

doi: 10.7690/bgzd.2025.08.014

基于卡尔曼滤波的杂波与噪声背景下红外小目标跟踪方法研究

高兴媛¹, 和铁行²

(1. 浙江长征职业技术学院, 杭州 310023; 2. 杭州医学院, 杭州 311300)

摘要: 为改进传统目标检测跟踪方法对特征的高层语义信息提取不够丰富, 导致目标跟踪结果误差大的问题, 提出基于卡尔曼滤波的复杂背景红外小目标跟踪方法。对获取到的弱小目标红外图像进行预处理, 增强其目标的对比度, 抑制背景对其的影响; 通过 YOLOv4 网络提取红外弱小目标特征, 将同一视觉特征的像素划入一个集合, 搜索集合内可疑目标, 在可疑目标中分割真实目标, 完成红外弱小目标检测; 采用卡尔曼滤波算法对红外弱小目标的运动轨迹进行追踪, 并引入合适的损失函数以提高目标跟踪能力。研究结果表明: 该方法的中心误差始终小于其他对比方法, 最小达到 0.53; 重叠率始终高于其他对比方法, 最高达到 1.0, 该方法优于对比方法, 能提高红外图像小目标的跟踪效果, 应用性能佳。

关键词: DeepSORT; 红外小目标; 红外图像; YOLOv4 网络; 跟踪算法

中图分类号: TP391.41 **文献标志码:** A

Research on Tracking Method of Infrared Small Target in Clutter and Noise Background Based on Kalman Filter

Gao Xingyuan¹, He Tiexing²(1. Zhejiang Changzheng Vocational and Technical College, Hangzhou 310023, China;
2. Hangzhou Medical College, Hangzhou 311300, China)

Abstract: In order to improve the problem that the traditional target detection and tracking method is not rich enough in extracting high-level semantic information of features, which leads to large error in target tracking results, a small infrared target tracking method based on Kalman filter in complex background is proposed. The method comprises the following steps of preprocessing an acquired infrared image of a dim target, enhancing the contrast of the target, and inhibiting the influence of a background on the target; extracting the characteristics of the infrared dim target through a YOLOv4 network, dividing pixels with the same visual characteristics into a set, searching a suspicious target in the set, segmenting a real target in the suspicious target, and detecting the infrared dim target; The Kalman filter algorithm is used to track the trajectory of infrared dim and small targets, and a suitable loss function is introduced to improve the target tracking ability. The results show that the center error of this method is always smaller than other contrast methods, the minimum is 0.53, and the overlap rate is always higher than other contrast methods, the maximum is 1.0. This method is superior to the comparison method, and improves the tracking effect of small target in infrared image, and has better application performance.

Keywords: DeepSORT; infrared dim target; infrared image; YOLOv4 network; tracking algorithm

0 引言

红外成像技术是利用红外探测器接受物体自身发射的红外热辐射来进行工作, 具有作用距离远、易于隐藏、可昼夜工作等特点。随着技术的进一步发展, 红外成像系统已被广泛应用于红外制导、红外报警及红外视频监控等领域。其应用领域宽泛, 并在各个领域发挥重要作用。红外图像的目标跟踪检测是目前较为重要的一项技术, 对于红外精准制导和预警非常重要。很多学者都在积极研究该项技术, 并已经取得一定研究成果。如王翔^[1]针对复杂海空背景难以进行红外小目标追踪检测的问题, 提出了一种新型红外小目标抗干扰检测跟踪算法。

采用不同分类器区分不同区域的杂波和小目标, 提升目标跟踪精度, 但其检测效率较低。徐亚杰等^[2]针对复杂杂波背景下红外弱小目标检测困难的现状提出了新特征算子的红外目标检测方法, 但其检测正确率和检测效率并不高。戴亚峰等^[3]针对红外视频目标进行检测和跟踪。采用 YOLOv3 网络模型完成目标跟踪的初始化目标选取, 然后采用 KCF 算法对后续视频帧进行目标跟踪及检测。存在跟踪检测精度不高的问题。

基于此, 笔者提出基于卡尔曼滤波的红外小目标跟踪检测方法, 希望通过本文中研究, 能够克服复杂背景, 实现红外小目标的精准检测。

收稿日期: 2024-09-10; 修回日期: 2024-10-17

基金项目: 浙江省教育厅项目(jg20240369)

第一作者: 高兴媛(1980—), 女, 四川人, 硕士。

1 红外图像弱小目标检测跟踪

1.1 红外弱小目标图像增强

红外探测器在进行目标数据提取时,会受到大气热辐射、较远作用距离及外界噪声等影响,导致其检测的信号偏弱,即使在正常天气下也容易受到外界干扰影响,导致图像模糊的状况。尤其是非平稳起伏的背景干扰下,待检测目标被大量复杂的背景杂波和噪声所淹没,所以其采集的目标红外图像存在对比度低、亮度低或信噪比低等情况,导致其小目标的形成。所谓负载背景就是红外目标所处的背景高灰度值成分较多,且灰度成分剧烈变化,其检测难度较大。针对复杂背景下弱小目标的红外图像,应先增强目标,抑制背景。

笔者先设定不同尺度下的分解级数,分解小目标的红外图像。采用拉普拉斯模型,描述分解后红外图像的剪切波系数分布,将大部分目标能量作为高频子带,将少部分目标能量和图像背景作为低频子带^[4]。其中拉普拉斯数学模型为:

$$Q=a+b. \quad (1)$$

式中: Q 为低频带中含有残余目标的图像信息; a 为低频带中少量目标信息; b 为低频带中大量背景信息^[5-6]。将少量目标信息看作低频段背景中的噪音,计算噪音信息中剪切波系数的概率密度 w :

$$w=\frac{1}{\sqrt{2\pi}c}\times\exp(-t^2/2c^2). \quad (2)$$

式中: c 为目标信息位置参数; t 为剪切波系数^[6]。利用式(1),对低频子带内剪切波系数进行建模,可得模型为:

$$\sigma_1=(d_1+d_2)Qw. \quad (3)$$

式中: σ_1 为低频段标准差; d_1 、 d_2 分别为低频子带内的背景剪切波系数和目标剪切波系数^[7]。

利用贝叶斯统计方法,在低频子带模型中,通过恢复大部分背景信息,对目标信息进行消除,使目标信息标准差 σ 达到设定阈值,实现红外图像低频子带的去噪^[8]。然后通过剪切波系数数量和幅值的差异,估计红外图像高频部分背景,当剪切方向的系数能量,占尺度能量的比重较大时,则判断红外高频子带包含较多细节信息,通过抑制剪切方向系数的方法,抑制细节成分中的噪声^[9]。其中剪切波系数阈值表达式为:

$$\sigma_2=|\xi h|-f/E. \quad (4)$$

式中: σ_2 为高频段标准差; f 为剪切方向的子带能量; ξ 为高频段尺度因子; h 为尺度 ξ 的剪切波方向

数; E 为剪切波将原图细分的层数,其中 E 与 σ_2 呈反比例相关^[10]。对可控制变量 E 进行控制,降低高频段标准差,从而降低剪切波系数的对应能量,实现红外图像高频子带的去噪。

利用逆变换重构的方法,处理去噪后的高低频子带,得到红外图像的估计背景,再利用原图法,使原始红外图像与背景图像相减,得到背景抑制后的图像,以此增强红外图像小目标^[11]。

1.2 红外图像目标检测

输入预处理后的红外图像,以 YOLOv4 网络为基础,检测红外图像小目标。笔者采用的 YOLOv4 网络结构大致可分为输入端、Backbone、Neck 和 Prediction 4 部分组成,如图 1 所示。

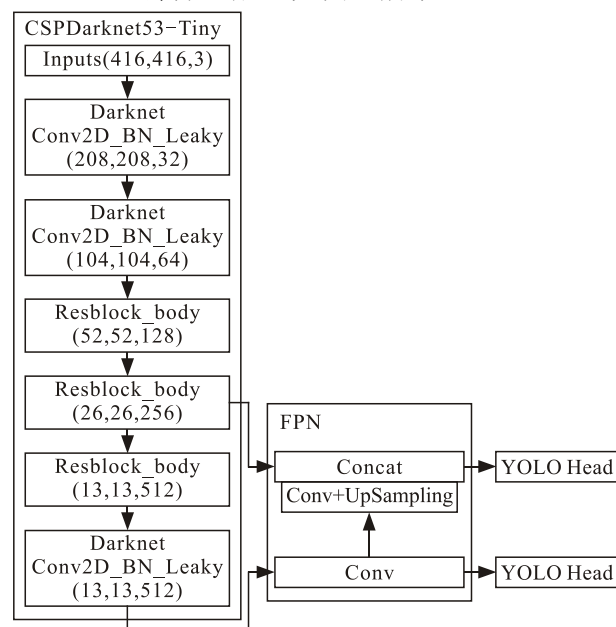


图 1 YOLOv4 网络结构

其中输入端可输入经过图像增强处理后的红外图像,Backbone 端通过插入图片特征的集合进行下采样,以提取更高的语义信息;Neck 部分主要可分为 SPP 和 PANet 2 个部分,SPP 结构利用不同大小池化核的最大池化对输入特征层进行池化后堆叠,增加主干特征的接收范围,而 PANet 则是主要进行下采样,仅以加深特征提取的性能;Prediction 主要将特征转化为预测结果,输出红外图像小目标的特征图。根据提取的特征信息,采用包含尺度因子的点扩散函数形式,对红外图像小目标进行描述,描述公式 T 为:

$$T=j\exp\left[-((x_1-x_2)^2+(y_1-y_2)^2)/2k\right]/g. \quad (5)$$

式中: g 为红外图像小目标大小直径; j 为目标峰值强度; k 为红外图像尺度因子; (x_1, y_1) 为红外图像检

测位置； (x_2, y_2) 为目标中心位置^[12]。

比较待处理像素的 T 值,对可疑目标进行分割,首先把红外图像小目标,看作红外图像中的细节成分,利用自适应邻域概念,对红外图像进行滤波处理,将每一个待处理像素作为种子像素,检测种子像素的邻接邻域像素^[13]。针对集合内的种子像素,规划集合区域,构成种子像素的拉普拉斯尺度空间,进行滤波响应,具体如图 2 所示。

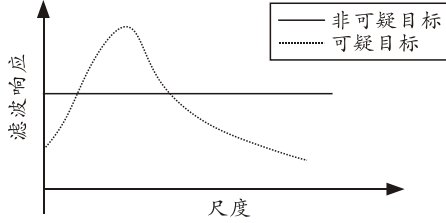


图 2 红外图像小目标的滤波检测

如图 2 所示,在拉普拉斯空间内,非可疑目标的滤波响应保持输出不变,而可疑目标的滤波响应,随空间尺寸的变大而变大,达到峰值后,随尺寸变大而变小。接下来对图像进行分割处理,获得红外小目标^[14]。根据目标大小划定背景窗口,使窗口能根据目标尺寸的变化而变化,计算背景窗口的标准差 M 为:

$$M = \frac{1}{n} \sum_{i=1}^n m_i. \quad (6)$$

式中: n 为背景区域中的像素个数; i 为背景区域像素; m_i 为像素 i 的像素值。则窗口的自适应阈值 B 为:

$$B = M + j \times \rho. \quad (7)$$

式中: j 为不同场景下的阈值系数^[15]; ρ 为背景自适应标准差。利用窗口的自适应阈值,抑制并剔除可疑目标中的虚假目标,准确分割出真实的红外图像小目标。

综上所述,式(5)构建了目标特征模型,可以根据目标的直径、峰值强度、尺度因子和检测位置等信息,对红外图像中的小目标进行初步描述。式(6)通过对背景窗口内像素值的统计分析,揭示了背景噪声的分布特性,为目标的准确分割提供了背景信息的支持。式(7)结合背景自适应标准差和阈值系数,计算出能够自适应目标尺寸和背景噪声变化的阈值,从而实现了可疑目标的有效筛选和真实目标的准确分割。

1.3 基于卡尔曼滤波算法的目标轨迹跟踪

利用 DeepSORT 算法中的卡尔曼滤波算法对目标的轨迹问题进行预测。设红外小目标的轨迹建模如下:

$$\mathbf{x} = [u, v, r, h, x, y, r, h]. \quad (8)$$

式中: \mathbf{x} 为目标的状态向量; u 为目标边界框中心的水平像素位置; v 为目标边界框中心的垂直像素位置; r 为目标边界框的长短比; h 为目标边界框的高度; x, y, r, h 为目标在图像坐标中的速度信息^[16]。在卡尔曼滤波算法中对红外图像进行检测非常重要的就是是否可检测到目标:如已经检测到目标且其与目标相关联时,检测到的边界框可被用于更新目标状态;如果没有检测到目标,则需要采用线性速度模型对目标的状态进行简单预测,而不会进行其余校正操作。

该算法运行共分为 2 个部分:

第 1 部分采用时间更新方程去对红外小目标进行描述,推测红外目标下一帧的状态,其时间更新方程如下:

$$\mathbf{X}_{(k|k-1)} = \mathbf{A}\mathbf{X}_{(k-1|k-1)} + \mathbf{B}\mathbf{U}_{(k)}; \quad (9)$$

$$\mathbf{P}_{(k|k-1)} = \mathbf{A}\mathbf{P}_{(k-1|k-1)}\mathbf{A}^T + \mathbf{Q}. \quad (10)$$

式中: $\mathbf{X}_{(k|k-1)}$ 为 $k-1$ 时预测到 k 时的状态向量; $\mathbf{X}_{(k-1|k-1)}$ 为 $k-1$ 时的最优状态向量; $\mathbf{P}_{(k|k-1)}$ 、 $\mathbf{P}_{(k-1|k-1)}$ 分别为 $\mathbf{X}_{(k|k-1)}$ 和 $\mathbf{X}_{(k-1|k-1)}$ 的协方差矩阵; $\mathbf{U}_{(k)}$ 为 k 时刻的真实状态向量; \mathbf{A} 、 \mathbf{B} 为状态更新矩阵; \mathbf{Q} 为噪声的协方差矩阵。

第 2 部分为校正,将上一步得到的预测轨迹和检测位置作差,求得误差协方差矩阵等参数,通过不断迭代,完成红外小目标追踪及检测^[17],其状态更新方程如下:

$$\mathbf{X}(k|k) = (1 - \mathbf{K}_k \mathbf{H}) \mathbf{P}_{(k|k-1)}. \quad (11)$$

式中: $\mathbf{X}(k|k)$ 为 k 时刻协方差的更新结果; \mathbf{H} 为预测部分高斯分布的协方差矩阵; \mathbf{K}_k 为 k 时刻的卡尔曼增益矩阵, $\mathbf{K}_k = \mathbf{P}_{(k|k-1)} \mathbf{H}^T (\mathbf{H} \mathbf{P}_{(k|k-1)} \mathbf{H}^T + \mathbf{R})^{-1}$ 。

为优化 YOLOv4 网络对红外小目标外观感知的能力,提高目标可分性与目标类内紧度和类间差异,对 DeepSORT 算法的深度外观描述部分进行改进。利用深度度量学习来评估先、后 2 帧的相似程度,通过引入损失函数 Margin Loss,以使 DeepSORT 算法更加符合红外小目标跟踪策略,提高跟踪器对相同红外小目标的辨别能力。

Margin Loss 的计算公式如下:

$$L = -\frac{1}{N} \sum_{i=1}^N \lg \frac{e^{d(\cos(\theta_{y_i} + p))}}{e^{d(\cos(\theta_{y_i} + p))} + \sum_{j=1, j \neq y_i}^q e^{s \cos \theta_j}}. \quad (12)$$

式中: N 为批次的大小; d 为目标特征缩放的特征尺度; θ_{y_i} 为特征层权重与目标特征之间的夹角; p

为附加角惩罚偏置项； q 为类别数量。Margin Loss 利用目标特征与特征层权重之间的角度的分析统计来构造损失函数，能够在不增加多余计算开销的前提下提高 DeepSORT 模型的目标鉴别能力以及训练过程的稳定性。

2 实验论证分析

2.1 实验准备

笔者采用实验室实测的红外图像进行仿真实验，本次研究在配备 Intel Xeon E5-2620 2.10 GHZ CPU 的 Ubuntu 系统上展开，其显卡型号为 NVIDIA RTX2080Ti，这是一款高性能显卡，采用 Turing 架构，拥有 4 352 个 CUDA 核心和 11 GB GDDR6 显存，能够提供出色的图形处理和并行计算能力。实验所用的编程语言使用 Python3.7 和 Pytorch1.4.0。以此为基础，完成实验环境搭建。以上述为基础，完成仿真模型的搭建，选取文献[1]、[2]、[3]为对比方法，与本文中方法进行对比实验。

2.2 评价指标

中心误差指标用于评价目标跟踪算法中目标位置的准确性，较小的中心误差值表示算法能够准确地估计目标的位置，反之则表示跟踪性能较差，表达式为：

$$CE_{fi} = \|c_{val}^{fi} - c_{gt}^{fi}\|^2 \quad (13)$$

式中 c_{val}^{fi} 与 c_{gt}^{fi} 为跟踪结果和目标真实位置的中心点坐标。

窗口重叠率是用于评价目标跟踪算法中目标边界框与真实目标边界框之间的重叠程度的指标。重叠率接近于 1 表示目标跟踪较为准确，而较低的重叠率值可能意味着目标跟踪装置存在误差或无法正确跟踪目标，表达式为：

$$OR_{fi} = \text{map}_{eval}^{fi} I \text{map}_{gt}^{fi} / \text{map}_{eval}^{fi} \quad (14)$$

式中： map_{eval}^{fi} 为目标跟踪区域； map_{gt}^{fi} 为真实目标所在区域。

2.3 对比实验分析

在实验的初始阶段，从实验室实测的红外图像数据集中选取实验样本。这些样本不仅具有代表性，而且涵盖了各种场景和目标类型，确保了实验的多样性和全面性。为了确保这些数据能够顺利输入到 YOLOv4 网络中，对选取的图像进行预处理，如尺寸调整、归一化等。

在模型搭建与训练阶段，使用 Python3.7 和

Pytorch1.4.0 搭建 YOLOv4 网络模型。根据消融实验设置，逐步去除网络中的不同组成部分，构建不同的消融模型。在训练过程中，使用选取的实验样本对基准模型和消融模型进行训练，记录训练过程中的损失值和准确率等指标。

训练完成后，使用训练好的模型对测试集预测，记录预测结果。根据中心误差和窗口重叠率等指标对预测结果进行评估，比较不同方法的跟踪性能。此外，对复杂背景下的目标跟踪效果进行特别分析，记录并比较各方法在复杂背景下的跟踪精度和误差。其中，YOLOv4 网络训练的代码如下：

```
num_epochs=50; //定义训练次数
for epoch in range(num_epochs)://开始训练
    model.train()
    for images,targets in train_loader://定义训练参数
        optimizer.zero_grad(); //将模型参数的梯度
        归零
        outputs =model(images); //输出模型处理后
        图像
        Loss=criterion(outputs,targets)// 调用 loss
        函数
        loss.backward(); //反馈
        optimizer.step(); //优化步骤
        print(f'Epoch {epoch+1}/{num_epochs} ',
        Loss: {loss.item()}); //输出结果
```

2.4 对比实验分析

为验证本文中 YOLOv4 网络的有效性，进行消融实验。YOLOv4 网络结构由输入端、Backbone、Neck 和 Prediction 4 部分组成。消融实验是通过逐步去除网络中的不同组成部分以观察其对网络性能的影响。以小目标跟踪精度为实验指标，进行消融测试，消融设定如表 2 所示。

表 1 消融实验设置

实验设置	输入端	Backbone	Neck	Prediction
基准模型	✓	✓	✓	✓
消融设置 1	×	✓	✓	✓
消融设置 2	✓	×	✓	✓
消融设置 3	✓	✓	×	✓
消融设置 4	✓	✓	✓	×

根据上述消融设定标准进行小目标跟踪精度对比测试，结果如图 3 所示。

根据图 3 可以看出：基准网络的跟踪精度效果最好，然后按照消融设置 1、3、4、2 的顺序依次下降，其中消融设置 2 准确率最低，说明 Backbone 对于跟踪精度影响最大；消融设置 1 准确率最高，

说明输入端对于跟踪精度影响最小。则影响精度按照由大到小进行对比，结果为：Backbone>Prediction>Neck>输入端。

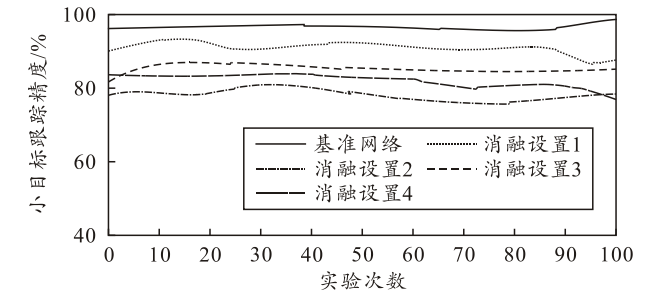


图 3 小目标跟踪精度对比结果

如图 4 所示，△为本文中方法、☆为文献[1]方法、○为文献[2]方法、□为文献[3]方法的小目标跟踪效果，以复杂背景各方法的跟踪效果进行对比分析。

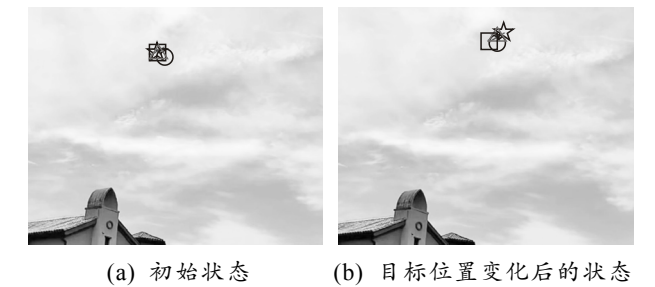


图 4 复杂背景下各方法的跟踪结果

从图 4 中可以看出：图 4(a)飞机处于缓慢移动的状态，在这种情境下，背景杂波和噪声相对稳定，

使得跟踪算法能够较为容易地区分目标和背景；图 4(b)飞机进行了较快的机动动作，导致背景杂波和噪声的相对位置也发生变化，增加了跟踪的难度。在图 4(a)的复杂背景小目标跟踪识别中，本文中方法、文献[1]方法与文献[3]方法有效识别出了目标；在图 4(b)的复杂背景小目标跟踪识别中，本文中方法仍能有效识别出小目标，说明本文中方法能够在复杂背景下具有精准的跟踪性能。

为清晰显示出跟踪精度结果，得到表 2 数据。

表 2 3 种方法的跟踪精度 %

实验 次数	跟踪精度			
	文献[1]方法	文献[2]方法	文献[3]方法	本文中方法
10	96	94	96	96
20	96	95	96	97
30	96	95	96	98
40	95	95	96	97
50	96	95	96	98
60	96	95	96	98
70	95	96	95	97
80	96	96	95	98
90	96	95	95	98
100	96	96	95	98

根据表 2 显示结果可以看出：随着实验次数的不断增加，本文中方法的跟踪精度始终高于对比方法，精度最高时达到了 98%，而文献[1]方法、文献[2]方法与文献[3]方法的跟踪精度最高分别为 96%、96%与 96%，说明本文中方法的跟踪效果较好。

利用中心误差和窗口重叠率 2 个指标评价本文中方法的跟踪性能，3 种方法的跟踪效果如表 3 所示。

表 3 3 种方法的跟踪效果

实验 次数	中心误差				窗口重叠率			
	文献[1]方法	文献[2]方法	文献[3]方法	本文中方法	文献[1]方法	文献[2]方法	文献[3]方法	本文中方法
10	1.06	0.83	1.01	0.53	0.96	0.94	1.00	1.00
20	1.06	1.22	1.23	0.55	0.96	0.92	0.95	1.00
30	1.08	1.35	1.54	0.61	0.95	0.90	0.93	1.00
40	1.09	1.64	1.64	0.62	0.95	0.89	0.92	0.99
50	1.11	1.78	1.95	0.64	0.94	0.89	0.90	0.99
60	1.11	2.03	2.32	0.64	0.94	0.88	0.89	0.99
70	1.12	2.19	2.54	0.67	0.92	0.85	0.87	0.99
80	1.19	2.54	2.98	0.69	0.92	0.85	0.80	0.98
90	1.20	2.56	3.31	0.69	0.91	0.84	0.78	0.98
100	1.20	2.98	3.33	0.71	0.90	0.84	0.78	0.98

根据上述表 3 可以看出：随着实验次数的不断增加，中心误差在不断增大，而重叠率在不断缩小，但本文中方法的中心误差始终小于其他对比方法，最小达仅为 0.53；重叠率始终高于其他对比方法，最高达到 1.0。具有较好的跟踪性能。

3 结束语

本文中方法基于卡尔曼滤波对红外图像进行了目标检测与跟踪的研究。在配备高性能计算资源的实验环境下，采用了实验室实测的红外图像数据集

进行了仿真实验，验证了所提出方法的有效性。

对比实验结果表明：所设计方法的跟踪效果较好，且随着实验次数的增加，本文中方法的中心误差始终小于其他对比方法，重叠率始终高于其他对比方法，具有较好的跟踪性能。此次研究仍存在一定不足，该方法对原始数据要求的较高，适用范围较小。在今后的研究中，会进一步改进原始数据预处理方法，增强目标检测算法的鲁棒性。