRB2022 数据集

为了获取更多实际拍摄数据，本文设计并制作了一套驱鸟装置，如图 7 所示。装置的采集部分，雷达和相机的硬件配置见表 1。该驱鸟装置配备了一台 800 万像素的单目摄像头，用于捕捉鸟类的视觉特征，同时配备了一台 TI AWR1443 型号的 76~77GHz 调频连续波 (frequency modulated continuous wave, FMCW) mm 波雷达，可在复杂环境下稳定采集鸟类的射频特征图。这两种传感器通过垂直布置实现数据的同步采集，从而确保视觉图像与射频图像在空间位置上的精确匹配。为扩大数据采集范围并提升采集效率，采集装置底部安装了一个可匀速旋转的云台，用于实现 360° 范围内的覆盖采集。此外，装置还配备了一部高精度激光，用于鸟类驱逐，减少鸟类对输电线路的潜在威胁。将驱鸟装置固定在杆塔周围，重点考虑装置的稳定性与视野覆盖范围。为确保采集装置在不同环境下的高效运行，对装置的硬件和软件进行多次调试，包括调整雷达和相机相对位置以优化射频信号和图像采集效果，以及校正云台的俯仰角以确保捕捉到清晰的鸟类图像。同时，对云台的旋转速度及范围进行精确设置，以实现数据采集的均匀性与全面性。实验测试表明，

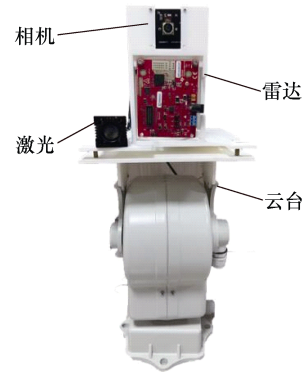


图 7 驱鸟装置

表 1 雷达和相机硬件配置

参数		数值
雷达	帧率/(帧/s)	30
	频率/GHz	76~77
	距离分辨率/m	0.25
	方位角分辨率/(°)	0~15
	发射器数量	2
	接收器数量	4
相机	图像分辨率	3 840×2 160
	信噪比/dB	36
	动态范围/dB	86

该驱鸟装置能够在复杂环境下稳定运行,可对杆塔周围的鸟类实现有效驱逐,同时高效采集多模态数据。

3.2 评价指标

为评估模型性能,本文采用广泛用于图像目标检测任务的交并比(intersection over union, IoU)作为评估标准,以衡量检测结果与真实目标之间的匹配程度。同时,选用平均精度(average precision, AP)和平均召回率(average recall, AR)作为主要评估指标^[17],分别用于衡量模型在不同阈值下的检测准确性与覆盖能力。此外,本文还重点关注模型的检测时间 T ,即完成一次目标检测所需的时间,通常以ms为单位。AP和AR的数值越高,说明模型在鸟类检测任务中的识别准确性和鲁棒性越强;检测时间越短,表明模型在目标检测速度方面更具优势,能够更好地满足实时检测需求。

3.3 CRB2022数据集上鸟类检测结果和算法对比

为了验证基于雷达-相机数据融合的检测框架

表2 使用RVFNet和基于单一模态的基线在CRB2022数据集上的执行结果

传感器类型	模型	T/ms	全部		鸟类		人	
			AP/%	AR/%	AP/%	AR/%	AP/%	AR/%
基于视觉	Faster-RCNN	87	48.41	57.10	46.60	56.61	50.28	57.63
	YOLOX	20	65.66	74.03	64.43	73.86	66.89	74.22
	YOLOv3	25	51.25	67.14	50.24	67.98	52.32	66.85
雷达视觉融合	RVFNet	70 (<100)	81.45	85.04	80.18	84.21	82.59	85.86

同时,使用多种多模态融合对比模型进行实验评估,包括文献[16]提出的基于雷达点云与视觉图像数据的雷达-相机融合网络(radar point attention mechanism-you only look once, RPAM-YOLO),以及文献[23]提出的结合YOLOv3的不确定性雷达-相机传感器融合目标检测网络(you only look once v3-based uncertainty radar-camera sensor fusion target detection network, YOdar)。多模态融合模型在CRB2022数据集上的执行结果见表3。根据表3的结果,RVFNet模型的检测准确度显著优于其余两种

表3 多模态融合模型在CRB2022数据集上的执行结果

传感器类型	模型	T/ms	全部		鸟类		人	
			AP/%	AR/%	AP/%	AR/%	AP/%	AR/%
雷达视觉融合	RPAM-YOLO	30	71.25	73.63	70.14	75.32	72.31	74.85
	YOdar	35	70.11	68.53	69.52	66.88	71.63	69.36
	RVFNet	70 (<100)	81.45	85.04	80.18	84.21	82.59	85.86

不同模型在不同天气条件下的检测结果如图8所示。从上至下,每一行分别为射频图像、RVFNet、

RVFNet,将其与先进的视觉识别方法及多传感器融合模型在增强后的CRB2022数据集上进行对比,所有模型均在配备有32GB内存的Nvidia RTX 4060 GPU上运行,设置初始学习率为 1×10^{-2} ,对网络进行500次训练并将CRB2022数据集按9:1的比例分为训练集和测试集。

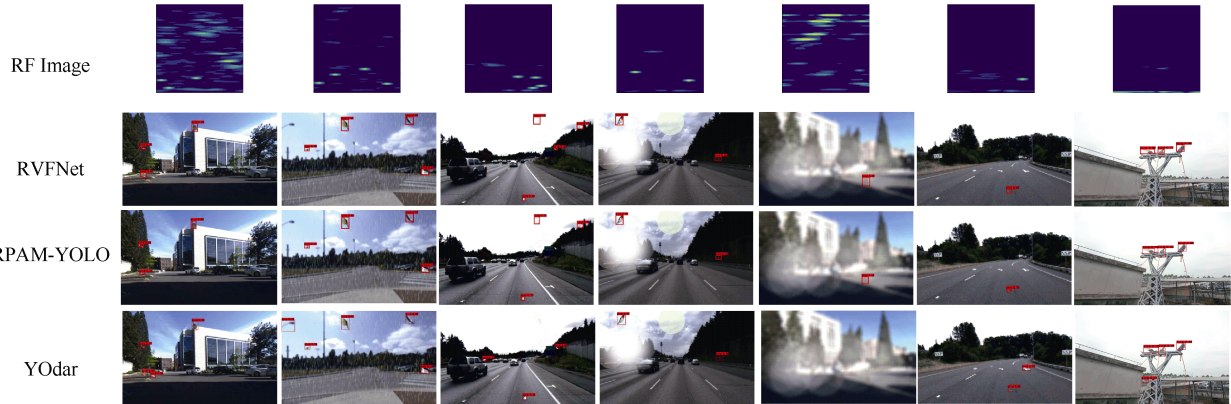
所对比的视觉识别算法包括工程实践中常用的更快的基于区域的卷积神经网络(faster region-based convolutional neural network, Faster-RCNN)^[20]模型、能够缓解正负样本不平衡问题的YOLOX(you only look once X)^[21]算法及具有较快目标检测速度的YOLOv3^[22]算法。使用RVFNet和基于单一模态的基线在CRB2022数据集上的执行结果见表2,所有基线和RVFNet模型均使用推荐的训练设置进行训练。结果显示,与基于视觉的识别方法相比,RVFNet融合方法的性能得到了显著提升。

先进的多模态融合方法。这进一步验证了RVFNet在处理复杂场景和多模态数据融合任务时的卓越性能。表2和表3还记录了不同模型在Nvidia RTX 4060 GPU上的检测时间。值得注意的是,YOLOX仅使用视觉图像作为输入,其检测时间为20ms,这归因于YOLOX对YOLO检测器的无锚点设计优化,有效减少了参数量和计算需求。尽管RVFNet的检测时间为70ms,相较于YOLOX有所增加,但依然满足实时检测时间阈值要求,同时在检测准确度和鲁棒性方面表现更优。

RPAM-YOLO和YOdar模型的检测结果;从左至右,每一列分别展示了正常天气、雨天、过曝场景、强

光照、雾天、单只鸟类及实拍图像场景下的检测结果。从图 8 可以看出，在复杂天气条件下，RVFNet 模型的检测效果明显优于其他模型，其能够完整、准确地检测被遮挡的目标，无明显漏检和误检现象。

这表明 RVFNet 在多尺度目标检测和抗遮挡能力方面表现出色，不受混杂信息的干扰。此外，RVFNet 还能在大视场图像中更精准地识别感兴趣区域，展现出较强的鲁棒性和适应性。



(a) 正常天气 (b) 雨天天气 (c) 强曝光效果 (d) 强光效果 (e) 雾天天气 (f) 单只鸟类 (g) 实拍图像

图 8 不同模型在不同天气条件下的检测结果

3.4 消融实验

为了验证 BPC 和 FFM 对输电线路杆塔鸟类检测模型 RVFNet 的改进效果，设计一系列消融实验，实验结果见表 4，其中“√”代表包含某个模块，“×”代表不包含某个模块。由表 4 可见，与仅使用视觉图像检测相比，雷达-相机交叉融合检测的识别 AP

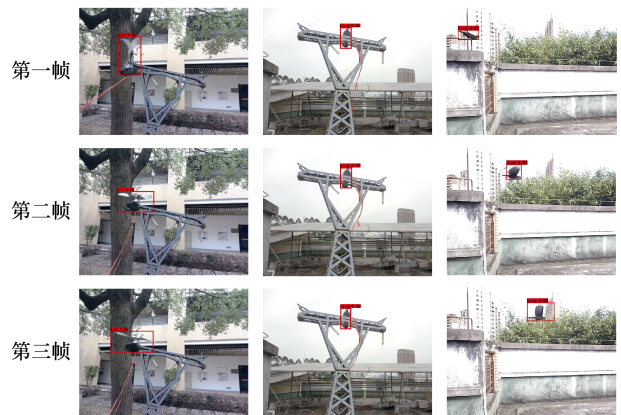
提升了约 6 个百分点。BPC 通过对鸟类翅膀变化的动态特征提取，进一步增强了网络对局部细节的捕捉能力，有效地保留了微小特征，提升了模型的检测性能。FFM 对不同层次的特征进行交互融合，提高了小目标的特征表达能力。

表 4 RVFNet 消融实验结果

传感器类型	BPC	FFM	T/ms	全部		鸟类		人	
				AP/%	AR/%	AP/%	AR/%	AP/%	AR/%
基于视觉	×	×	20	69.59	75.82	68.82	74.55	70.32	75.23
雷达视觉融合	×	×	25	75.11	77.53	74.20	76.34	76.98	78.66
	√	×	60	78.43	81.51	77.94	80.96	79.88	82.02
	√	√	70 (<100)	81.45	85.04	80.18	84.21	82.59	85.86

3.5 鸟类驻留识别分析

不同姿态下的三帧预测框数据如图 9 所示。当鸟类停留在监控区域内时，其姿态和位置较为固定，预测框的重叠程度较高，此时 $S_{3F-GIoU}$ 的值较大，并随着鸟类姿态的稳定性增强而进一步增大，如图 9 (a) 和图 9 (b) 所示。这表明三帧差分广义交并比的计算方法能够准确捕捉鸟类在静止状态下的行为特征，为识别鸟类停留行为提供可靠依据。当鸟类飞越监控区域时，预测框呈现出一定的运动轨迹，重叠程度较低，此时 $S_{3F-GIoU}$ 的值较小，并随着鸟类飞行速度的增加而逐渐减小，如图 9 (c) 所示。这表明三帧差分广义交并比的计算方法对动态监测鸟类飞行状态也具有较好的灵敏性和响应能力。基于



(a) $S_{3F-GIoU}=0.39$ (b) $S_{3F-GIoU}=0.97$ (c) $S_{3F-GIoU}=-0.56$

图 9 不同姿态下的三帧预测框数据

此,利用三帧广义交并比方法不仅可以辅助判断鸟类的活动状态,还能够实现驱鸟设备的自动起停,从而在节约能源的同时,有效预防鸟类对输电线路安全运行的影响。

4 结论

本文提出了一种基于雷达和相机数据融合的输电线路杆塔鸟类多目标识别方法 RVFNet。针对鸟类图像目标检测领域中存在的挑战,如小目标众多、遮挡问题严重及对检测速度的需求,设计了 BPC 和 FFM 等模块,并论证了雷达射频图像和视觉图像的匹配关系,有效改善了目标检测模型在识别微小鸟类目标、抵御图像混淆及遮挡等方面的表现。其中,BPC 针对鸟类飞行姿态变化,提高了特征提取的准确性和检测效率;FFM 融合不同层次的特征来增强特征表达能力,从而提高了模型对多尺度目标的检测能力。实验结果表明,RVFNet 在 CRB2022 数据集上的表现优于其他方法,其对鸟类识别的平均精度达到 80.18%,证明了该方法在复杂天气、不同光照背景条件下对电力输电线路鸟类识别的有效性。

未来须进一步扩充实际场景中不同鸟群的数据集,以实现针对不同目标鸟类的精准检测,并采用不同措施实现针对性的驱逐。同时,还须对驱鸟装置进行现场应用测试,以进一步验证其性能。

参考文献

[1] 杨鑫,周攀,李斌,等. 10kV 配电线路鸟害防治措施研究[J]. 科技创新与应用, 2021, 11(18): 109-113.

[2] 戴宇辰,叶青,许安杰,等. 配电线路鸟害故障预测模型研究[J]. 电气技术, 2018, 19(3): 100-102.

[3] 王翼虎,徐浩. 330kV 紧凑型铁塔鸟害故障研究[J]. 电工技术, 2023(22): 202-205.

[4] 付豪,张东,董新胜,等. 输电线路涉鸟故障防范措施研究[J]. 电气技术, 2016, 17(3): 132-135.

[5] 李秋平. 关于 20kV 线路裸导线运行的防鸟害研究[J]. 电气技术, 2016, 17(8): 136-139.

[6] 张传民,张东. 防鸟刺针刺长度选择方法[J]. 电气技术, 2023, 24(4): 48-51, 56.

[7] 刘希和,刘佳俊,柯贤友,等. 220kV 输电线路感应电压驱鸟试验和应用[J]. 电气技术, 2013, 14(9): 44-48.

[8] 周文平,李永茂,张君,等. 电力线路驱鸟防治系统研究[J]. 电工技术, 2020(5): 58-59, 62.

[9] 刘传洋,吴一全,刘景景. 无人机航拍图像中绝缘

子缺陷检测的深度学习研究方法研究进展[J]. 电工技术学报, 2025, 40(9): 2917-2930.

[10] 郑含博,胡思佳,梁炎燊,等. 基于 YOLO-2MCS 的输电线路走廊隐患目标检测方法[J]. 电工技术学报, 2024, 39(13): 4164-4175.

[11] 仲林林,胡霞,刘柯妤. 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.

[12] 廖宁生,曹天秀,刘科言,等. 复合特征与多尺度融合的无人机小目标检测算法[J]. 计算机工程与应用, 2025, 61(3): 111-120.

[13] 杨波,曹雪虹,焦良葆,等. 改进实时目标检测算法的电力巡检鸟巢检测[J]. 电气技术, 2020, 21(5): 21-27, 32.

[14] LIANG Haijun, ZHANG Xiangwei, KONG Jianguo, et al. SMB-YOLOv5: a lightweight airport flying bird detection algorithm based on deep neural networks[J]. IEEE Access, 2024, 12: 84878-84892.

[15] WANG Yizhou, JIANG Zhongyu, LI Yudong, et al. RODNet: a real-time radar object detection network cross-supervised by camera-radar fused object 3D localization[J]. IEEE Journal of Selected Topics in Signal Processing, 2021, 15(4): 954-967.

[16] 吴洋铭,洪翠,高伟. 基于雷达点云与视觉图像融合的输电线路探鸟驱鸟技术[J]. 高电压技术, 2023, 49(8): 3446-3457.

[17] GAO Wei, WU Yangming, HONG Cui, et al. RCVNet: a bird damage identification network for power towers based on fusion of RF images and visual images[J]. Advanced Engineering Informatics, 2023, 57: 102104.

[18] LIANG Tianyi, LI Baopu, WANG Mengzhu, et al. A closer look at the joint training of object detection and re-identification in multi-object tracking[J]. IEEE Transactions on Image Processing, 2022, 32: 267-280.

[19] ZHENG Zhaohui, WANG Ping, LIU Wei, et al. Distance-IoU loss: faster and better learning for bounding box regression[C]//AAAI Conference on Artificial Intelligence, New York, USA, 2020: 12993-13000.

[20] REN Shaoqing, HE Kaiming, 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.

[21] DAI Jifeng, QI Haozhi, XIONG Yuwen, et al. Defor-

- [16] NZAMANAN N, SYKULSKI J. The evolution of high impedance fault modeling[C]//2014 16th International Conference on Harmonics and Quality of Power (ICHQP), Bucharest, Romania, 2014: 77-81.
- [17] 钟逸涵, 邓丰, 史鸿飞, 等. 基于动态电阻串联的高阻接地故障精确建模[J]. 电工技术学报, 2024, 39(7): 2046-2059.
- [18] 孙月琴, 倪江, 王宾, 等. 应用于输电线路单端测距的高阻接地故障电弧模型分析[J]. 电力系统自动化, 2016, 40(22): 86-92.
- [19] 梁晨光, 王宾, 李凤婷, 等. 高压线路单相弧光接地对数仿真模型数据修正算法[J]. 电力系统自动化, 2019, 43(4): 126-131.
- [20] 邵庆祝, 崔鑫, 谢民, 等. 弧光高阻接地故障建模及数据修正算法[J]. 电力系统自动化, 2021, 45(11): 120-125.
- [21] WEI Mingjie, LIU Weisheng, SHI Fang, et al. Distortion-controllable arc modeling for high impedance arc fault in the distribution network[J]. IEEE Transactions on Power Delivery, 2021, 36(1): 52-63.
- [22] 梁栋, 徐丙垠, 唐毅, 等. 10kV 架空导线单相触树接地故障模型及其检测方法[J]. 中国电机工程学报, 2021, 41(15): 5221-5232.
- [23] 余颖, 刘亚东, 李维, 等. 配电线路针式绝缘子早期故障动态特性研究[J]. 电工技术学报, 2023, 38(1): 71-82.
- [24] 韦明杰. 基于宽频同步测量的配电网高阻故障诊断技术与应用[D]. 济南: 山东大学, 2022.
- [25] 伍家驹, 冯上贤, 徐杰, 等. 电阻负载单相全波整流电路无功分量的一种图解方法[J]. 中国电机工程学报, 2019, 39(5): 1505-1516.
- [26] IZADI M, MOHSENIAN-RAD H. Characterizing synchronized Lissajous curves to scrutinize power distribution synchro-waveform measurements[J]. IEEE Transactions on Power Systems, 2021, 36(5): 4880-4883.
- [27] 任万滨, 金建炳, 郭继峰, 等. 碳电极交流电弧伏安特性的实验研究[J]. 电工技术学报, 2014, 29(1): 18-22, 30.
- [28] 梁军, 屈智巍, 贾秀贞, 等. 树体电容的生理学研究[J]. 林业科学, 2006, 42(1): 90-95.
- [29] 王仁甫. 电弧现象模型的发展[J]. 高压电器, 1991, 33(4): 39-46.
- [30] 秦际轲, 宁鑫, 范松海, 等. 架空导线单相触树接地故障的放电引燃机制研究[J]. 电网技术, 2023, 47(3): 1289-1299.
- [31] 贾曜霆, 王军, 王棣生, 等. 配电线路单相触树接地故障建模分析与仿真[J/OL]. 电网技术, 1-10 [2025-03-11]. <http://kns.cnki.net/kcms/detail/11.2410.tm.20241204.1550.009.html>.

收稿日期: 2025-01-20

修回日期: 2025-02-18

作者简介

林万里(2000—), 男, 福建泉州人, 硕士, 主要研究方向为配电网及其自动化技术。

(上接第37页)

mable convolutional networks[C]//2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017: 764-773.

[22] REDMON J, FARHADI A. YOLOv3: an incremental improvement[EB/OL]. <https://arxiv.org/abs/1804.02767v1>.

[23] KOWOL K, ROTTMANN M, BRACKE S, et al. Yodar: uncertainty-based sensor fusion for vehicle detection

with camera and radar sensors[EB/OL]. <https://arxiv.org/abs/2010.03320v2>.

收稿日期: 2025-01-09

修回日期: 2025-02-25

作者简介

范程涛(1999—), 男, 福建省三明市人, 硕士研究生, 主要从事多模态信号输电线路探鸟驱鸟技术研究工作。