

文章编号:1007-2780(2024)04-0532-11

基于主干增强和特征重排的反无人机目标跟踪

郑滨汐*, 杨志钢, 丁钰峰

(哈尔滨工程大学 信息与通信工程学院, 黑龙江 哈尔滨 150001)

摘要: 视频图像中面向无人机的目标跟踪是反无人机任务中的重要一环。无人机低空飞行背景复杂, 同时在视频图像中目标像素占比较小, 都给目标跟踪增加了难度。针对以上问题, 以 SiamRPN++ 为基础, 提出了一种引入改进的主干网络和特征重排的孪生神经网络目标跟踪算法 (SiamAU)。首先, 在主干网络中加入 ECA-Net 注意力机制网络, 同时对激活函数进行改进, 以提升复杂背景下的特征表征能力; 然后, 对主干网络输出的浅层特征进行浅层降维并与后三层深层特征进行融合, 得到更适合无人机等小目标跟踪的改进深度融合特征。在 DUT Anti-UAV 数据集上, SiamAU 算法的成功率和精确率达到了 60.5% 和 88.1%, 相比基准算法提升了 5.6% 和 8.1%。在两个公开数据集上的测试结果表明, 在反无人机场景中 SiamAU 算法的跟踪表现优于目前主流的算法。

关键词: 反无人机; 目标跟踪; 孪生网络; 注意力机制; 特征重排

中图分类号: TP391 **文献标识码:** A **doi:** 10.37188/CJLCD.2023-0150

Anti-UAV object tracking with enhanced backbone and feature rearrangement

ZHENG Binxi*, YANG Zhigang, DING Yufeng

(College of Information and Communication Engineering, Harbin Engineering University,
Harbin 150001, China)

Abstract: Object tracking for the unmanned aerial vehicle (UAV) in videos is an important part of the Anti-UAV task. The complex background during low-altitude flight and the small imaging size are two difficulties for UAV object tracking. A Siamese neural network object tracking algorithm (SiamAU) is proposed, which is based on SiamRPN++ in combination with an improved backbone and a feature rearrangement technique. Firstly, ECA-Net attention module is integrated into the backbone network, while the activation function is improved to enhance the representation ability of convolution features in complex background. Then, channel number of the last three convolution features is rearranged in order to make full use of low-level features that are conducive for small object tracking. The rearranged features are further fused to obtain the improved feature map. Finally, On the DUT Anti-UAV dataset, SiamAU algorithm achieves success and precision scores of 60.5% and 88.1%, an improvement of 5.6% and

收稿日期:2023-04-21;修订日期:2023-05-05.

基金项目:航空科学基金(No.201801P6002);中央高校基本科研业务费(No.3072022CF0802)

Supported by Aeronautical Science Foundation of China(No.201801P6002);Fundamental Research Funds for the Central Universities(No.3072022CF0802)

*通信联系人, E-mail:914016304@qq.com

8.1% in comparison with the baseline algorithm. Extensive experimental results on two public datasets validate that the proposed SiamAU achieves better UAV tracking performance and outperforms previous methods, especially in small object and complex background scenarios.

Key words: Anti-UAV; object tracking; siamese network; attention mechanism; feature rearrangement

1 引言

近年来,小型无人机在各行各业都得到了广泛的应用,与此同时也给低空安全带来了巨大的隐患。为了防止无人机威胁公共安全,反无人机已经成为工业和军事应用中的关键问题,对非法入侵的无人机进行有效的探测跟踪显得尤为重要。目前常见的反无人机方法有音频探测^[1]、雷达探测^[2]、射频探测^[3]以及视频探测^[4]等。音频探测技术在城市环境下使用时易受到背景噪声干扰,且只适用于大型无人机。雷达探测技术针对民用无人机而言散射面积不足。射频探测技术无法针对所有型号的无人机。相比较于前3种方法,基于视频的反无人机探测技术具有距离适中、成本低以及使用场景广泛等特点。

目标跟踪是视频反无人机任务中的重要一环,可以分成基于生成模型^[5]、基于相关滤波^[6]和基于深度网络^[7]的方法。随着SINT^[8]和SiamFC^[9]的出现,一系列基于孪生网络结构的目标跟踪算法取得了长足的发展。SiamRPN^[10]算法在SiamFC算法的基础上,引入了区域生成网络RPN,使跟踪器可以回归目标的位置和形状,提高了尺度估计精度,在去除了多尺度测试的同时,也达到了提升算法跟踪性能和跟踪速度的效果,是典型的基于锚点的目标跟踪方法。与基于锚点的跟踪方法不同,SiamBAN^[11]、SiamFC++^[12]、SiamCAR^[13]、SiamAPN++^[14]等无锚方法直接对目标位置的前景和背景进行分类,并回归对应4个方向的偏移,避免了预定义锚的影响,具有良好的通用性。SiamRPN++^[15]在SiamRPN的基础上引入了深层网络作为特征提取器,解决了深层网络带来的平移不变性问题。在ATOM^[16]、DiMP^[17]算法中,采用共轭梯度策略的同时结合使用深度学习框架对跟踪算法进行快速优化,具有良好的跟踪性能。在目标跟踪过程中,由于跟踪目标、背景、光照等变量的存在,常出现光照变化、目标旋转、目标尺寸巨变、背景相似物干扰、遮

挡等情况,对目标的有效、准确跟踪造成了挑战。SiamNL^[18]、SiamAttn^[19]和TransT^[20]等分别在孪生网络的基础上引入了注意力机制,在提升特征提取效率的同时还能够提升跟踪算法的精确率。以上算法的改进对跟踪算法的精确率有一定的提高,但同时也会增加额外的运算消耗。

针对上述目标跟踪算法存在的问题,本文提出了一种更适合对无人机等小目标进行跟踪的算法——SiamAU,在改进的Resnet-50^[21]主干网络基础上结合了具有良好跨通道信息获取能力的ECA-Net^[22]注意力机制,同时使用效果更好的激活函数对主干网络进行改进。改进的HE-ResNet-50主干网络可以提升复杂场景下的特征提取能力,针对反无人机工作中以低空飞行的无人机等小目标为跟踪目标的情况对算法进行调整,对网络中的浅层特征进行浅层降维操作,同时与深度特征通道数重新排列再融合。改进后的新融合特征可以解决跟踪算法对无人机等像素占比较小、目标跟踪效果不佳的问题,分别在无人机视角数据集UAV123^[23]和反无人机数据集DUT Anti-UAV^[24]平台上进行了测试。实验结果表明,提出的SiamAU算法和加入注意力机制的基准算法在无人机视角跟踪及反无人机场景中的跟踪性能均优于目前的基准算法。

2 反无人机目标跟踪算法SiamAU

2.1 反无人机目标跟踪的难点分析

对于无人机目标的有效跟踪是反无人机过程成功的先决条件。针对无人机的目标跟踪之所以特殊是因为无人机目标独特的飞行特点、外观以及工作环境:在视频当中目标占比小、飞行移动速度快、飞行高度较低以及飞行环境相对复杂等。

在实际的反无人机任务中,通用的目标跟踪算法常会遇到跟错、跟丢等情况。这是由于无人机等小目标在飞行过程中像素占比较少,在经过深度网络特征提取后,不能有效地利用浅层信

息,过度关注深层信息可能会导致跟踪效果下降。对于SiamRPN++等使用了深度网络进行特征提取的算法,无人机等小目标的浅层特征相比深层特征具有更好的跟踪效果,因此合理利用浅层特征会对跟踪无人机目标产生良好的效果。

此外,无人机目标的飞行场景常包含树木、楼房等低空复杂场景,在跟踪算法当中加入合适的注意力机制对特征提取网络进行改进有助于降低复杂低空场景下的跟踪失败概率。

2.2 总体框架

针对反无人机场景,提出了一种引入注意力机制和特征重排机制的目标跟踪算法——SiamAU。以SiamRPN++作为基准算法,在ResNet-50主干网络中加入ECA-Net模块,同时使用效果更好的激活函数,得到改进的HE-ResNet-50,

可以解决复杂背景下低效特征提取导致的跟踪表现不佳问题。又根据反无人机工作中无人机目标的飞行特点,对浅层特征进行浅层降维并与深层特征融合,新的融合特征可以对无人机等小目标有更好的跟踪效果。整体算法框图如图1所示,以去除池化层、加入注意力机制、改进激活函数的HE-ResNet-50作为主干网络,两条支路的设计结构和参数完全相同,分别对模板帧和检测帧进行特征提取,得到两组特征图 $\{conv_1, conv_2, conv_3, conv_4, conv_5\}$ 。在特征提取阶段,选取后3层深度特征 $conv_3, conv_4$ 和 $conv_5$,经过ECA-Net注意力机制增强后得到新的深度特征,再进行特征重排级联融合,实现对图像的特征提取。最后,通过RPN、边界回归、背景分类等处理输出跟踪结果。

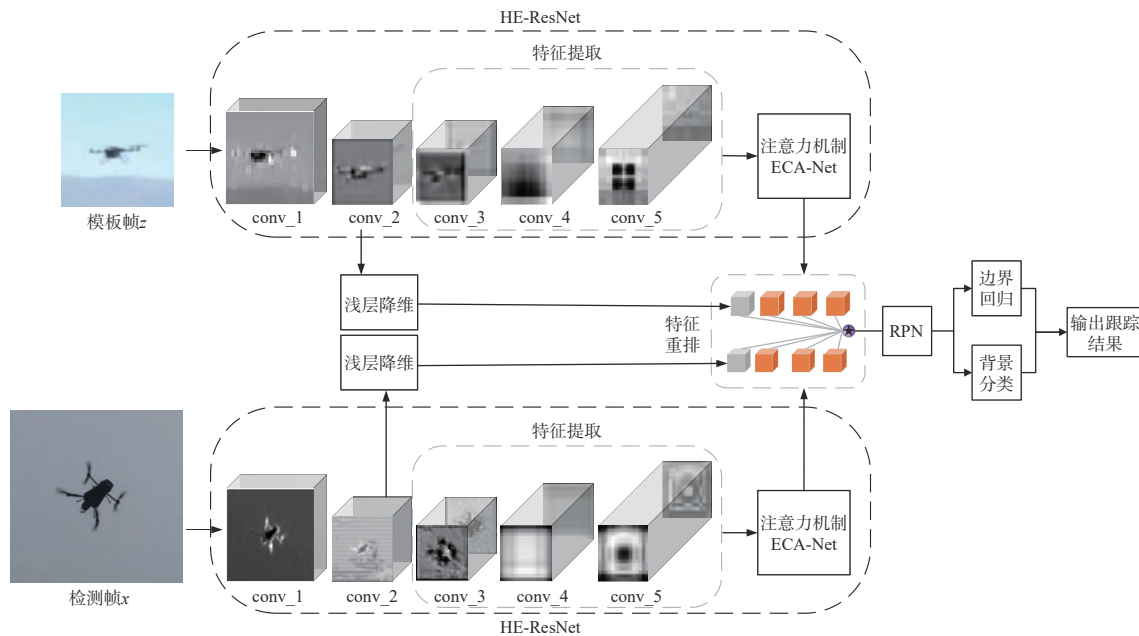


图1 SiamAU总体框架

Fig. 1 Overall structure of SiamAU

2.3 对主干网络的改进

对主干网络的改进包含两个部分:在ResNet-50各层加入ECA-Net注意力机制;对内部激活函数进行改进。最终得到改进增强的主干网络,称为HE-ResNet50。

2.3.1 加入ECA-Net注意力机制

为了提升在复杂背景下的跟踪效果,在SiamAU算法中引入了ECA-Net注意力机制,通过提出一种不降维的局部跨通道的交互策略,有效避

免了维度降低对通道注意力学习效果的影响。

ECA-Net结构框架如图2(a)所示。虽然只涉及少数几个参数,但效果提升明显。相比通道注意力机制由于捕获所有通道的依赖关系带来的效率低下的问题,ECA-Net去除了原模块中的全连接层,在经过全局平均池化的特征上通过一维卷积进行学习。ECA-Net可以通过快速一维卷积有效实现,其中 k 是自适应选择核大小,代表本地跨通道交互的覆盖范围,卷积交互作用的覆

盖范围 k 与通道维度成正比^[22]。如图 2(b)所示,在 ResNet-50 的后 3 层深度特征 conv_3、conv_4 和 conv_5 输出分别加入 ECA-Net 注意力机制,经过输出得到 3 个新的深度特征 E_conv_3、E_conv_4 和 E_conv_5,新特征与原本特征相比,分辨率和特征通道数保持不变。

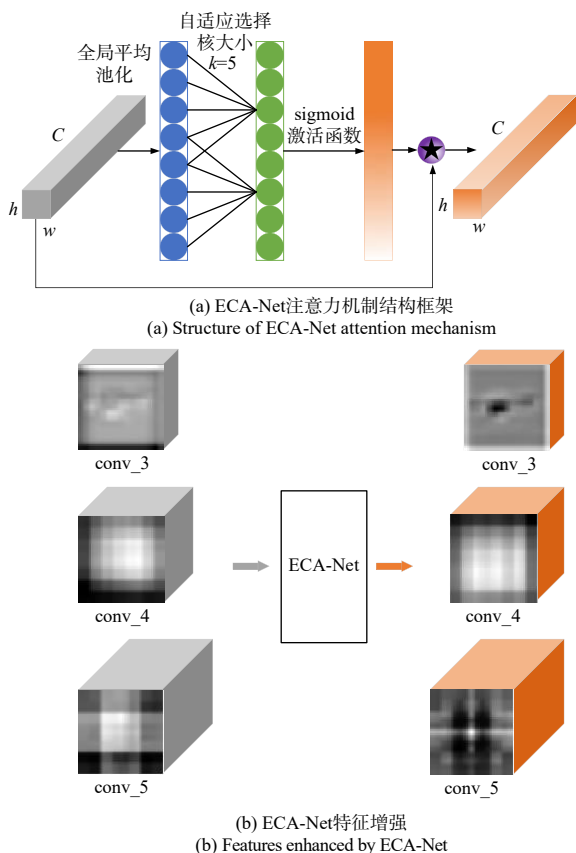


图 2 ECA-Net 注意力机制

Fig. 2 ECA-Net attention mechanism

分别选取原 ResNet-50 和结合注意力机制的 ResNet-50+ECA-Net 在无人机飞行场景下提取目标特征,特征热图如图 3 所示,最左侧代表可见

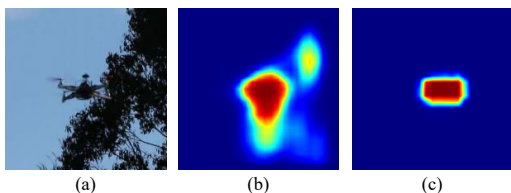


图 3 特征热图。(a)原图;(b)原特征热图;(c)增强特征热图。

Fig. 3 Feature heatmap. (a) Original image; (b) Original feature heatmap; (c) Enhanced feature heatmap.

光图像,中间代表原 ResNet-50 提取目标特征的效果,右侧代表结合注意力机制后提取目标特征的效果。可以看出,加入注意力机制后,特征提取时的高响应部分更集中于目标附近,降低了飞行场景下树木干扰的影响。

2.3.2 改进激活函数

原主干网络中应用的 ReLU 激活函数由于其优秀的特性得到了广泛应用。虽然已经提出了若干新的激活函数作为替代品,但由于增益效果不一,ReLU 激活函数始终无法被完全取代。通过基于强化学习的搜索发现了多种新颖的激活函数,在许多具有挑战性的数据集中,Swish 在更深的模型上往往比 ReLU 更好。然而 Swish 只有在深层网络中才能发挥作用,且计算量相对复杂。在此基础上提出的 HardSwish 激活函数首先消除了潜在数值精度损失。其次,HardSwish 激活函数可以实现分段功能,减少了内存访问次数,从而大幅降低了等待时间成本。3 种激活函数的表达式见式(1)、(2)、(3),函数图像如图 4 所示。

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

$$\text{Swish}(x) = \frac{x}{1 + e^{-x}}, \quad (2)$$

$$\text{HardSwish}(x) = x \times \frac{\text{ReLU6}(x+3)}{6}. \quad (3)$$

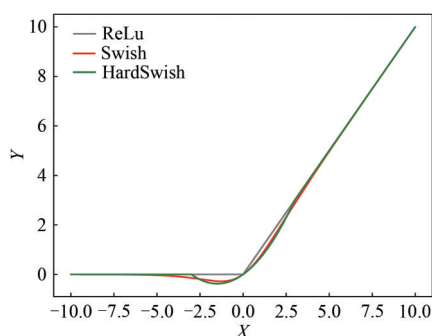


图 4 3 种激活函数

Fig. 4 Three activation functions

与原主干网络相比,结合了 HardSwish 的 HE-ResNet-50 主干网络更适用于反无人机跟踪中常见的目标遮挡、小目标跟踪等需要丰富语义特征的跟踪场景,可以在不增加网络复杂度的同时更有效地提取特征。ResNet-50 与 HardSwish 激活函数的结合方式如图 5 所示。

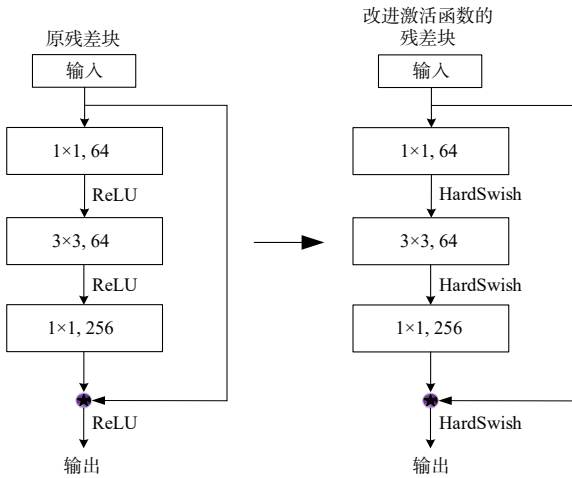


图5 ResNet-50 与 HardSwish 结合示意图(以 conv_2 层为例)

Fig. 5 Integration of ResNet-50 and HardSwish (conv_2 as an example)

2.4 特征重排机制

特征表示在孪生网络目标跟踪中起到关键的

作用。主干卷积网络提取的特征图会极大地影响跟踪器的性能。具体而言,浅层特征主要集中在颜色、形状等定位所必需的底层信息上,相比于深层特征缺乏必要的语义信息;深层特征则具有更加丰富的语义信息,可以在运动模糊、目标形变等更具有挑战的场景中发挥作用。在反无人机目标跟踪中,由于无人机目标具有像素占比较小的特点,在跟踪过程中应该更加关注有价值的浅层信息,并合理地聚合各层次的深度特征才可以更好地提高目标识别和定位的精确率。

为了提升对小目标的跟踪效果,提出了特征重排机制(FR)。SiamAU中使用HE-ResNet-50作为主干网络用于特征提取。由表1可知,以输入大小是 $127 \times 127 \times 3$ 为例,经过HE-ResNet-50的E_conv_2是典型的浅层特征。经过浅层降维模块,使用步距为2的浅层降维层,维度可由原来的 $31 \times 31 \times 256$ 变为 $15 \times 15 \times 1\ 024$,其过程如图6所示。再经过 1×1 卷积可继续得到 $15 \times 15 \times 512$ 的Ep_conv_2。

表 1 各个卷积层输出的特征图的大小

Tab. 1 Size of the output feature maps for each convolutional layer

输入	输出					
	conv_1	MaxPool	conv_2	conv_3	conv_4	conv_5
$127 \times 127 \times 3$	$61 \times 61 \times 64$	$31 \times 31 \times 64$	$31 \times 31 \times 256$	$15 \times 15 \times 512$	$15 \times 15 \times 1\ 024$	$15 \times 15 \times 2\ 048$

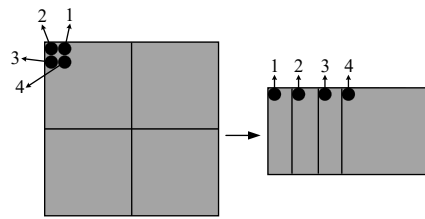


图6 浅层降维模块

Fig. 6 Low-level dimension reduction module

对最后3个残差块中提取的多层次特征进行分层聚合实现对目标的位置确定。如图7所示,由于经过ECA-Net注意力机制的Ep_conv_2、E_conv_3、E_conv_4和E_conv_5是4个空间分辨率相同但通道数不同的由浅入深的特征,通道数分别是512、512、1 024和2 048。对于原有基准算法SiamRPN++,后3层通道数经过调节后变为256、256和256,虽然在保证计算量的前提下能得到良好的特征提取效果,但并没有充分利用浅层特征。在所提出的SiamAU算法中,针对小

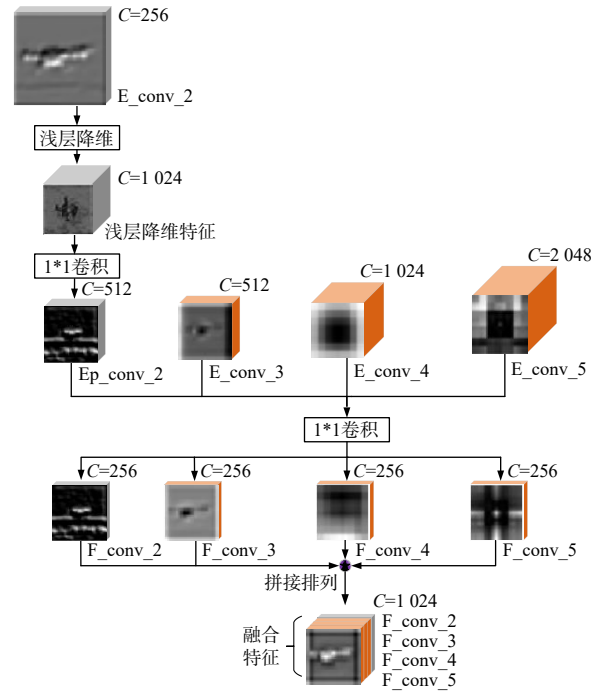


图7 特征重排原理图

Fig. 7 Diagram of feature rearrangement

目标的跟踪,利用特征重排机制将后4层特征通道数调节为256、256、256和256,再进行特征融合。

3 测量实验与结果

3.1 实验设置

为了测试所提出的SiamAU算法性能,选用近年来具有代表性的7种跟踪算法进行实验,包括KCF^[6]、SiamFC^[9]、SiamRPN^[10]、SiamFC++^[12]、SiamAPN++^[14]、SiamRPN++^[15]和Atom^[16]。SiamAU算法训练过程采用SiamRPN++作为基础跟踪器,选用经过裁剪的LaSOT^[25]、DET^[26]和YOUTUBE-BB^[27]等多个数据集进行联合训练。

实验平台硬件环境如下:CPU为Intel i7-11700,2.5 GHz;内存为12G;显卡为NVIDIA GTX3060。软件环境如下:Linux18.04.6操作系统,CUDA 11.1,Python3.6+Pytorch1.8.0。

3.2 数据集选取

选用两个公开数据集进行跟踪算法测试,分别是无人机视角目标跟踪数据集UAV123^[23]和反无人机目标跟踪数据集DUT Anti-UAV^[24]。

UAV123有123个无人机视角的视频序列,具有俯视视角下的背景复杂、干扰、目标尺度变化大等特点,广泛地应用于无人机等需要部署轻量化算法的测试与评估。在此数据集上测试,可以较全面地反映ECA-Net注意力机制对于跟踪效果的提升。

DUT Anti-UAV是专门针对反无人机目标跟踪而提出的可见光数据集,以飞行的无人机作为跟踪目标,背景主要以低空为主。数据集包含短时、长时共20个无人机飞行视频序列,合计24804帧图片,对于针对无人机目标设计的跟踪算法有良好的评估效果,对反无人机工作意义重大。DUT Anti-UAV相比于其他反无人机数据集具有场景多变、图像分辨率高的特点,同时可以很好地模拟反无人机目标跟踪任务当中常见的难点,有效地验证加入ECA-Net注意力机制以及特征重排机制后的跟踪性能提升。

3.3 评价指标

使用精确率图和成功率图对跟踪的结果进行量化评估,分别对应中心位置误差以及区域重叠率这两个常见的跟踪算法关键指标。

精确率注重跟踪算法在跟踪过程当中预测的目标中心位置与实际标注的中心位置是否接近,当两者距离小于设定的阈值代表精确,并评估所有图像中精确的比率。其中,中心位置误差是在目标跟踪过程当中算法预测的目标中心位置与实际标注的目标中心位置在视频图像上的欧氏距离:

$$e = \sqrt{(x_t - x_0)^2 + (y_t - y_0)^2}, \quad (4)$$

其中: (x_t, y_t) 表示预测目标的中心坐标, (x_0, y_0) 表示实际标注的中心坐标。通常设定阈值 $t=20$ 作为精确率的评价标准,即中心位置误差为20个像素点作为评价依据。

区域重叠率即成功率图计算公式如式(5)所示:

$$os = \frac{A_i \cap A_{gt}}{A_i \cup A_{gt}}, \quad (5)$$

其中: A_i 表示跟踪器预测的区域, A_{gt} 是目标的真实区域。区域重叠率可以定义为区域重叠度 os 大于该阈值的图像帧数占序列总帧数的比例,规定的阈值一般设定为0.5。

3.4 实验结果

3.4.1 定量实验

采用10种算法进行对比实验,包括实验设置中提到的7种对比算法和所提出的SiamAU算法及其2种消融形式(SiamRPN++ + HE, SiamRPN++ + FR)。

图8展示了10种算法在UAV123数据集上的测试结果。SiamAU的成功率和精确率分别达到了61.8%和83.2%,高于其他的孪生网络类跟踪算法以及相关滤波类跟踪算法,但相较于排名第一的Atom取得的65.1%和85.7%仍存在明显差距。这是因为UAV123是常规数据集,包含的跟踪目标种类多且目标尺度相差较大,针对常规目标跟踪设计的Atom算法相比针对反无人机任务设计的SiamAU更加有效。

图9展示了10种算法在DUT Anti-UAV数据集上的测试结果。同时加入了改进主干网络和特征重排的SiamAU算法,其成功率和精确率分别达到了60.5%和88.1%,相比排名第二的Atom分别提升了2.8%和5.1%。这是因为DUT Anti-UAV是典型的反无人机数据集,使用无人机作为跟踪目标,可以很好地模拟反无人机

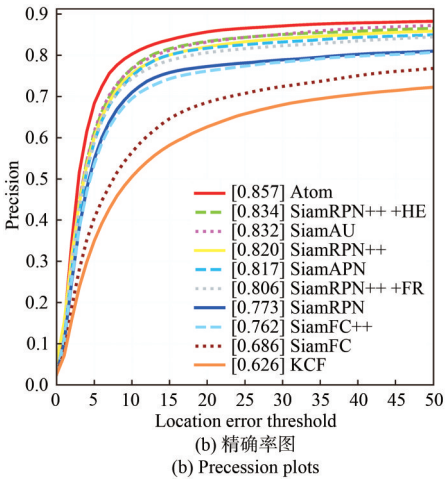
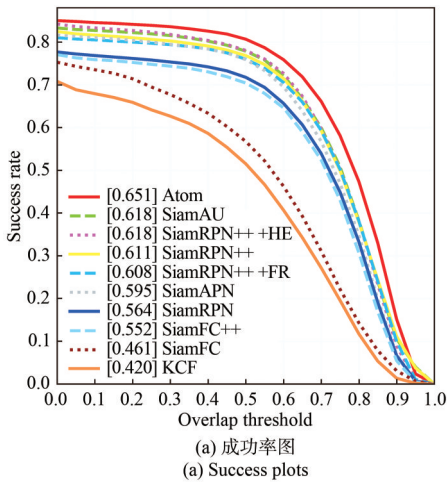


图 8 不同算法在 UAV123 上的测试结果
Fig. 8 Test results of different algorithms on UAV123

目标跟踪任务中的复杂背景以及小目标跟踪情况。SiamAU 算法在复杂场景下对无人机目标的跟踪优势相比其他对比算法更加明显。

3.4.2 消融实验结果分析

为了验证改进主干网络和特征重排机制对总体跟踪性能的贡献,选取 SiamRPN++ 作为基准算法,在 DUT Anti-UAV 数据集上开展了 4 组消融实验。SiamRPN++、SiamRPN++ + HE、SiamRPN++ + FR 和 SiamAU 在消融实验中的成功率和精确率图如图 9 所示。更细致的消融实验结果如表 2 所示。

由表 2 实验结果可知,相比基准算法,对激活函数进行改进后,算法在数据集上的测试结果中成功率和精确率分别提升了 0.1% 和 0.4%,验证了改进激活函数对跟踪算法性能提升的有效性。加入 ECA-Net 注意力机制改进后,在数据集上的测试结果中成功率和精确率分别提升了 0.8% 和

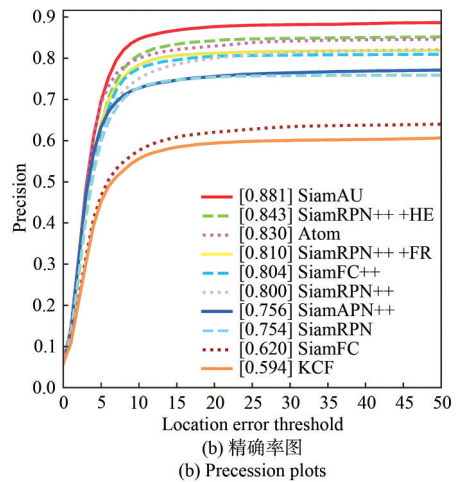
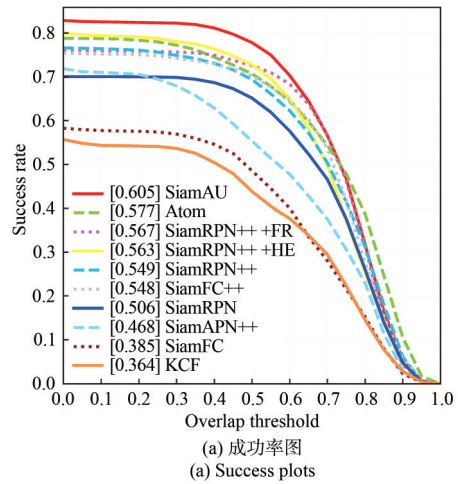


图 9 不同算法在 DUT Anti-UAV 上的测试结果
Fig. 9 Test results of different algorithms on DUT Anti-UAV

表 2 DUT Anti-UAV 消融实验结果

Tab. 2 Results of ablation experiments on DUT Anti-UAV

算法	成功率/%	精确率/%
SiamRPN++	54.9	80.0
+HardSwish	55.0	80.4
+ECA-Net	55.7	84.1
+HardSwish +ECA-Net (HE)	56.3	84.3
+FR	56.7	81.0
SiamAU	60.5	88.1

4.1%, 验证了加入 ECA-Net 注意力机制对跟踪算法性能提升的有效性。相比基准算法,单独加入改进的主干网络在成功率指标上提升了 1.4%,在精确率指标上提升了 4.3%,表明了改进的主干网络对复杂背景下跟踪的有效性。单独加入特征重排机制在成功率指标上提升了 1.8%,在精确率指标上提升了 1.0%,表明了特征重排对

于小目标跟踪的有效性。当同时加入改进的主干网络和特征重排机制后,成功率和精确率分别提升了5.6%和8.1%,在DUT Anti-UAV反无人机数据集上达到了最佳的跟踪效果。

对于UAV123数据集,SiamRPN++、SiamRPN++ +HE、SiamRPN++ +FR和SiamAU在图8中对应的曲线同样也展示了消融实验的结果。其中SiamRPN++ +HE跟踪效果最优,进一步说明了ECA-Net对于复杂背景下的有效性。SiamRPN++ +FR跟踪效果最差,这是因为UAV123数据集中包含了多种尺寸的目标,而特征重排机制仅是针对小目标跟踪的优化策略,并不能提升非小目标的跟踪效果,这也从侧面验证了特征重排机制对于反无人机小目标跟踪场景下的有效性。

3.4.3 跟踪序列对比

为了进一步定性分析所提出方案的有效性,选取DUT Anti-UAV反无人机数据集中4个不同场景下的视频序列,给出了可视化跟踪结果。这4个序列分别是v05、v09、v12和v17,包含了树木、楼房等典型的无人机低空飞行背景,并且其中的无人机目标较小。4个序列的分辨率依次为 $720 \times 1\,280$ 、 $1\,080 \times 1\,920$ 、 $1\,080 \times 1\,920$ 和 $1\,080 \times$

1 920,序列帧数分别为450、2 500、1 486和1 300。可视化结果如图10所示,由上而下分别对应4个序列的实际跟踪情况。图11给出了标注的目标运动轨迹和算法预测的目标跟踪轨迹对比。其中橙色代表实际标注的目标运动轨迹,蓝色代表SiamAU算法预测的目标跟踪轨迹。

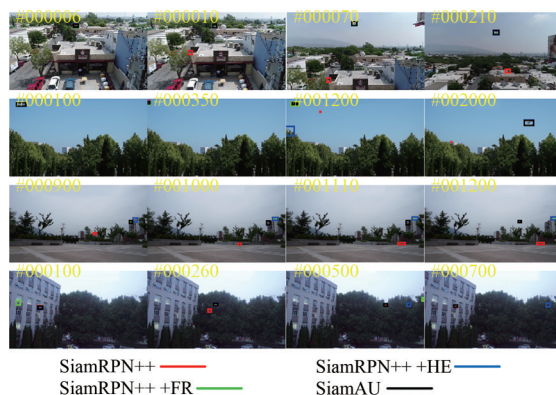


图10 DUT Anti-UAV中4个序列的可视化跟踪结果(由上至下依次为v05、v09、v12和v17)。

Fig. 10 Visualized results of four sequences on DUT Anti-UAV (from top to down: v05, v09, v12 and v17).

在v05序列中,无人机目标是自下而上的运动。在第6帧无人机从浅色楼房背景向上移动到

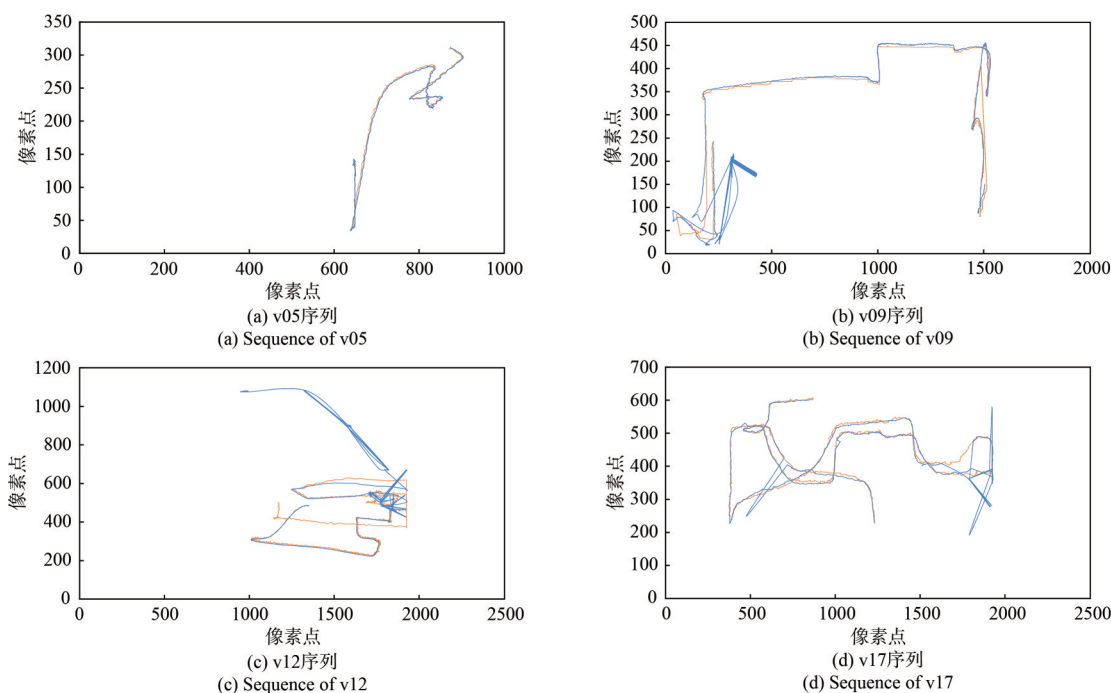


图11 DUT Anti-UAV中4个序列的跟踪轨迹

Fig. 11 Position statistics of four sequences on DUT Anti-UAV

树木背景的过程中,原 SiamRPN++ 算法跟踪出现漂移导致了跟错现象,相比之下 SiamAU 算法仍然可以在目标运动复杂背景发生变换时得到准确的跟踪结果。由图 11(a)可知,橙色线和蓝色线高度重合代表跟踪效果良好。在 v09 视频序列中的 350 帧出现了无人机目标飞出取景范围的情况,之后的 1 200 帧无人机目标又重新飞回取景范围之内。SiamAU 和 SiamRPN+++HE 算法可以在短时间对无人机目标进行重新跟踪,由图 11(b)可知,橙色线和蓝色线在左下角部分重合部分较少,是由于无人机在此过程中飞出取景范围导致的。在无人机目标重新出现之后,SiamAU 仍然可以保持良好的跟踪效果。在 v12 视频序列中,无人机目标是典型的小目标。对于小目标跟踪情况,加入了特征重排机制的 SiamRPN+++FR 和 SiamAU 算法相比其他两种算法均体现了一定的跟踪优势。由图 11(c)可知,橙色线和蓝色线在中上部分重合部分较少,可以看出出现了明显的跟丢。这是由于无人机在此过程中飞出取景范围且镜头拉近出现明显的尺度变换导致的跟踪漂移现象出现。在该序列的前半阶段,SiamAU

可以保持良好的跟踪效果。在 v17 序列中,无人机目标反复在楼房背景和树木背景间快速移动。从图 10 可以看出,SiamAU 算法取得了良好的跟踪效果,相比之下余下的 3 种算法在跟踪过程中均产生了不同程度的偏移。表 3 给出了这 4 个序列的定量跟踪结果,与图 10 展示的定性结果相一致。由图 11(d)可知,橙色线和蓝色线重合较为完全,代表跟踪器发挥了良好的跟踪效果。

由跟踪序列对比可知,SiamAU 算法对于短时无人机的目标跟踪效果提升更加明显。实验中,常规的 SiamRPN++ 目标跟踪算法表现较差,SiamRPN+++HE 算法、SiamRPN+++FR 算法表现稍好,同时加入特征重排机制和注意力机制的 SiamAU 算法具有最好的跟踪效果。在 DUT Anti-UAV 数据集上的测试结果表明,加入 ECA-Net 的注意力网络对于常见低空复杂场景的目标跟踪效果有所提升,同时也表明了特征重排机制可以在反无人机任务中对无人机等小目标具有良好的跟踪效果。虽然所提出的 SiamAU 在短时跟踪场景下效果提升明显,但对于长时跟踪(如图 11(b~c)所示的目标飞出视场)仍需做出相应的改进。

表 3 DUT Anti-UAV 中 4 个序列的跟踪结果

Tab. 3 Tracking results of four sequences on DUT Anti-UAV

序列名称	SiamRPN++		SiamRPN+++HE		SiamRPN+++FR		SiamAU	
	成功率/%	精确率/%	成功率/%	精确率/%	成功率/%	精确率/%	成功率/%	精确率/%
v05	27.6	33.7	80.1	94.7	81.6	94.7	83.7	95.2
v09	7.0	45.3	43.1	90.3	49.0	94.6	47.2	92.8
v12	34.2	57.2	35.5	57.6	41.7	67.5	43.4	70.2
v17	61.0	79.1	33.7	42.5	31.5	39.9	68.4	86.2

4 结 论

本文提出一种面向反无人机场景的目标跟踪算法 SiamAU。针对复杂背景下跟踪算法表现不佳的问题,提出了结合注意力和改进激活函数的 HE-ResNet-50 作为主干网络,有效提升了复杂背景下的特征表征能力。同时,针对无人机等小目标难以追踪的问题,对主干框架输出的浅层特征进行浅层降维并与后 3 层深层特征进行融合,更

好地利用了对小目标具有良好表征能力的浅层特征。在公开数据集 UAV123 和 DUT Anti-UAV 上进行了定量实验、消融实验和可视化分析,相比于其他主流的目标跟踪算法,所提出的 SiamAU 算法在反无人机场景下具有更加优秀的表现,在反无人机数据集上的成功率和精确率达到了 60.5% 和 88.1%,相比基准算法提升了 5.6% 和 8.1%。下一步将对反无人机场景中的长时跟踪和目标遮挡问题开展进一步的改进和完善。

参 考 文 献:

- [1] 薛珊,卫立炜,顾宸瑜,等. 采用混合域注意力机制的无人机识别方法[J]. 西安交通大学学报,2022,56(10):141-150.

- XUE S, WEI L W, GU C Y, *et al.* Drone identification method based on mixed domain attention mechanism [J]. *Journal of Xi'an Jiaotong University*, 2022, 56(10): 141-150. (in Chinese)
- [2] 陈小龙,陈唯实,饶云华,等. 飞鸟与无人机目标雷达探测与识别技术进展与展望[J]. *雷达学报*, 2020, 9(5): 803-827.
- CHEN X L, CHEN W S, RAO Y H, *et al.* Progress and prospects of radar target detection and recognition technology for flying birds and unmanned aerial vehicles [J]. *Journal of Radars*, 2020, 9(5): 803-827. (in Chinese)
- [3] 杨小伟,文清丰,杨雪,等. 基于卷积神经网络的无人机射频信号识别[J]. *无线电工程*, 2022, 52(3): 456-462.
- YANG X W, WEN Q F, YANG X, *et al.* UAV RF signal recognition based on convolutional neural network [J]. *Radio Engineering*, 2022, 52(3): 456-462. (in Chinese)
- [4] 杨名宇,王浩,王含宇. 基于深度学习与光电探测的无人机防范技术[J]. *液晶与显示*, 2021, 36(9): 1323-1330.
- YANG M Y, WANG H, WANG H Y. A counter-UAV system based on deep learning and EO detection [J]. *Chinese Journal of Liquid Crystals and Displays*, 2021, 36(9): 1323-1330. (in Chinese)
- [5] DANELLJAN M, HÄGER G, KHAN F S, *et al.* Accurate scale estimation for robust visual tracking [C]. *British Machine Vision Conference*. Nottingham: BMVA Press, 2014.
- [6] HENRIQUES J F, CASEIRO R, MARTINS P, *et al.* High-speed tracking with kernelized correlation filters [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(3): 583-596.
- [7] 苗宗成,高世严,贺泽民,等. 基于孪生网络的目标跟踪算法[J]. *液晶与显示*, 2023, 38(2): 256-266.
- MIAO Z C, GAO S Y, HE Z M, *et al.* Single-objective tracking algorithm based on Siamese networks [J]. *Chinese Journal of Liquid Crystals and Displays*, 2023, 38(2): 256-266. (in Chinese)
- [8] TAO R, GAVVES E, SMEULDERS A W M. Siamese instance search for tracking [C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016: 1420-1429.
- [9] BERTINETTO L, VALMADRE J, HENRIQUES J F, *et al.* Fully-convolutional Siamese networks for object tracking [C]. *European Conference Computer Vision*. Amsterdam: Springer, 2016: 850-865.
- [10] LI B, YAN J J, WU W, *et al.* High performance visual tracking with Siamese region proposal network [C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018: 8971-8980.
- [11] CHEN Z D, ZHONG B N, LI G R, *et al.* SiamBAN: target-aware tracking with Siamese box adaptive network [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, 45(4): 5158-5173.
- [12] XU Y D, WANG Z Y, LI Z X, *et al.* SiamFC++: towards robust and accurate visual tracking with target estimation guidelines [C]//*Proceedings of the 34th AAAI Conference on Artificial Intelligence*. New York: AAAI, 2020: 12549-12556.
- [13] GUO D Y, WANG J, CUI Y, *et al.* SiamCAR: Siamese fully convolutional classification and regression for visual tracking [C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2020: 6268-6276.
- [14] CAO Z A, FU C H, YE J J, *et al.* SiamAPN++: Siamese attentional aggregation network for real-time UAV tracking [C]. *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Prague: IEEE, 2021: 3086-3092.
- [15] LI B, WU W, WANG Q, *et al.* SiamRPN++: evolution of Siamese visual tracking with very deep networks [C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach: IEEE, 2019: 4277-4286.
- [16] DANELLJAN M, BHAT G, KHAN F S, *et al.* ATOM: accurate tracking by overlap maximization [C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach: IEEE, 2019: 4655-4664.
- [17] BHAT G, DANELLJAN M, VAN GOOL L, *et al.* Learning discriminative model prediction for tracking [C]//*Proceedings of the IEEE/CVF International Conference on Computer Vision*. Seoul: IEEE, 2019: 6181-6190.
- [18] 洪培钦,罗灵鲲,刘冰,等. 引入轻量注意力的孪生神经网络目标跟踪算法[J]. *计算机工程与应用*, 2022, 58(12): 112-121.
- HONG P Q, LUO L K, LIU B, *et al.* Siamese neural network target tracking algorithm with lightweight attention [J].

- Computer Engineering and Applications*, 2022, 58(12): 112-121. (in Chinese)
- [19] YU Y C, XIONG Y L, HUANG W L, *et al.* Deformable Siamese attention networks for visual object tracking [C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2020: 6727-6736.
- [20] MA S H, DING J H, JIA W J, *et al.* TransT: type-based multiple embedding representations for knowledge graph completion [C]. *European Conference on Machine Learning and Knowledge Discovery in Databases*. Skopje: Springer, 2017: 717-733.
- [21] HE K M, ZHANG X Y, REN S Q, *et al.* Deep residual learning for image recognition [C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016: 770-778.
- [22] WANG Q L, WU B G, ZHU P F, *et al.* ECA-Net: efficient channel attention for deep convolutional neural networks [C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2020: 11531-11539.
- [23] MUELLER M, SMITH N, GHANEM B. A benchmark and simulator for UAV tracking [C]. *14th European Conference Computer Vision*. Amsterdam: Springer, 2016: 445-461.
- [24] ZHAO J, ZHANG J S, LI D D, *et al.* Vision-based anti-UAV detection and tracking [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2022, 23(12): 25323-25334.
- [25] FAN H, LIN L T, YANG F, *et al.* LaSOT: a high-quality benchmark for large-scale single object tracking [C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach: IEEE, 2019: 5369-5378.
- [26] RUSSAKOVSKY O, DENG J, SU H, *et al.* ImageNet large scale visual recognition challenge [J]. *International Journal of Computer Vision*, 2015, 115(3): 211-252.
- [27] REAL E, SHLENS J, MAZZOCCHI S, *et al.* Youtube-BoundingBoxes: a large high-precision human-annotated data set for object detection in video [C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017: 7464-7473.

作者简介:



郑滨汐,男,硕士研究生,2020年于哈尔滨工程大学获得学士学位,主要从事目标检测、跟踪和图像处理方面的研究。E-mail:914016304@qq.com