

DOI: 10.19650/j.cnki.cjsi.J2413509

融合梯度改进 YOLO 和 KCF 模型的无人机 目标识别跟踪算法*

王文胜^{1,2}, 何君尧¹, 黄 民^{1,2}, 吴国新²

(1. 北京信息科技大学机电工程学院 北京 100192; 2. 北京信息科技大学现代测控技术教育部重点实验室 北京 100192)

摘 要:针对无人机目标小、目标不显著的情况以及目标被遮挡后的再跟踪问题,提出一种将改进 YOLO 和改进 KCF 模型融合的无人机识别跟踪算法 YOLO-G-KCF。该算法在特征处理方面将多通道梯度特征和原图特征通过特征级联的方式进行融合,并将融合特征引入 YOLOv10 算法之中,使改进算法对强光照、阴影等复杂光照条件下的目标有更好的检测效果;同时将多通道梯度特征信息引入 KCF 目标跟踪算法之中,通过设计一种多尺度特征检测,使 KCF 算法具有良好的尺度自适应;在头侧引入 KCF 跟踪结果,计算得新的损失函数 Inner-IoU,更准确的识别跟踪目标。经实验表明,在由多个开源无人机视频目标跟踪组成的数据集上进行测试,YOLO-G-KCF 算法取得 95.3% 的准确率;与 YOLOv10 原始模型相比,改进模型的 $mAP@0.5$ 提高了 1.37%,平均精度 $mAP@0.5$ 达到了 94.28%,且识别速度达到了 112 FPS,能以 100 FPS 以上的速度运行,满足无人机目标识别跟踪的实时性需求。通过引入识别机制的跟踪并进行改进,在不损失速度的基础上,对比其他算法具有更好地识别跟踪效果。YOLO-G-KCF 算法实现了对无人机在目标小、不显著以及遮挡后再跟踪等情况下的识别跟踪,识别准确率高、抗干扰能力强、硬件开发实时性好,具有一定的理论研究和工程应用价值。

关键词: 目标跟踪;目标识别;无人机跟踪;YOLOv10

中图分类号: TP391.41 TH89 **文献标识码:** A **国家标准学科分类代码:** 510.4050

Fusion gradient improved YOLO and KCF models for UAV target recognition and tracking algorithm

Wang Wensheng^{1,2}, He Junyao¹, Huang Min^{1,2}, Wu Guoxin²

(1. College of Mechanical and Electrical Engineering, Beijing Information Science and Technology University, Beijing 100192, China; 2. Key Laboratory of Modern Measurement and Control Technology, Ministry of Education, Beijing Information Science and Technology University, Beijing 100192, China)

Abstract: In view of the small and unobvious target of the UAV and the re-tracking problem after the target is occluded, an anti-UAV recognition and tracking algorithm, YOLO-G-KCF, which integrates improved YOLO and improved KCF models, is proposed. This algorithm introduces multi-channel gradient features and original image features into the YOLOv10 algorithm by means of feature concatenation in feature processing. Therefore, the improved algorithm has a better detection effect on targets under strong light, shadow, and other complex lighting conditions. Meanwhile, the multi-channel gradient features are introduced into the KCF target tracking algorithm, and a multi-scale detection is designed to make the KCF algorithm have good scale adaptability. The KCF tracking results are introduced after the Head, and the new loss function Inner-IoU is calculated to more accurately identify the tracking target. The experimental results show that the YOLO-G-KCF algorithm achieves a 95.3% accuracy rate when tested on the dataset comprising multiple open-source UAV video target tracking. This is in comparison with the original model of YOLOv10, wherein the improved model's $mAP@0.5$ has an increase of 1.37%, and the average precision $mAP@0.5$ reaches 94.28% and the recognition speed reaches 112 FPS, which can operate at more than 100 FPS to satisfy the real-time requirements of UAV target recognition and tracking. Compared with other algorithms without sacrificing speed, introducing recognition mechanisms for tracking and improving them has better

收稿日期: 2024-11-19 Received Date: 2024-11-19

* 基金项目: 国家重点研发计划(2020YFB1713203)、北京市教育委员会科研计划(KM202411232023)、北京信息科技大学“青年骨干教师”支持计划(YBT202403)项目资助

recognition and tracking effects. YOLO-G-KCF algorithm realizes the recognition and tracking of low-speed, small-sized, and low-altitude unmanned aerial vehicles in situations where the target is small, not prominent, and occluded. It has high recognition accuracy, strong anti-interference ability, good real-time hardware development, and certain theoretical research and engineering application value.

Keywords: target tracking; target recognition; UAV tracking; YOLOv10

0 引言

无人机(unmanned aerial vehicle, UAV)由于其具有操作简单、体积小和成本低等优势,在民用和军用领域都有着广泛的应用。近些年来无人机的使用不断融入到生产生活中的各个方面。根据民航局所公布的数据,截止到2024年上半年,新注册无人机约60.8万架,对比2023年年底增长48%,无人机飞行小时数累计达到981.6万小时,较去年同期增加13.4万小时,无人机应用和研发正在慢慢成为未来的主流方向。无人机给人们生产生活带来便利的同时,“黑飞”现象的发生对国家安全和公共安全造成了严重的威胁^[1]。近两年来“黑飞”现象对机场区域、铁路电气化线路管控区域、列车行驶区域等区域造成了严重的安全隐患,2023年珠海机场受“黑飞”影响;2024年9月天津机场受到“黑飞”影响导致航班起降受到影响,导致大面积航班延误,多数乘客滞留机场。上述情况不胜枚举,为确保无人机的正确使用、保证使用过程中的安全稳定,现在迫切需要通过无人机的识别和跟踪实现对无人机的管控^[2]。

目前,无人机的探测方法从信号形式上可以分为:雷达探测、激光探测、金属探测、光电探测、无线电探测和声学探测,主要基于雷达、图像和无线电等方式。由于无人机体积小、雷达反射面积小,导致雷达难以探测到无人机,且无人机产生的回波与飞鸟等干扰物特征相近,增加了识别难度;其次无线电和雷达探测的技术成本高昂且配置严格。对比其他的探测方式,基于深度学习和机器视觉的方法具有成本低、配置简单、准确度高等优点。复杂背景下的无人机目标检测和跟踪任务具有检测背景复杂、检测目标不显著的特点。

刘璐琦等^[3]提出了基于HSV(hue, saturation, value)色彩空间的阴影识别算法,能够对检测对象阴影区域进行分割识别,从而排除阴影对目标检测的干扰。薛远亮等^[4]针对跟踪过程中目标的尺寸小、尺度变化大和相似物干扰等问题,提出了一种基于多尺度融合的目标跟踪算法。李悦等^[5]针对粗瞄准过程中如何快速、准确识别出激光通信设备、辅助完成激光通信,提出将不同目标检测算法与长时跟踪核化相关滤波器(kernelized correlation filters, KCF)算法相结合。以上方法都有效解决了无人机检测相关问题,但是实时性上都相对较差。

薛珊等^[6-7]针对在复杂飞行背景的公共安全区域内无人机实时检测困难的问题,提出了基于注意力机制和尺度自适应特征融合的YOLOv5(you only look once, YOLO)反无人机系统目标检测算法。但他们的算法无法对小目标和不显著目标做到有效识别。张聪等^[8]利用红外小目标在有限的区域内拥有比邻近区域更高的灰度值特点,提出密度与距离概念,建立密度距离空间来检测目标。Xu等^[9]对YOLOv8算法提出了改进,以提高其在复杂环境中对小型、部分遮挡物体的检测精度和速度,但此算法平均类别精度 $mAP@0.5$ (mean of average precision, mAP)仅有37.6%,无法提供足够的检测精确度。以上方法解决了遮挡或者背景复杂问题下的无人机识别,但精度有待进一步提高。

上述提出方法大多是在解决常规识别跟踪问题,且都是独立的跟踪或者识别问题^[10]。经过归纳,多是通过引入注意力机制、替换损失函数实现对YOLO算法的改良,这些改进会使得模型参数量增大,使得算法的帧率无法满足实时性要求。普遍存在检测速度低、实时性差、无法对小目标和不显著目标做到有效识别、存在漏检、检测的成功率和准确率较低的问题。

针对现有无人机识别跟踪算法无法解决小目标、不显著目标的准确检测问题及目标被遮挡后的再跟踪问题,文中提出一种引入梯度特征改进并融合深度学习识别与跟踪算法的目标识别跟踪方案。通过将YOLOv10和KCF算法的优点进行有效融合,将多通道梯度特征有效的融入到算法中,使得无人机识别跟踪更加有效鲁棒。本研究改进算法实时性高,可实现再跟踪和短时间被遮挡时的持续跟踪。通过设计一种多尺度检测,使KCF算法具有良好的尺度自适应,缓解了KCF算法尺度变化缺陷;将多通道梯度特征引入YOLOv10算法中,使用特征级联的方法保留原图特征和多通道梯度特征。特征级联可形成一个更丰富的输入,增加了输入YOLO算法中特征的维度,使YOLOv10网络能够同时利用原图的语义信息和梯度信息;使得YOLOv10算法具有更好的特征提取和检测能力。

1 算法设计

1.1 梯度改进的YOLOv10模块

YOLOv10算法将输入图像划分为网格,每个网格预测一定数量的边界框和对应的类别概率^[11]。传统的YOLO

模型使用非极大值抑制 (non-maximum suppression, NMS) 来过滤重叠的预测,这导致了推理时间的延长^[12-13]。为缩短推理时间,算法引入了一种双重分配策略,消除了 NMS 的需求,实现了更快、更高效的目标检测。

算法采用的双重标签分配策略,用于不需 NMS 的 YOLO 训练,从而在推理时不需要使用 NMS。此策略既保留了丰富的监督信号,又大幅提高推理效率。这样既保证了性能优势,也进一步降低了推理延迟。新算法引入双重标签分配策略,结合了两者的优势,扬长避短。实现了高效和准确的端到端部署。为了实现一致的监督,新算法在一对多标签分配和一对一标签分配中采用了统一的匹配度量,如式 (1) 所示。

$$m(\alpha, \beta) = s \times p^\alpha \times IoU(\hat{b}, b)^\beta \quad (1)$$

其中, p 为分类得分, $\hat{b}$ 和 b 分别代表预测框和实例的

边界框, s 代表空间先验, α 和 β 是平衡语义预测任务和位置回归任务影响的超参数。

同时,采取了新的轻量级分类头设计,减少了计算时冗余从而提高了模型的计算效率。此改进不仅降低了计算复杂度,还增强了特征的提取能力,实现更高效的参数利用和更加优异的性能。通过空间-通道解耦下采样策略,算法可以在保持精度的同时,降低计算复杂度。并且为了增强模型的特征提取能力,算法引入了大核卷积和部分自注意模块,在较低计算性能下提高了识别效率。紧凑反转块 (compact inverted block, CIB) 模块结构图如图 1 中 CIB 模块所示,深度卷积 (depth wise convolution, DW) 可对每个输入通道分别执行卷积操作,从而减少计算量;部分自注意模块 (partial self-attention module, PSA) 结构图如图 1 中右下 PSA 模块所示。

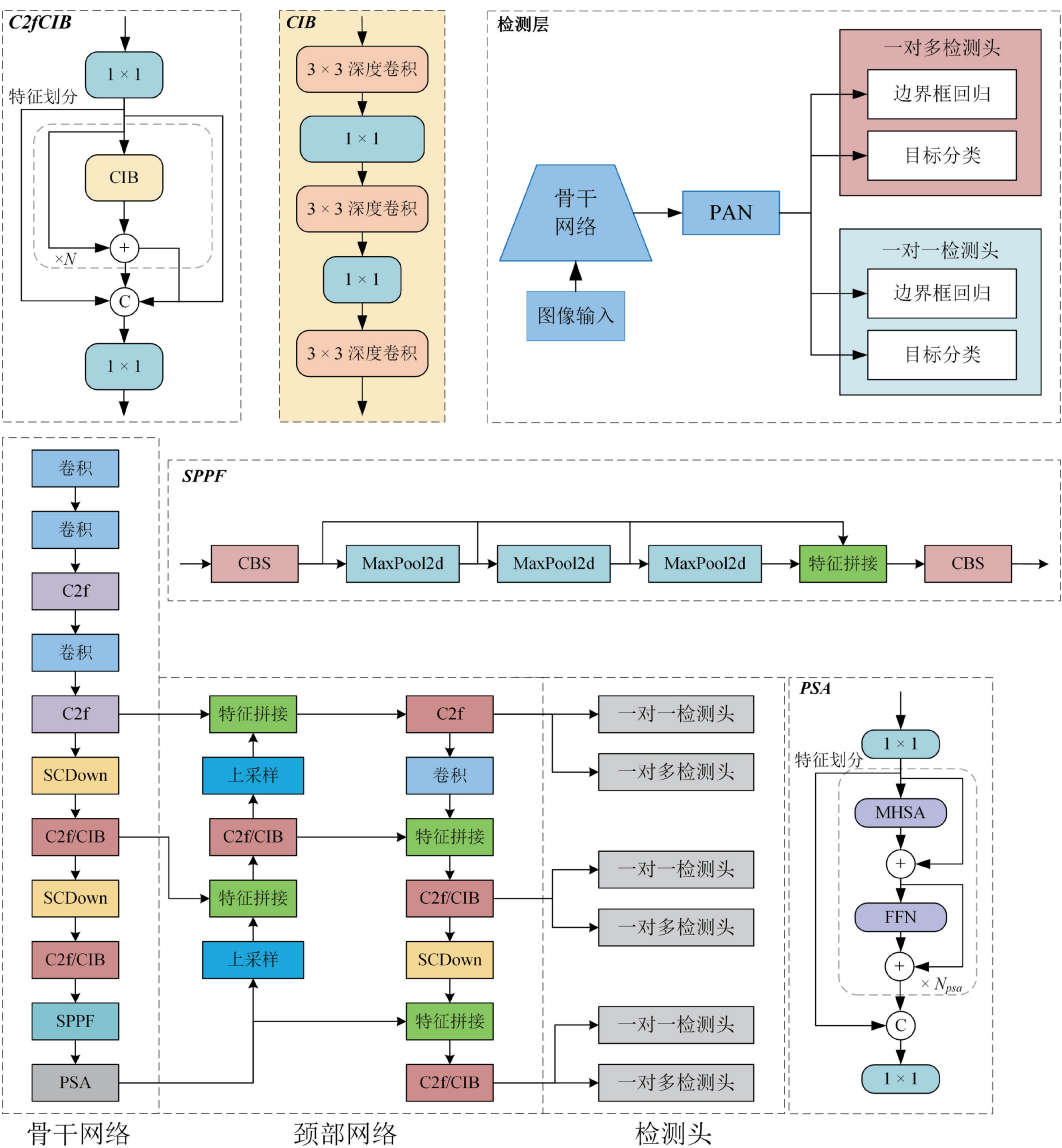


图 1 CIB 模块结构(上)和部分自注意模块(下)

Fig. 1 CIB module structure (top) and some self-attention modules (bottom)

在图 1 中,骨干网络中的跨阶段部分特征融合(cross-stage partial fusion, C2f)采用特征拆分与特征拼接,可优化计算效率;快捷下采样模块(short cut down sampling, SCDn)是一种下采样方法,能够快速做到下采样,在减少计算量的同时保留更多的细节信息;快速空间金字塔池化(spatial pyramid pooling-fast, SPPF)是空间金字塔池化(spatial pyramid pooling, SPP)的优化版本;逐点空间注意力(point-wise spatial attention, PSA)是一种注意力机制,强调对特征图中逐个点进行权重分配,用来增强重要特征。在逐点空间注意力中的多头自注意力(multi-head self-attention, MHSA)和前馈神经网络(feed-forward network,

FFN)分别负责全局特征提取能力和特征表达能力的增强。在强光照、阴影状态和待检测目标与背景区分不显著的情况下, YOLOv10 算法的识别具有局限性,在城市中对进行无人机跟踪的过程中,上述状态并不少见^[14-16]。YOLOv10 算法在背景情况下对待检测目标的识别率低,故对 YOLOv10 算法进行改进。将多通道梯度特征进行提取后使用特征级联这一思想将其与原图进行拼接,特征级联是一种在机器学习和计算机视觉中常用的技术,尤其在目标检测领域。它的基本思想是通过逐步筛选和处理特征来提高检测效率,减少计算量,其核心优势在于其高效性。梯度特征提取流程如图 2 所示。

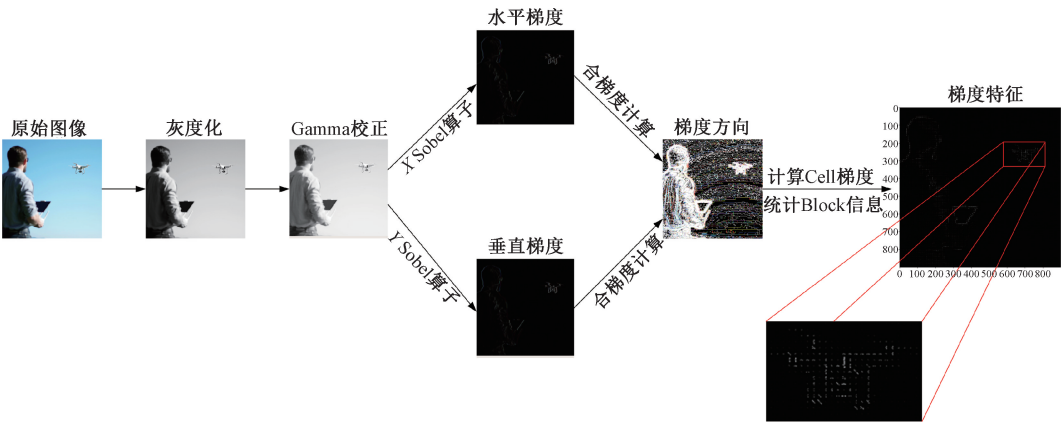


图 2 梯度特征提取流程
Fig. 2 Flowchart of gradient feature extraction

使用特征级联将梯度特征图 and 原图的特征图拼接形成一个新的多通道特征图并输入至骨干网络中。YOLOv10 的网络结构通常包含多个卷积层,能处理更高维度的输入。所以,将梯度特征图 and 原图的特征图沿着通道维度进行拼接,得到一个具有更多通道和细节的新特征图。通过特征级联, YOLOv10 网络可以同时利用原图信息和梯度信息,充分挖掘图像中的边缘、纹理等特征,且这种方式不会丢失原图中的信息。此改进添加在骨干网络进行第 1 次卷积前,将更多的细节和特征信息引入 YOLOv10 算法之中,使算法对强光照或阴影下的目标有更好的检测效果,多特征提取和特征级联的效果图如图 3 所示,结构改进图如图 4 所示。

在梯度特征使用过程中,会将三通道的彩色图像转换成一定长度的特征向量,梯度算法会对图像进行预处理,将图像进行裁剪并且缩放到固定尺寸,再计算三通道颜色值的梯度,之后取三通道颜色值中梯度值最大的作为像素的梯度。取得梯度值之后,再进行伽马校正、调节对比度等。伽马校正公式如式(2)所示。其中输出图像 $f(x)$ 是输入图像 x 的幂函数,指数为 γ 。

$$f(x) = x^\gamma \tag{2}$$

梯度直方图(histogram of oriented gradients, HOG)是一种通过计算图像局部区域的梯度并将其量化为直方图的特征描述方法,用于捕捉图像中的形状信息。为得到梯度直方图,需要计算水平梯度及垂直梯度,由行列式 $\begin{bmatrix} -1 & 0 & 1 \end{bmatrix}$ 即可计算水平梯度;由行列式 $\begin{bmatrix} 0 & 1 & 1 \end{bmatrix}$ 可以计算垂直梯度,使用内核大小为 1 的索贝尔算子来实现水平梯度和垂直梯度的计算。

在计算完水平梯度和垂直梯度后,使用式(3)和(4)计算梯度的大小和方向。

$$g = \sqrt{g_x^2 + g_y^2} \tag{3}$$
$$\theta = \arctan \frac{g_y}{g_x} \tag{4}$$

其中, g 代表梯度的模长, g_x 和 g_y 分别代表水平梯度和垂直梯度, θ 代表梯度的夹角。水平方向的梯度图会强化垂直边缘特征,垂直方向的梯度图会强化水平边缘特征。强化垂直边缘特征和水平边缘特征可以使重要的特征得以保留,去除无关的信息。

计算梯度直方图并且进行归一化,将图像划分为若干个 8×8 单元,计算每个单元的梯度直方图;梯度将

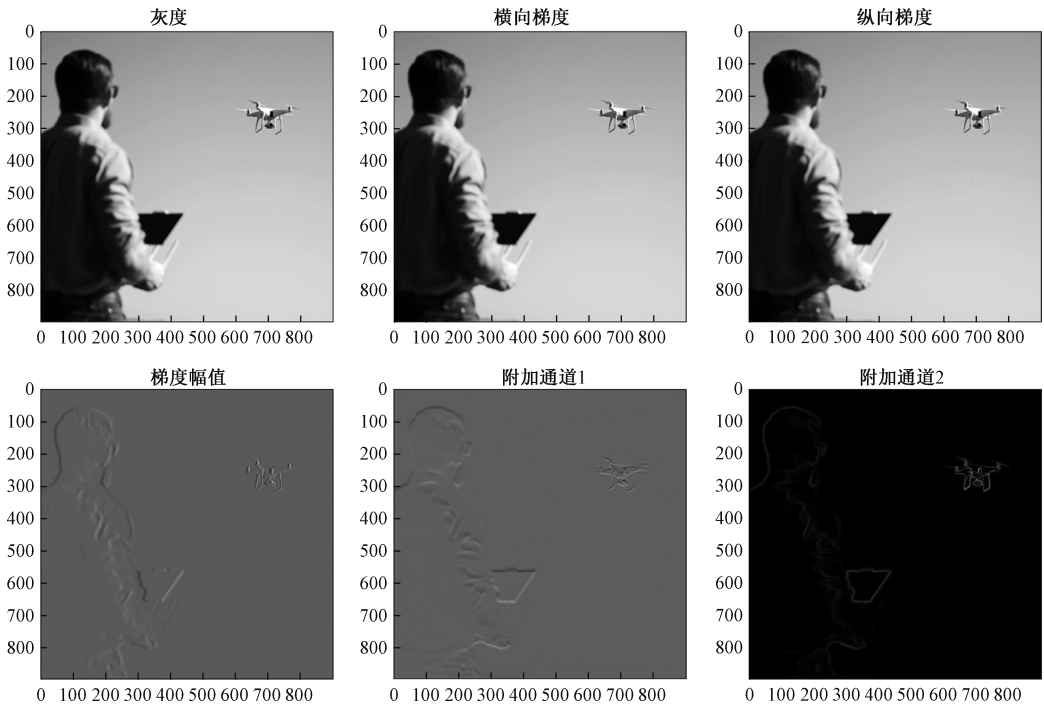


图 3 多特征提取和特征级联

Fig. 3 Multi-feature extraction and feature concatenation

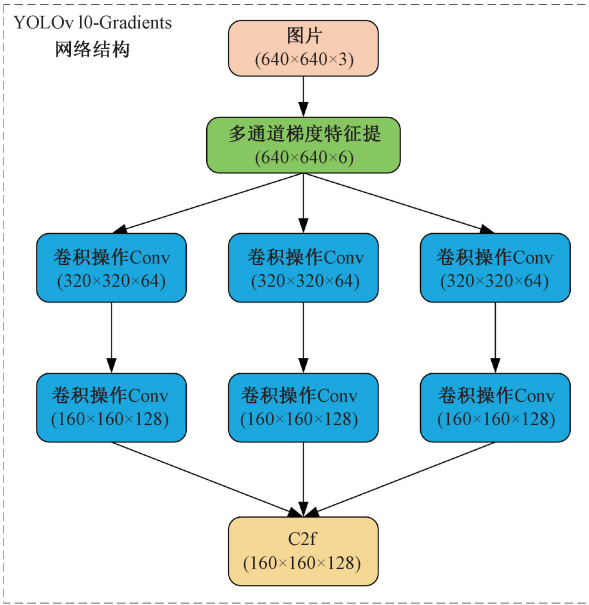


图 4 YOLOv10 改进结构

Fig. 4 Improved YOLOv10 architecture

2×2 个单元合并为一个块 (block)。为减少光照变化对特征描述符的影响,将直方图进行归一化操作。归一化之后计算梯度特征向量,将整幅图像划分成 8×16 单元。每个 block 有 2×2 个单元,block 的个数为 (16-1)×(8-1)=105 个,即有 7 个水平方向的 block 和 15 个垂直方向的 block。

将这 105 个 block 合并,就得到了整个图像的特征描述符,维度为 105×36=3 780。

1.2 梯度改进的 KCF 目标跟踪模块

KCF 算法是一种基于相关滤波器的目标跟踪算法,它将目标的外观建模为一个核化的相关滤波器问题,该算法具有高效率 and 较强实时性的优点。KCF 算法采用了核技巧,使其能在频域中高效地进行计算。KCF 算法的流程主要分为:初始化目标、提取目标特征、训练相关滤波器、傅里叶变换优化滤波器、目标跟踪与位置预测和更新模型。初始化目标是指在视频的第 1 帧中,需要选择目标区域,即标记目标的初始位置和大小。此区域被用作训练样本,提取特征进行目标建模。提取目标特征是指对第 1 帧图像中的目标区域进行特征提取,目的是捕获目标的外观信息,用于后续滤波器的训练。训练相关滤波器是指通过训练一个相关滤波器来建模目标的外观。KCF 算法使用一种基于岭回归的方法,通过最小化一个损失函数来优化滤波器。当样本数不足,输入数据不是满秩矩阵时,此时将导致非满秩矩阵 $X^T X$ 无法直接求得逆矩阵,岭回归是在矩阵 $X^T X$ 的基础上加入一个正则项 λI ,使矩阵非奇异,对 $X^T X + \lambda I$ 求逆,从而求得回归模型中的回归系数 $\hat{w}$,如式(5)所示。

$$\hat{w} = (X^T X + \lambda I)^{-1} X^T y \quad (5)$$

在 KCF 算法中,使用傅里叶变换优化滤波器可以使用循环矩阵来辅助运算的,循环矩阵可以减少大量的运

算量。循环矩阵是一种特殊形式的托普利兹矩阵,满足 $X_{i,j} = X_{i+1,j+1} = x_{i-j}$ 的矩阵即可成为托普利兹矩阵。托普利兹矩阵示例如式(6)所示。

$$X = \begin{bmatrix} x_0 & x_{-1} & x_{-2} & \cdots & x_{-(n-1)} \\ x_1 & x_0 & x_{-1} & \cdots & x_{-(n-2)} \\ x_2 & x_1 & x_0 & \cdots & x_{-(n-3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n-1} & x_{n-2} & x_{n-3} & \cdots & x_0 \end{bmatrix} \quad (6)$$

在训练样本中的基样本为正样本,其他样本为虚构出来的负样本,这样的样本构成可以很方便的利用快速傅里叶变换(fast Fourier transform, FFT)和傅里叶对角化的性质来进行计算,从而不需要知道负样本的具体形式。

KCF 算法利用目标所在区域形成循环矩阵,再利用循环矩阵傅里叶对角化的性质,通过岭回归得到滤波模板。循环矩阵 X 可以被离散傅里叶矩阵 F 对角化处理,如式(7)所示。

$$X = C(x) = F \cdot \text{diag}(\hat{x}) \cdot F^H \quad (7)$$

其中, $\hat{x}$ 是 x 的离散傅里叶变换,如式(8)所示。 $C(x)$ 是由基样本 x 生成的循环矩阵, F^H 是 F 的厄米特转置, F 是离散傅里叶变换的常量矩阵,形式如式(9)所示。

$$\hat{x} = F(x) = \sqrt{n} \times F \cdot x \quad (8)$$

$$F = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & \omega & \cdots & \omega^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{n-1} & \omega^{(n-1)(n-2)} & \omega^{(n-1)^2} \end{bmatrix} \quad (9)$$

将循环矩阵 X 代入至式(5)中,可得到回归系数 $\tilde{w}$ 如式(10)所示。再利用反对角化性质,可得式(11)。进一步使用循环矩阵卷积性质 $F(C(x)y) = F(x) * F(y)$ 得到式(12)。

$$\tilde{w} = \left(F \cdot \text{diag} \left(\frac{\hat{x}^*}{\hat{x}^* * \hat{x} + \lambda} \right) F^H \right) y \quad (10)$$

$$w = C \left(F^{-1} \left(\frac{\hat{x}^*}{\hat{x}^* * \hat{x} + \lambda} \right) \right) y \quad (11)$$

$$F(w) = \left(\frac{\hat{x}^*}{\hat{x}^* * \hat{x} + \lambda} \right)^* * F(y) = \frac{\hat{x}^* \hat{y}}{\hat{x}^* * \hat{x} + \lambda} \quad (12)$$

在使用滤波器进行检测的过程中,KCF 算法利用训练好的跟踪器对选定区域进行滤波计算,得到分布图^[17]。之后将分布图中的最大响应位置作为预测目标的中心位置。由上式可解得 α , 输入样本和模板样本在多维空间的核矩阵形式如式(13)所示。

$$\alpha = (K + \lambda I)^{-1} y \quad (13)$$

记构造循环矩阵式(14),将式(15)进行化简并求其傅立叶变换,可推导求得式(16)。

$$K^2 = C(k^{xz}) \quad (14)$$

$$f(z) = \begin{bmatrix} k^{x_1 z} \\ k^{x_2 z} \\ \vdots \\ k^{x_n z} \end{bmatrix}^T \alpha \quad (15)$$

$$\hat{f}(z) = \hat{k}^{xz} * \hat{a} \quad (16)$$

式(14)中构造的 x 和 z 的核相关矩阵 k^{xz} ,用于进行回归检测。其中 $f(z)$ 是岭回归的结果, $\hat{f}(z)$ 为输出响应方程。

在 KCF 算法中分别引入单通道灰度特征和多通道特征,两个改进算法经过对比,其速度差异较小。但单通道灰度算法对光照变化较为敏感,容易导致跟踪或识别不稳定;忽略了方向性和边缘信息,导致区分能力差;在复杂背景中,灰度特征容易受到噪声和背景干扰的影响,难以应对复杂场景中的目标等缺点。作为对比,多通道梯度特征鲁棒性强、对光照变化的敏感性较低、区分能力强并且表现力更丰富。综合考虑应用背景,选取多通道特征对 KCF 算法进行改进。多通道特征本身不具有自动适应目标大小变化的能力,因此在面对特征目标尺寸变化时,多通道 KCF 算法无法准确的跟踪目标,会导致产生较大的误差。这是因为相关滤波器的训练过程是基于特定尺度下的样本,当目标尺度变化时,目标框尺度无法自适应,会造成无关特征进入识别区域造成目标框漂移,从而导致跟踪失败。为解决尺度自适应的问题,在多通道特征改进的 KCF 算法中引入多尺度检测,使得识别框对目标远近移动产生的尺度缩放具有良好的尺度自适应性。

在多尺度检测中,算法会在不同的缩放级别下对目标进行检测。这是通过在不同尺度上分别提取模板并计算相应的特征来实现的。具体而言,算法会使用不同的缩放系数来调整目标框的大小,对每一个尺度下的目标,都会计算一个匹配度。该匹配度主要评估目标框是否需要调整。在识别过程中,如果某个尺度下的匹配度更好,目标框的大小会按比例进行调整。例如:如果目标在较小尺度上的匹配效果较好,目标框会被缩小,从而更精确地围绕目标;反之,如果目标在较大尺度上的匹配效果较好,目标框会被放大,以适应目标的实际大小。

多通道特征通过增强对目标的描述能力,确保目标在图像中产生变化时,目标仍然能被有效识别。多尺度检测使得目标框能够根据不同尺度下的匹配结果动态调整目标框的大小,使目标框始终保持在最佳的尺寸范围内。多通道特征和多尺度检测的结合,解决了 KCF 算法在识别时的检测框漂移和尺度自适应的问题。该改进确保了目标框能够灵活且精确地随着目标的变化进行缩小或放大,从而提高了目标跟踪的准确性。

YOLOv10 算法在跟踪过程中遭受相似目标干扰或多目标识别时会输出多个结果,多输出结果时使用坐标计算和坐标比较时会造成误跟踪,无法持续跟踪某一目标,YOLOv10 改进算法与 KCF 算法可以弥补此缺点。KCF 跟踪算法的优势是采取在线训练策略,此策略不需要实现对大量目标样本进行训练,只需要在初始帧中框选出被跟踪的目标,在跟踪的过程中不断更新迭代滤波器,使用该滤波器确定下一帧的目标位置即可实现跟踪。由于该算法是基于多尺度的目标检测,在一定程度上克服了检测过程中目标尺度变化和目标遮挡带来的影响,在目标检测过程中具有很好的准确性和鲁棒性。如上文所述,YOLO-G-KCF 算法对比 YOLOv10 算法和 KCF 算法可提高跟踪系统的鲁棒性。

2 实验验证

2.1 实验环境和参数

使用 YOLOv10m 模型对数据集进行训练,训练时使用命令行界面命令 (command-line interface, CLI) 进行训练。本研究的实验环境为自主搭建电脑,系统环境为 Ubuntu 18.04 系统,处理器为 AMD R7-5800X@ 3.8 GHz, 48 GiB 运行内存,GPU 使用 NVIDIA GeForce RTX3060Ti 8 GB 显存。使用 Visual Studio Code 编写程序,Qt Creator 编写窗口界面,调用 YOLOv10、OpenCV-Python、Torch、NumPy、PySerial、Uservo 等第三方库。其中 Keftracker 经过对开源代码的重构,可在 Python3 环境下将梯度特征融入至 KCF 算法中。为训练数据集所构建的网络模型基于机器学习和深度学习的开源深度学习框架 PyTorch,开发环境为 PyTorch 1.12.1、CUDA11.3、OpenCV 4.3.0.36、Python3.8。

2.2 数据集和训练结果

为保证 UAV 在公共环境等干扰较大的环境中 and 以上情况时,依旧可以保持良好的检测效果,故需要大量不同种类的无人机照片、无人机被遮挡和检测目标与背景区分不明显 (检测目标不显著) 的照片。由这些照片组成本研究所需的无人机数据集,用于 YOLO-G-KCF 模型的训练与评估,改进算法选用 YOLOv10m 模型。

本研究数据集来源于 GitHub 上的开源数据集 DUT-Anti-UAV、A Deep Learning Approach to Drone Monitoring、Anti-UAV 和其他开源数据集中符合要求的部分图片^[20]。本研究数据集包含不同角度的无人机图片、与背景相似的无人机图片和小目标的无人机图片共计 21 027 张,将图片使用软件 Labelimg 进行标注从而组成数据集。数据集按 7:2:1 的比例划分为训练集、测试集和验证集,其中

14 719 张作为训练集,4 205 张作为验证集,2 103 张作为测试集。经过实验,数据集训练参数为轮次 (epochs) 设置为 400 epochs、耐心值 (patience) 设置为 100 轮、一次性处理的图片数量 (batch) 设置为 -1 (自动模式)、输入图像大小 (imgsz) 设置为 640 pixels、数据加载线程数 (workers) 设置为 24 workers。在数据集训练时,由于设置耐心值 (patience) 为 100 轮,当 100 轮内训练结果没有进步时,程序会提前结束训练,数据集训练 327 轮时自动停止训练。在训练过程中,随着轮次的增加精确率上升速度很快,根据数据显示在 26 轮时达到 86.55%,根据轮数的增加,精确率波动上升,最终稳定在 95.3%。使用 $mAP@0.5$ 作为指标表示平均精度,YOLO-G-KCF 算法的 $mAP@0.5$ 达到了 94.28%。YOLO-G-KCF 模型对比 YOLOv10m 模型在训练相同轮次的情况下,YOLO-G-KCF 的精准率高出 YOLOv10m 约 5.5%,平均精度 $mAP@0.5$ 高出 1.37%。在检测精度上,YOLO-G-KCF 算法高于 YOLOv10m,在 $mAP@0.5$ 、精确率、召回率上更具优势,检测精度更高。综合对比两算法,本研究算法对小目标识别准确率高、识别精度高的优点。

选取小目标、与背景区分度不高的照片作为验证照片,验证集共有 4 205 张图像。模型检验结果如图 6 所示,数据集训练结果如图 7 所示。在图 7 中,边界框遗漏损失 (bounding box omission loss, box_om) 代表模型遗漏目标的边界框损失;边界框重叠损失 (bounding box overlap loss, box_oo) 常用于衡量预测边界框与真实框之间的重叠误差;分类遗漏损失 (classification omission loss, cls_om) 指在模型分类时某些目标没有被正确分类;分类重叠损失 (classification overlap loss, cls_oo) 代表目标类别被错误预测为其他相似类别的重叠损失。模型评估指标 (metrics) 用于衡量模型检测和分类的综合表现,其数值越高越好。

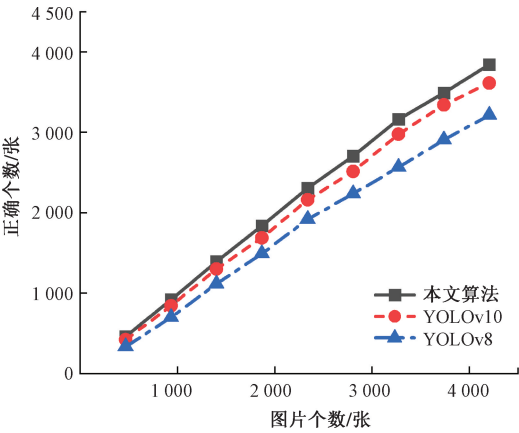


图 6 各算法识别结果

Fig. 6 The recognition results of each algorithm

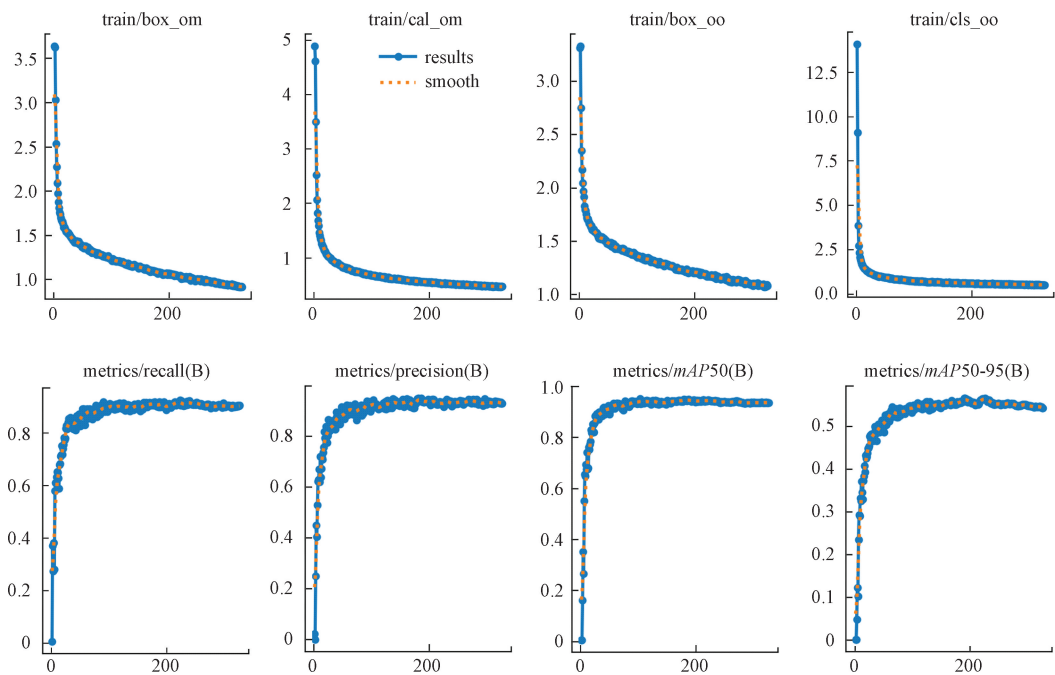
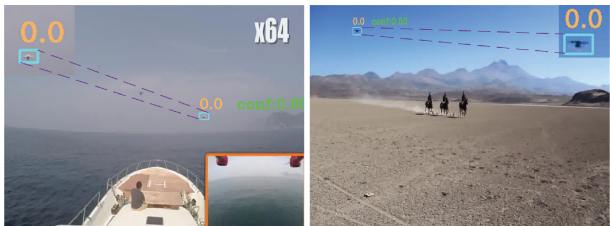
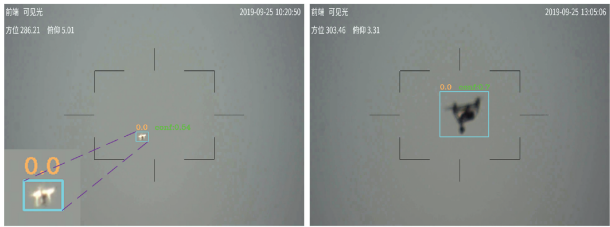


图 7 数据训练结果
Fig. 7 Data training results

由验证结果可知改进模型在小目标和检测目标不显著的情况下有很好的检测效果。验证集的实验结果如图 8 所示,验证数据如图 9 所示,经验证集验证该模型在小目标、阴雨背景下可以实现较好的检测效果且置信度均高于 0.5。



(a) 小目标图片验证效果
(a) Small objective validation



(b) 光电吊舱验证效果
(b) Optical pod validation

图 8 验证集实验效果
Fig. 8 Data validation results

2.3 实验过程和结果

为方便进行测试和使用,将各算法编写进 QT 窗口中统一调用。为检测算法的跟踪效果,故使用多段视频对算法进行检测,系统图像交互界面如图 10 所示。将

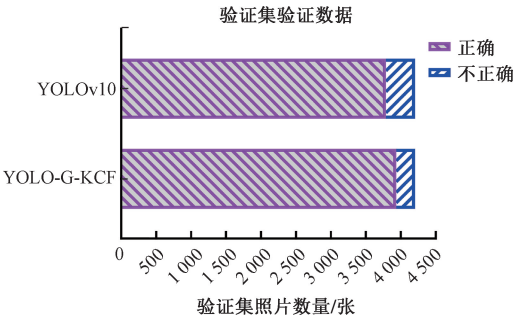


图 9 验证数据
Fig. 9 Validating the results of the dataset

算法与现有方法进行了对比,比较了算法运行精度和运行速度,如表 1 所示,给出了不同检测算法性能的对比。由表 1 可知,在小目标和检测目标不显著的情况和其他场景下,YOLO-G-KCF 算法单个类别平均精确度 (average precision, AP) 高于其他算法、识别误差较小、跟踪正确率较高、稳定性好,抗干扰能力强。各算法平均精确度从低至高排序分别为:快速区域卷积网络 (fast region convolutional neural networks, Fast R-CNN)、单发多箱探测器 (single shot multi-box detector, SSD)、YOLOv8、YOLOv10 和本研究 YOLO-G-KCF 算法。

平均帧率 (frames per second, FPS) 反应各算法的识别速度, FPS 从低至高排序分别为: Fast R-CNN、SSD、YOLOv8、本研究 YOLO-G-KCF 算法和 YOLOv10。因为本研究数据集仅有无人机单个类别,故本模型只需要使



图 10 系统图像交互界面

Fig. 10 System image interaction interface

表 1 不同检测算法性能比较

Table 1 Comparison results of performance of different detection algorithms

检测算法	AP/%	FPS
SSD	79.23	62
Fast R-CNN	74.64	42
YOLOv8	89.85	118
YOLOv10	92.91	131
YOLO-G-KCF	94.28	112

用单类别识别。由结果可知, YOLO-G-KCF 在无人机跟踪过程中平均精确度高于 YOLOv10, 平均识别速度低于 YOLOv10 算法, 表明了本研究算法在些许牺牲识别速度的基础上相比其他算法提高了识别精确度, 体现了优秀的实时性和准确性。

改进算法和原算法检测效果如图 11 所示。在目标不显著、小目标检测、阴天小目标和城市阴天背景下, 改进算法可以有更好的置信度和识别速度。同时, 图 12 还给出了目标快速移动和遮挡干扰下的再跟踪情况结果图。如图 12 所示, 在跟踪过程中目标发生快速移动、被遮挡和干扰时, 本研究改进算法可以准确和快速的重新跟踪目标。由实验表明, 该算法抗干扰能力强, 实时性好; 目标被遮挡后再跟踪时可快速响应。由实验结果可以表明在预设场景下融合算法 YOLO-G-KCF 在遇到小目标、目标与背景区分不显著、干扰物干扰和快速移动的情况下可以准确识别, 同时在跟踪错误时可以实现自动纠正, 体现了较高的鲁棒性。

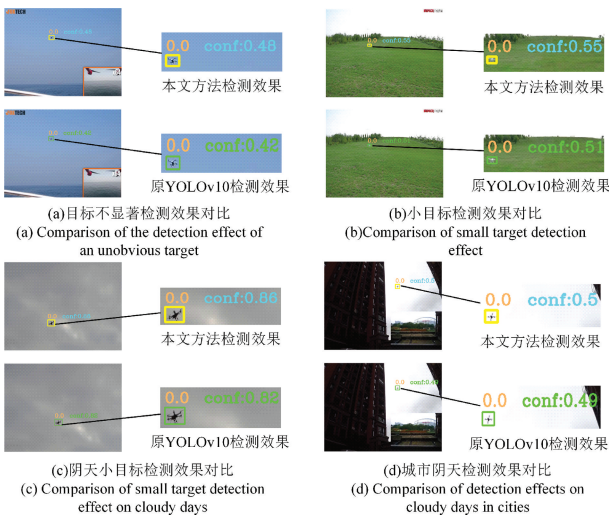


图 11 算法检测效果

Fig. 11 Algorithm detection effect images

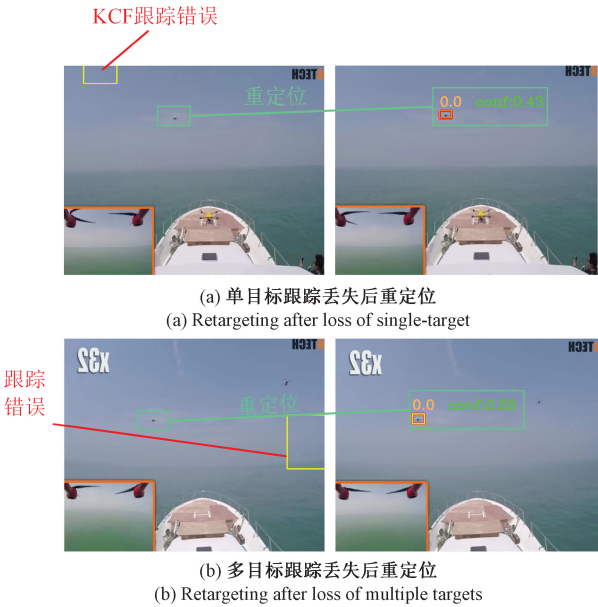


图 12 跟踪丢失后再跟踪

Fig. 12 The result images of tracking after loss

为了更好地评估本研究算法各项改进模块的有效性, 使用自制的无人机数据集对基于 YOLOv10 模型的各个改进部分进行了消融实验^[21-22]。逐步添加各项改进, 共设计 6 组实验对不同的改进部分进行分析, 实验结果如表 2 所示。本研究 YOLO-G-KCF 算法的精确率、召回率和平均精确度相对于原 YOLOv10 算法得到了提高, 在小目标、检测目标不显著的情况及目标被遮挡后的再跟踪情况下, YOLO-G-KCF 算法相比于 YOLOv10 分别提高了 7.5%、7.1%、1.37%。使用模型对含遮挡物的连续视频进行测试, 测试结果如图 13 所示^[23-24]。

从图 13 中可知,YOLO-G-KCF 算法具有更快的识别速度以及在遮挡干扰下良好的目标持续跟踪效果。对比原 YOLOv10 算法,可以发现本改进算法在目标被遮挡跟踪时,置信度虽不进行更新,但依旧可保持跟踪;且本改进算法虽帧数低于 YOLOv10 算法,但依旧具备良好的实时性。

表 2 消融实验

Table 2 The results of ablation experiment

算法	精确度 /%	Recall /%	mAP@ 0.5 /%	FPS
YOLOv10	89.8	88.1	92.91	131
YOLOv10(梯度)	90.7	86.2	90.47	124
YOLOv10+KCF	92.4	90.2	91.85	128
YOLOv10+KCF(梯度)	95.1	91.7	92.14	119
YOLOv10(梯度)+KCF	93.3	93.4	92.76	121
YOLO-G-KCF	95.3	95.2	94.28	112

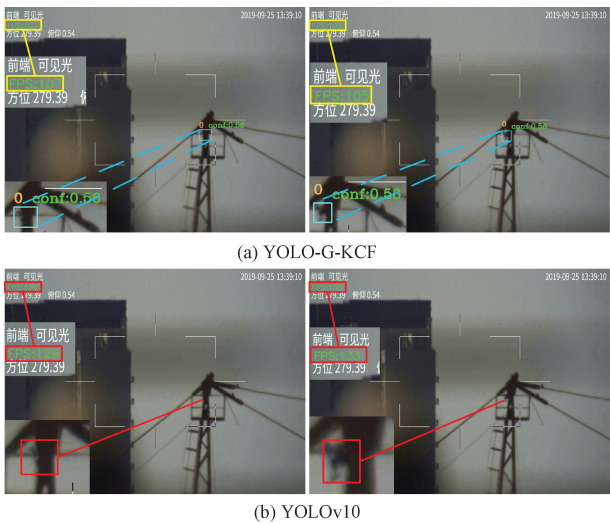


图 13 遮挡情况下跟踪效果

Fig. 13 Algorithm detection effect image Tracking effect diagram under occlusion

3 结 论

本研究设计了一种针对小目标、检测目标不显著的情况和目标被遮挡后的再跟踪问题的算法 YOLO-G-KCF。该算法基于 YOLOv10 算法,可实现对复杂背景下无人机的检测和跟踪。该算法将梯度特征和原图特征通过特征级联的方式引入 YOLOv10 算法之中,使改进算法对强光照或阴影下的目标有更好的检测效果;同时将梯度特征引入 KCF 目标跟踪算法之中,通过设计一种多尺度检测,使 KCF 算法具有良好的尺度自适应;

在 Head 头后引入 KCF 跟踪结果,计算得新的损失函数 Inner-IoU。YOLO-G-KCF 算法可弥补传统视觉在目标被遮挡或其他原因导致的目标丢失情况,也可弥补深度学习算法在小目标检测时某一帧目标丢失的情况,降低在跟踪过程中被跟踪目标某一帧丢失情况的发生概率。

测试结果表明,跟踪系统运行稳定,准确率达到 95.3%。在小目标、背景与目标区分度不高和目标遮挡后的再识别方面具有很好的鲁棒性,能够实现设定场景下对无人机的检测和跟踪。YOLO-G-KCF 算法识别准确率高,抗干扰能力强,实时性好;目标被遮挡后再跟踪时可快速响应,对其他飞行物的抗干扰能力强,鲁棒性高,具有一定的理论研究和工程应用价值。

参考文献

[1] 姚 研,刘 琼,谭 智 诚. 基于软件无线电的无人机入侵检测方法研究[J]. 电子测量技术,2023,46(2):101-110.

YAO X, LIU Q, TAN ZH CH. Research of UAV intrusion detection method based on software defined radio[J]. Electronic Measurement Technology, 2023, 46(2):101-110.

[2] 吴一全,童 康. 基于深度学习的无人机航拍图像小目标检测研究进展[J/OL]. 航空学报, 1-28[2025-02-13].

WU Y Q, TONG K. Research advances on deep learning-based small object detection in UAV aerial images[J]. Acta Aeronautica et Astronautica Sinica, 1-28[2025-02-13].

[3] 刘 珞 琦,王 红 雨,韩 佼 志. 面向无人机目标的检测与实时跟随[J]. 计算机工程与应用,2024,60(11):319-327.

LIU R Q, WANG H Y, HAN J ZH. Target detecction and real-time following for unmanned aerial vehicle[J]. Computer Engineering and Applications, 2024, 60(11):319-327.

[4] 薛 远 亮,金 国 栋,谭 力 宁,等. 基于多尺度融合的自我适应无人机目标跟踪算法[J]. 航空学报, 2023, 44(1):209-226.

XUE Y L, JIN G D, TAN L N, et al. Adaptive UAV target tracking algorithm based on multi-scale fusion[J]. Acta Aeronautica et Astronautica Sinica, 2023,44(1):209-226.

[5] 李 悦,马 志 勇,俞 建 杰,等. 融合 YOLOV5+DSST+KCF 的目标检测跟踪算法在激光通信中的应用[J]. 计算机测量与控制, 2023,31(10):49-53,75.

LI Y, MA ZH Y, YU J J, et al. Application of target detection and tracking algorithm combining YOLOV5+DSST+KCF in laser communication[J]. Computer

- Measurement & Control, 2023,31(10):49-53,75.
- [6] 薛珊, 张亚亮, 吕琼莹, 等. 复杂背景下的反无人机系统目标检测算法[J]. 吉林大学学报(工学版), 2023,53(3):891-901.
- XUE SH, ZHANG Y L, LYU Q Y, et al. Anti-unmanned aerial vehicle system object detection algorithm under complex background[J]. Journal of Jilin University (Engineering and Technology Edition), 2023,53(3):891-901.
- [7] 薛珊, 王亚博, 吕琼莹, 等. 基于 YOLOX-drone 的反无人机系统抗遮挡目标检测算法[J]. 工程科学学报, 2023,45(9):1539-1549.
- XUE SH, WANG Y B, LYU Q Y, et al. Anti-occlusion target detection algorithm for anti-UAV system based on YOLOX-drone[J]. Chinese Journal of Engineering, 2023,45(9):1539-1549.
- [8] 张骢, 韩自强, 岳明凯, 等. 反“低慢小”无人机红外检测方法研究[J]. 兵器装备工程学报, 2023,44(7):203-208.
- ZHANG C, HAN Z Q, YUE M K, et al. Research on the infrared detection method of anti-“low-slow-small” UAVs[J]. Journal of Ordnance Equipment Engineering, 2023,44(7):203-208.
- [9] XU L Y, ZHAO Y F, ZHAI Y H, et al. Small object detection in UAV images based on YOLOv8n[J]. International Journal of Computational Intelligence Systems, 2024,17(1):223.
- [10] 谭亮, 赵良军, 郑莉萍, 等. 基于 YOLOv5s-AntiUAV 的反无人机目标检测算法研究[J]. 电光与控制, 2024,31(5):40-45,107.
- TAN L, ZHAO L J, ZHENG L P, et al. An Anti-UAV target detection algorithm based on YOLOv5s-AntiUAV[J]. Electronics Optics & Control, 2024,31(5):40-45,107.
- [11] WANG L H, MIAO ZH, LIU EN D. UAV remote sensing detection and target recognition based on SCP-YOLO[J]. Neural Computing and Applications, 2024,36(28):17495-17510.
- [12] QIU Z F, BAI H H, CHEN T Y. Special vehicle detection from UAV perspective via YOLO-GNS based deep learning network[J]. Drones, 2023,7(2):117.
- [13] CAO J SH, BAO W SH, SHANG H X, et al. GCL-YOLO: A GhostConv-based lightweight YOLO network for UAV small object detection[J]. Remote Sensing, 2023,15(20):4932.
- [14] ZHAO ZH H, HE P. YOLO-U: Multi-task model for vehicle detection and road segmentation in UAV aerial imagery[J]. Earth Science Informatics, 2024,17(4):3253-3269.
- [15] 肖选杰, 张浩天, 艾剑良. 空中“低慢小”目标检测跟踪算法的应用研究[J]. 复旦学报(自然科学版), 2023,62(5):605-614.
- XIAO X J, ZHANG H T, AI J L. Application of airborne “low, slow and small” object detection and tracking algorithm[J]. Journal of Fudan University (Natural Science), 2023,62(5):605-614.
- [16] JIANG CH CH, REN H ZH, YE X, et al. Object detection from UAV thermal infrared images and videos using models[J]. International Journal of Applied Earth Observation and Geoinformation, 2022(112):102912.
- [17] 谢家阳, 王行健, 史治国, 等. 动态云台摄像机无人机检测与跟踪算法[J]. 智能系统学报, 2021,16(5):858-869.
- XIE J Y, WANG X J, SHI ZH G, et al. Drone detection and tracking in dynamic pan-tilt-zoom cameras[J]. CAAI Transactions on Intelligent Systems, 2021,16(5):858-869.
- [18] 崔勇强, 李嘉轩, 侯林果, 等. 改进 YOLOv7 的城市小型无人机目标检测方法[J]. 计算机工程与应用, 2024,60(10):237-245.
- CUI Y Q, LI J X, HOU L G, et al. Improved YOLOv7 target detection method for small urban UAVs[J]. Computer Engineering and Applications, 2024,60(10):237-245.
- [19] 方鑫, 朱婧, 黄大荣, 等. 低 SNR 场景下微型无人机跟踪-检测融合方法[J]. 仪器仪表学报, 2022,43(4):79-88.
- FANG X, ZHU J, HUANG D R, et al. Integrated tracking and detection of micro UAV under low SNR environment[J]. Chinese Journal of Scientific Instrument, 2022,43(4):79-88.
- [20] DELLEJI T, SLIMENI F, FEKIH H, et al. An Upgraded-YOLO with object augmentation: Mini-UAV detection under low-visibility conditions by improving deep neural networks[J]. Operations Research Forum, 2022,3(4):60.
- [21] 刘芳, 王洪娟, 黄光伟, 等. 基于自适应深度网络的无人机目标跟踪算法[J]. 航空学报, 2019,40(3):179-188.
- LIU F, WANG H J, HUANG G W, et al. UAV target tracking algorithm based on adaptive depth network[J]. Acta Aeronautica et Astronautica Sinica, 2019,40(3):179-188.
- [22] 孙备, 孙晓永, 钱翰翔, 等. 动态大视角场景融合帧间信息与模板匹配的低慢小无人机目标检测[J]. 仪器仪表学报, 2024,45(7):64-74.
- SUN B, SUN X Y, QIAN H X, et al. Low slow small

UAV targets detection by fused using inter-frame information and emplate matching in dynamic large-view scene[J]. Chinese Journal of Scientific Instrument, 2024,45(7):64-74.

[23] 郝晋渊,张家明,张少康,等. 基于改进 YOLO 的无人机入侵检测方法[J]. 电子测量与仪器学报, 2024,38(7):143-151.

HAO J Y, ZHANG J M, ZHANG SH K, et al. UAV intrusion detection method based on improved YOLO[J]. Journal of Electronic Measurement and Instrumentation,2024,38(7):143-151.

[24] 史雨馨,朱继杰,凌志刚. 基于特征增强 YOLOv4 的无人机检测算法研究[J]. 电子测量与仪器学报, 2022,36(7):16-23.

SHI Y X, ZHU J J, LING ZH G. Research on UAV detection method based on feature enhanced YOLOv4 algorithm[J]. Journal of Electronic Measurement and Instrumentation,2022,36(7):16-23.

作者简介



王文胜(通信作者),2013 年于哈尔滨工业大学获得学士学位,2018 年于中国科学院大学获得博士学位,现为北京信息科技大学副教授,主要研究方向为计算机视觉。

E-mail:ws_wang1128@126.com

Wang Wensheng(Corresponding author) received his B. Sc. degree from Harbin Institute of Technology in 2013, received his Ph. D. degree from University of Chinese Academy of Sciences in 2018. He is currently an associate professor at Beijing Information Science and Technology University. His main research interest includes computer vision.



何君尧,2023 年于北京信息科技大学获得学士学位,现为北京信息科技大学硕士研究生,主要研究方向为图像处理与机器视觉。

E-mail:qiuyu0726@163.com

He Junyao received his B.Sc. degree from Beijing Information Science and Technology University

in 2023. He is currently a M.Sc. candidate at Beijing Information Science and Technology University. His main research interests include image processing and machine vision.



黄民,1987 年于南京理工大学获得学士学位,1990 年于南京理工大学获得硕士学位,1997 年于中国矿业大学获得博士学位,现为北京信息科技大学教授,主要研究方向为机器人智能感知与控制。

E-mail:huangmin@bistu.edu.cn

Huang Min received his B.Sc. degree from Nanjing University of Science and Technology in 1987, received his M.Sc. degree from Nanjing University of Science and Technology in 1990, received his Ph.D. degree from China University of Mining and Technology in 1997. He is currently a professor at Beijing Information Science and Technology University. His main research interests include robot intelligent perception and control.



吴国新,2000 年于北京机械工业学院获得学士学位,2003 年于北京机械工业学院获得硕士学位,2011 年于北京理工大学获得博士学位,现为北京信息科技大学现代测控技术教育部重点实验室研究员,主要研究方向为增强现实在工业领域的应用、机械系统测控与信息化、系统状态监测与故障诊断。

E-mail:wgx1977@bistu.edu.cn

Wu Guoxin received his B. Sc. degree from Beijing Institute of Machinery Industry in 2000, M.Sc. degree from Beijing Institute of Machinery Industry in 2003, and Ph. D. degree from Beijing Institute of Technology in 2011. He is currently an associate researcher at the key laboratory of the Ministry of Education of Beijing Information Science and Technology university. His main research interests include augmented reality in industrial applications, mechanical system measurement and control and information, system status monitoring and fault diagnosis.