

蒋虹西, 李智. 基于改进孪生网络的地对空红外小目标探测识别研究[J]. 智能计算机与应用, 2026, 16(4): 173-179. DOI: 10.20169/j.issn.2095-2163.25112802

基于改进孪生网络的地对空红外小目标探测识别研究

蒋虹西, 李智

(四川大学 电子信息学院, 成都 610065)

摘要: 视觉目标探测与定位技术经过数十年的发展, 主流研究仍以可见光视频为主要对象。然而在红外成像场景中, 尤其是远距无人机平台对飞行器目标的观测任务中, 由于红外热辐射成像本身缺乏丰富的表面纹理, 目标往往呈现出细节不足的问题。同时, 成像设备的动态偏移与无人机自主飞行产生的运动叠加, 会形成复杂的复合运动干扰, 从而显著提升传统追踪算法的失效风险。针对上述挑战, 本研究提出了一种高效、稳定的红外目标探测与定位框架。该框架将特征增强机制与动态搜索区域自适应策略有机结合, 创新性地优化了孪生网络在复杂红外场景下的表示与匹配能力。在权威的抗干扰无人机红外数据集上的实验验证表明, 该方法不仅具备良好的实时处理性能, 还能在多源干扰条件下保持优异的鲁棒性。

关键词: 孪生网络; 红外空中小目标; 目标探测; 搜索策略; 图像处理

中图分类号: TP391

文献标志码: A

文章编号: 2095-2163(2026)04-0173-07

Research on ground-to-air infrared small target detection and recognition based on improved Siamese Network

JIANG Hongxi, LI Zhi

(College of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China)

Abstract: After decades of development, the mainstream research on visual target detection and localization technology still mainly focuses on visible light video. However, in infrared imaging scenarios, especially in the observation tasks of aircraft targets from long-range unmanned aerial vehicle (UAV) platforms, due to the lack of rich surface textures in infrared thermal radiation imaging itself, the targets often present insufficient detail problems. At the same time, the dynamic offset of the imaging equipment and the motion generated by the autonomous flight of the UAV are superimposed, forming complex compound motion interference, which significantly increases the failure risk of traditional tracking algorithms. In response to these challenges, this study proposes an efficient and stable infrared target detection and localization framework. This framework organically combines the feature enhancement mechanism with the adaptive strategy of dynamic search regions, innovatively optimizing the representation and matching capabilities of the Siamese Network in complex infrared scenarios. Experimental verification on the authoritative anti-interference UAV infrared dataset shows that this method not only has good real-time processing performance but also maintains excellent robustness under multi-source interference conditions.

Key words: Siamese Network; infrared small aerial target; target detection; search strategy; image processing

0 引言

在复杂动态环境中, 红外探测^[1]系统需要应对多重技术挑战: 包括在远距离成像条件下从低分辨率目标中提取有效特征的能力, 以及在复杂背景干扰与局部遮挡^[2]情况下保持稳健性的能力。由于此类目标在红外热成像中往往缺乏显著的形态结构特征, 同时易被高动态背景^[3]噪声与多类干扰源淹

没, 导致传统追踪算法在特征建模与轨迹预测方面面临显著困难。

为突破传统方法的限制, 基于孪生网络^[4]的追踪范式提出了一种由首帧标注引导的相似性度量机制, 构建起动态的目标匹配框架。该类算法通过端到端的离线训练获得判别性特征表达, 不仅缓解了分类网络对未见目标泛化能力不足的问题, 还显著提升了在线追踪^[5]的实时性。在典型方法的演进

作者简介: 蒋虹西(2000—), 男, 硕士研究生, 主要研究方向: 红外目标探测。

通信作者: 李智(1975—), 男, 博士, 教授, 博士生导师, 主要研究方向: 无人系统与群体智能, 频谱压缩感知。Email: lizhi@scu.edu.cn。

收稿日期: 2025-11-28

哈尔滨工业大学主办 ◆ 专题设计与应用

过程中, SINT 首次构建了模板与候选区域间的显式匹配函数,但其全局搜索策略带来了较高的计算开销; SiamFC 引入深度特征互相关操作,通过响应图峰值实现目标定位,大幅提升了速度;而 SiamRPN 则将区域建议网络融入孪生结构,通过多尺度预测模块增强了对目标形变的适应性。

本文对红外小目标探测定位研究问题主要集中在以下几个方面:

1) 图像特征提取是否高效全面,是否能包含红外图像中的大部分信息(前景、背景、目标)。

2) 网络提供的候选区域是否能够有效利用,与模板目标进行正确匹配。

3) 模板与搜索分支产生的候选区域进行相似度匹配后输出的相似度分数是否可信。

针对上述问题,本文在实验部分对所提出的孪生网络进行全面分析,通过选取具有代表性的红外小目标数据集在不同的神经网络模型上进行实验,对比实验数据能够说明改进网络展现出了更为优秀的性能。

1 相关技术

1.1 全卷积孪生网络

全卷积孪生网络(Fully Convolutional Siamese Network, FC-Siamese Network)^[6]是一类基于孪生结构设计的深度学习模型,广泛应用于视觉目标跟踪任务^[7]。该方法通过构建模板特征与搜索区域特征之间的匹配机制,实现对目标位置的快速估计^[8]。其核心思想是在离线训练阶段学习稳定的相似性度量空间,使网络能够在后续帧中利用固定模板与搜索区域进行前向匹配,无需在线更新模型参数,从而保持较高的跟踪速度和稳健性。

由于采用静态模板与固定特征提取策略, FC-Siamese Network 有效避免了传统在线更新方法中易出现的模板漂移问题,即便目标在跟踪过程中受到短时遮挡,也能凭借未被污染的初始模板保持较好的判别能力和一致性。

SiamFC 架构通过冻结网络参数的非自适应机制,在保证运算速度的同时避免了模板更新带来的误差累积问题,因此被确立为本研究的基准框架。

如图 1 所示,全卷积孪生网络目标跟踪算法由模板分支和检测分支组成。模板分支主要以图像 z 作为模板输入,目标模板 z 是跟踪图像序列第一帧已给定的目标框裁剪出的图像。检测分支主要负责接收当前帧的待搜索区域 x , 是以上一帧目标位置

为中心裁出的图像。

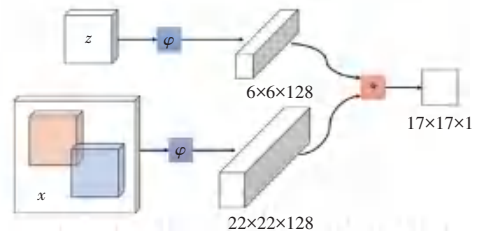


图 1 全卷积网络架构图

Fig. 1 Architecture of the Fully Convolutional Network

1.2 SwinTransformer 特征提取头

Transformer^[9]最初在自然语言处理(Natural Language Processing, NLP)领域取得突破性成功,但其直接应用于计算机视觉任务仍面临多重困难。传统 Transformer 在处理图像时通常将图像划分为多个较小的 patch,并将其展平后作为输入序列。为了获得更丰富的视觉特征,需要使用更细的 patch 划分方式,但这会导致序列长度大幅增长。当序列变长时,自注意力机制的计算复杂度呈 $O(n^2)$ 增长,因此在处理高分辨率图像时计算量会迅速膨胀,带来显著的效率问题。

此外,图像中的许多关键视觉信息依赖于局部结构,而标准 Transformer 的全局注意力机制难以有效建模局部关系,从而限制了其在视觉任务中的表现。针对这些缺陷, SwinTransformer^[10]引入了基于滑动窗口的注意力机制与分层结构,以替代传统 Transformer 的长序列处理方式。类似于 CNN 中的感受野概念, SwinTransformer 通过逐层合并 token 的方式构建层级化特征表示:第一层包含较多的 token,随后各层逐步进行两两合并,使 token 数量逐级减少(例如从 400 \rightarrow 200 \rightarrow 100),同时注意力窗口的尺寸随层数增加而扩张。这种分层特征提取机制与卷积网络中的卷积与池化过程具有相似性。

与之相对,传统的 ViT^[11]在所有层中均采用固定的 patch 尺寸和注意力结构,缺乏层级变化。如图 2 所示, SwinTransformer(左)与传统 ViT(右)在架构设计上呈现出明显差异。

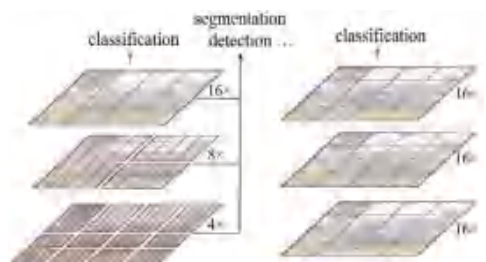


图 2 SwinTransformer 与传统 ViT 对比图

Fig. 2 Comparison diagram of SwinTransformer and traditional ViT

将 SwinTransformer 用作特征提取器具有显著优势,这主要得益于其独特的层次化结构和高效的注意力机制设计。SwinTransformer 通过局部窗口划分与逐层合并策略构建多尺度特征表示,既保留了类似 CNN 特征金字塔的结构化优势,又凭借自注意力机制在捕捉长距离依赖和全局上下文方面表现更为突出。其核心的移位窗口机制有效平衡了计算开销

与特征表达能力:局部窗口内的自注意力显著降低了计算复杂度(尤其适用于高分辨率图像),而窗口间的周期性平移则突破了固定窗口的局限,使得深层特征能够实现跨区域的信息交互,从而在保持全局感知能力的同时实现高效计算^[12]。

如图 3 所示, SwinTransformer 的网络架构体现了上述关键特性^[13]。

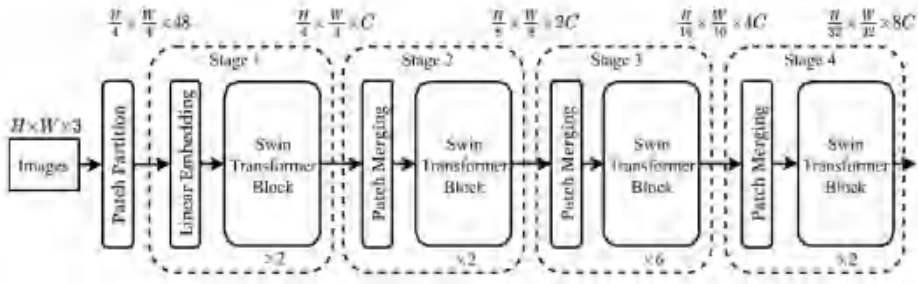


图 3 SwinTransformer 网络架构图

Fig. 3 Architecture of the SwinTransformer Network

1.3 RPN 网络

SiamRPN^[14] 作为 SiamFC 的重要扩展,在原有框架基础上融入了区域建议网络(Region Proposal Network, RPN)模块,从而构建出更高效的目标跟踪架构。与基础版本相比, SiamRPN 通过引入 RPN 避免了多尺度测试(Multi-Scale Test)的繁琐操作,显著提升了推理速度。同时, RPN 中 anchor box 的使用使得网络能够预测具有更高精度的目标边界框,大幅增强了定位能力。

此外, SiamRPN 的另一关键优势在于其能够利用稀疏标注数据进行训练。这意味着模型可以使用数量更庞大的弱标注数据集,而无需完全依赖高成本的精确标注数据。在目标跟踪任务中,训练数据规模往往直接影响模型的鲁棒性与泛化能力,因此 SiamRPN 在可扩展数据上的训练能力进一步提升了其实际表现。如图 4 所示, SiamRPN 的整体结构充分体现了上述改进设计。

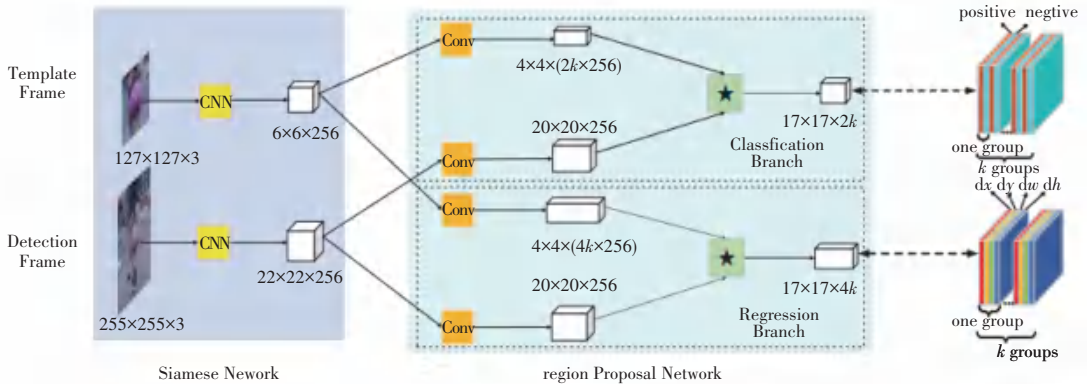


图 4 SiamRPN 网络架构图

Fig. 4 SiamRPN network architecture diagram

2 方法

本文提出一种新型的孪生网络架构 SiamGG,将 SwinTransformer 与区域建议网络相结合,并采用改进 R-CNN^[15] 网络计算模板分支与搜索分支的相似性分数^[16],将相似性得分最高的搜索框视为下一帧的目标位置。此外,还引入一种背景抑制策略,有效

降低语义背景对目标的干扰。

2.1 改进孪生网络

对于红外小目标检测与跟踪^[17],将 SwinTransformer 与区域建议网络(RPN)相结合,能够通过多层次特征融合、动态注意力机制以及高效的计算架构,显著提升复杂背景下小目标的检测精度与跟踪稳定性^[18]。这种结合不仅优化了传统卷积神经网络

(CNN)在特征提取上的局限性,还通过 Transformer 的全局建模能力增强了小目标的语义信息表达,从而在低信噪比、高背景噪声的极端场景中展现出独特优势,图 5 为改进孪生网络结构图。

SwinTransformer 的分层架构与移动窗口自注意力机制为解决小目标检测中的多尺度特征融合问题提供了新思路。传统的 CNN 在深层网络中容易丢失小目标的细节信息,而 SwinTransformer 通过分层的特征金字塔结构和跨窗口信息交互,能够在不同层级捕

捉目标的局部边缘、纹理特征及全局上下文信息。例如,在红外弱小目标检测中,SwinTransformer 通过自底向上的跨层特征融合模块(如 SFPN)将浅层的高分辨率特征与深层的语义特征结合,有效保留了微小目标的几何细节,同时抑制背景噪声干扰。这种多尺度特征提取能力为 RPN 生成高质量候选框提供了更丰富的空间与语义信息,尤其适用于像素占比极低(如红外场景中目标像素占比<1%)的弱小目标检测任务。

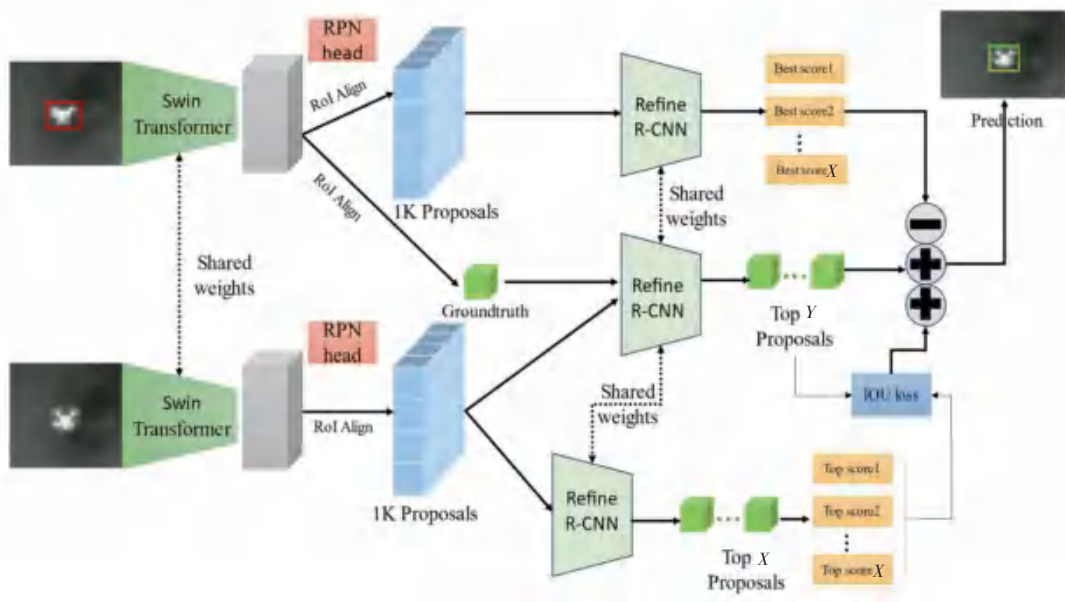


图 5 改进网络结构图

Fig. 5 Architecture of the Improved Network

2.2 背景抑制策略

在红外小目标跟踪领域^[19],传统区域建议网络(RPN)的设计理念与红外小目标所面临的严苛环境存在根本性冲突,尤其在应对复杂背景干扰方面表现出明显的结构性缺陷。其核心问题源于 RPN 最初是为可见光场景中具有丰富纹理和较大尺寸的通用目标检测而优化的,当直接应用于信噪比极低、目标仅占据几个像素的红外小目标场景时,其内在机制会放大背景干扰的影响,导致跟踪性能严重下降。

传统 RPN 高度依赖预设锚框(Anchor Boxes)来生成候选区域,这种机制在红外小目标上遭遇了维度层面的灾难性错配。为检测常见物体(如行人、车辆)设计的锚框尺寸(如 128×128 或 256×256 像素)远远超出了红外小目标(通常仅 3×3 至 9×9 像素)的实际尺度。一个微小目标会被数十个远大于自身的锚框密集覆盖,导致锚框内绝大部分像素被背景区域占据。当 RPN 对这些锚框进行“目标性”

评分时,背景像素的复杂噪声和杂波^[20](如云层边缘、地表热辐射斑块)成为主导因素。背景干扰的强响应会显著拉低真正包含小目标的锚框得分,使其难以在候选框中脱颖而出。更严重的是,即使某个锚框被标记为正样本,其内部极高的背景占比也迫使模型学习“目标+大块干扰背景”的混合模式,而非纯粹的小目标判别特征,极大地削弱了模型的区分能力。

因此,本文提出了一种动态缓解区分背景和目标的背景抑制策略^[20],本文使用以下公式计算预测结果与背景干扰物的相关性得分:

$$G(D, P) = \sum_i R - \text{CNNHead}(\Psi_{\text{corr}}(D, p_i)) \quad (1)$$

其中, $G(D, P)$ 计算预测候选 P 和背景干扰项 D 之间的相似性,使用 R-CNN 预测的最大置信度分数 $G(D, P_i)$ 作为 P_i 到 D 的相似性输出。 $G(D, P_i)$ 值越高,表明 P_i 是背景干扰者的概率越高。因此,本文利用 $G(D, P_i)$ 作为负反馈来调整最终预测

结果。集合 D 使用来自第一帧的干扰项进行初始化,并且每个背景干扰项的特征存储在序列内存中。如果本文在某个帧获得可靠的跟踪结果,则相应地更新 D ; 否则, D 保持不变。

3 实验结果与分析

3.1 实验环境与设置

3.1.1 实验环境

本文采用 Python 3.7 版本,选用 PyTorch 作为深度学习框架。所有实验均在 Ubuntu 16.04 操作系统上进行,CPU 为 intel E5-1650 v4,GPU 型号为 GTX1080ti。模型训练算法伪代码如下。

算法 1

输入:

```
Cfg_file      //配置文件
Ckp_file      //权重文件
Video_path    //数据集文件
Subset       //数据集类型
Params       //模型及更新过程中的其他参数
```

输出:

```
Tracking_results //定位结果
Number          //序号
Sequence_name   //数据集序列名称
IR Fixed Measure //每一序列平均跟踪覆盖
```

盖率

//初始化

```
 $T_{template} \leftarrow \text{Extractor}(\text{datasets})$  //初始模板
```

```
 $T_{online} \leftarrow T_{template}$  //初始化在线模板
```

```
FeatureExtractor  $\leftarrow$  SiamGG (weight shared)
```

Initialize ClassificationHead //初始化在线模板分类头//主跟踪循环

```
For each frame  $t = 1$  to  $T$  do:
```

```
 $S \leftarrow \text{ExtractSearchRegion}(T_{search}, \text{previous target})$ 
```

```
 $T + 1_{template} \leftarrow \text{FeatureExtractor}(T_{online})$ 
```

```
 $T + 1_{search} \leftarrow \text{FeatureExtractor}(S)$ 
```

//在线模板更新

```
If fixed score  $>$  th_score then:
```

```
 $T_{online} \leftarrow \text{OnlineTemplateUpdate}(T_{online},$ 
```

```
 $T_{search}, \text{target\_box})$ 
```

```
End If
```

```
Tracking Result[ $t$ ]  $\leftarrow$  target_box
```

```
End For
```

```
Return Tracking_Result
```

3.1.2 实验设置

本文模型基于 PyTorch 框架,在 AntiUAV410 数据集上进行训练与测试,评估模型在该数据集上的性能指标。分别训练 10epochs、20epochs、25epochs,每轮随机抽取 30 000 个图像。SiamGG 骨干网络的学习率设置为 0.002,采用 SGD 作为优化器,并设置 1×10^{-4} 的权重衰减。搜索区域大小设置为 256×256 像素,模板大小设置为 128×128 像素。

3.2 数据集与评价指标

AntiUAV410 数据集总计 410 个视频序列,包含超过 438 000 帧手动标注的边界框 (Bounding Boxes) 图像。每个序列平均 1 000+ 帧,时长 10~60 S,图像分辨率为 640×512 像素(热红外灰度图像),帧率 25~50 FPS。

该网络使用的评价指标主要包含两个: Precision Plot(精度)、Success Plot + AUC(成功率曲线 + AUC)。其中 Precision Plot 衡量预测位置与真实位置中心误差(CLE)在 20 像素阈值内的帧占比,Success Plot+AUC 计算成功率且 AUC 用作最终排名指标。

3.3 实验部分

3.3.1 对比实验

本实验将改进网络与 SiamFC、SiamRCNN 进行对比,对比指标为整个实验过程中目标的探测定位成功率与探测精度,效果对比数据如图 6—图 7、表 1 所示。

表 1 探测精度对比

Table 1 Comparison of detection accuracy

网络	探测精度 (Precision)
SiamFC	0.754 3
SiamRCNN	0.790 1
SiamGG	0.803 1

图 6 中黄色线为 SiamGG 数据,蓝色线为 SiamFC 数据,图 7 中蓝色线为 SiamGG 数据,黄色线为 SiamRCNN 数据。从对比图中能够得知,对于数据集中的大部分序列,改进网络均能够实现 75% 以上的探测定位效果且优于 SiamFC、SiamRCNN 网络。

3.3.2 探测定位结果定性实验

对不同的网络结构分别进行对比试验如图 8—图 9 所示。

由图 8—图 9 可知,对于运动中的红外视角下无人机,改进网络 SiamGG 在多种环境下相较于 SiamFC、SiamRCNN 网络,均能够实现更高水平的探测定位。

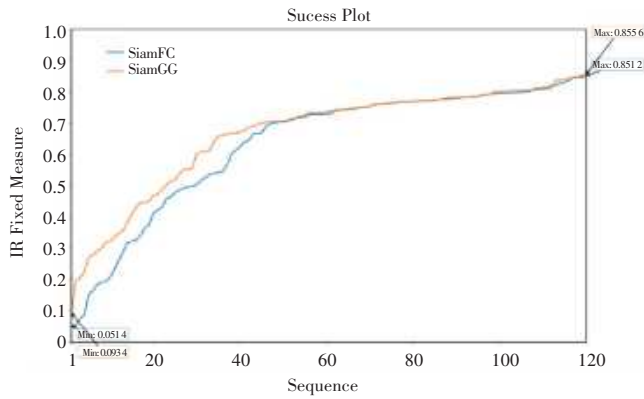


图6 改进网络 SiamGG 与 SiamFC 探测成功率对比图

Fig. 6 Detection success rate comparison between SiamGG and SiamFC

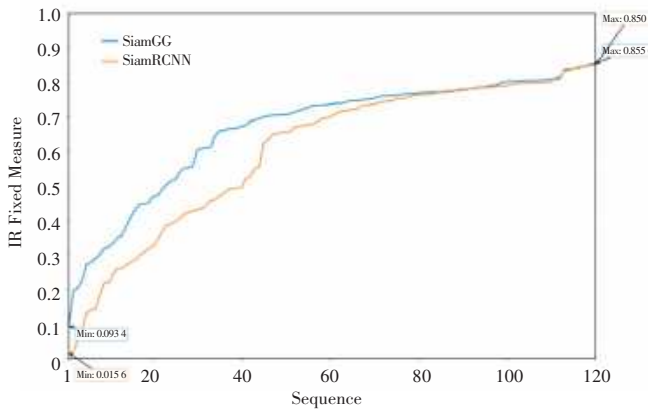


图7 改进网络 SiamGG 与 SiamRCNN 探测成功率对比图

Fig. 7 Detection success rate comparison between the improved SiamGG network and SiamRCNN



图8 改进网络 SiamGG 与 SiamFC 定位效果对比图

Fig. 8 Localization performance comparison between SiamGG and SiamFC



图9 改进网络与 SiamRCNN 定位效果对比图

Fig. 9 Localization performance comparison between SiamGG and SiamRCNN

4 结束语

本文采用 Siamese 网络进行红外小目标的探测定位研究,结合 SwinTransformer 特征提取头与 RPN 网络生成候选目标区域,通过 R-CNN 计算目标与模板的相似度分数。Siam 结构将“目标外观”以模板形式显式引入,匹配阶段天然偏向与模板外观一致的候选区域,可有效抑制海天线、云边缘、地面热斑等背景伪热点带来的误检,并且分支共享权重,连续帧以同一模板/记忆库做关联,有助于在低信噪、外观弱纹理条件下保持轨迹稳定。该两阶段范式先以 RPN 粗召回,再由相似度精筛,结构清晰、可控度高;通过小尺度锚或低层特征参与,能较好覆盖微小目标,提高候选的覆盖率。将模板-候选的相似度学习(R-CNN)放到第二阶段的 RoI 特征上,判别器能利用更细致的区域内统计与上下文信息,最终精度比一阶段直出更有优势。

同样值得注意的问题是,在 SwinTransformer 中,窗口化注意力机制是其最核心的结构设计之一,其通过将整幅图像划分为固定大小的局部窗口,只在每个窗口内部计算自注意力,从而大幅降低计算复杂度。然而,这种机制天然地限制了不同窗口之间的特征交互。每个窗口在注意力层中被作为一个独立的区域处理,注意力权重的计算和特征聚合都只发生在这个局部范围内,窗口间的信息无法直接流通。

在红外小目标检测任务中^[21],目标通常仅有几个像素,能量极弱、纹理信息极少。当一个目标恰好落在窗口边界时,其会被切分到不同窗口中,每个窗口仅看到目标的一部分或被背景信号淹没,而窗口之间的注意力无法协同感知这一碎片属于同一个目标。这意味着模型在注意力层面无法建立对小目标^[22]的整体表征,只能在局部范围内学习孤立的响应,导致红外目标的显著性被削弱,检测能力下降。

参考文献

[1] 周莉莉,姜枫. 图像分割方法综述研究[J]. 计算机应用研究, 2017,34(7):1921-1928. DOI:10.3969/j.issn.1001-3695.2017.07.001.

[2] 邓博. 基于深度学习的红外弱小目标相关滤波跟踪算法研究[D]. 四川:电子科技大学,2019.

[3] 范明明. 复杂背景条件下的红外小目标检测算法研究[D]. 陕西:西安电子科技大学,2021.

[4] 宋国鹏. 基于孪生网络的单目标跟踪算法研究及其应用[D]. 北京:北京交通大学,2020.

[5] 许美琪. 复杂背景下红外弱小目标检测方法研究[D]. 北京:中

国科学院大学,2020.

[6] BERTINETTO L, VALMADRE J, HENRIQUES F J, et al. Fully-convolutional siamese networks for object tracking [C]// Proceedings of the 14th European Conference on Computer Vision. Cham:Springer, 2016:850-865.

[7] WU Han, LI Weiqiang, LI Wanqi, et al. A real-time robust approach for tracking UAVs in infrared videos [C]//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Piscataway, NJ:IEEE,2020:4448-4455.

[8] ZOU Zhengxia, CHEN Keyan, SHI Zhenwei, et al. Object detection in 20 years: A survey [J]. Proceedings of the IEEE, 2023,111(3):257-276. DOI:10.1109/JPROC.2023.3238524.

[9] WANG Ning, ZHOU Wengang, WANG Jie, et al. Transformer meets tracker: Exploiting temporal context for robust visual tracking [C]// Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ:IEEE,2021:1571-1580.

[10] LIU Z, LIN Y, CAO Y, et al. SwinTransformer: Hierarchical vision transformer using shifted windows [C]//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Piscataway, NJ:IEEE,2021:9992-10002. DOI:10.1109/ICCV48922.2021.00986.

[11] 孟晓燕,段建民. 基于相关滤波的目标跟踪算法研究综述[J]. 北京工业大学学报,2020,46(12):1393-1416. DOI:10.11936/bjtxb2019030011.

[12] 汤一明,刘玉菲,黄鸿. 视觉单目标跟踪算法综述[J]. 测控技术,2020,39(8):14. DOI:10.19708/j.ckjs.2020.08.003.

[13] SHAH R, SÄCKINGER E, BENTZ J W, et al. Signature verification using a “Siamese” time delay neural network [J]. International Journal of Pattern Recognition and Artificial Intelligence,1993,7(4):669. DOI:10.1142/S0218001493000339.

[14] LI Bo, YAN Junjie, WU Wei, et al. High performance visual tracking with siamese region proposal network [C]// Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ:IEEE,2018:8971-8980.

[15] VOIGTLAENDER P, LUITEN J, TORR P H S, et al. Siam R-CNN: Visual tracking by re-detection [C]// Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ:IEEE,2020:6577-6587.

[16] REN Shaoqing, HE Kaiming, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [C]// Proceedings of Advances in Neural Information Processing Systems. NeurIPS,2015:91-99.

[17] CHEN Fei, ZHANG Fuhuan, WANG Xiaodong. Two stages for visual object tracking [C]// Proceedings of 2021 International Conference on Intelligent Computing, Automation and Applications. Piscataway, NJ:IEEE,2021:165-170.

[18] 陈琳,刘允刚. 面向无人机的视觉目标跟踪算法:综述与展望[J]. 信息与控制,2022,51(1):23-40. DOI:10.13976/j.cnki.xk.2022.1144.

[19] 王立功. 基于孪生网络的目标跟踪算法研究[D]. 陕西:西安电子科技大学,2021.

[20] 汪大宝. 复杂背景下的红外弱小目标检测与跟踪技术研究[D]. 陕西:西安电子科技大学,2010.

[21] 李博豪,李智. 基于改进 YOLOv8n 的无人机航拍小目标检测算法[J]. 智能计算机与应用,2025,15(5):173-179. DOI:10.20169/j.issn.2095-2163.24120902.

[22] 羊森海,陈丹. 基于改进 YOLOv7 的抓取图像小目标检测算法[J]. 智能计算机与应用,2025,15(7):99-103. DOI:10.20169/j.issn.2095-2163.250714.