DOI:10.19652/j.cnki.femt.2204455

基于深度学习的红外过采样扫描图像小目标跟踪算法*

姚迎乐[1] 赵 娟[2]
(1.郑州工业应用技术学院信息工程学院 郑州 451150;2.武汉轻工大学数学与计算机学院 武汉 430023)

摘 要:红外小目标跟踪过程中由于背景、外界杂波等干扰,导致跟踪精确度和实时性欠佳,为此,提出基于深度学习的红外过采样扫描图像小目标跟踪算法。首先构建了红外过采样扫描图像模型,通过背景估计、形态学开运算,对图像中背景以及外界杂波进行多级滤除;然后增加设计特征融合模块和区域选取模块来改进孪生网络,生成融合特征图输入目标区域,通过分类和回归计算提高图像的特征表征能力和跟踪精度;最后建立损失函数训练孪生网络,输出红外过采样扫描图像小目标跟踪结果。实验结果表明,利用所提算法进行图像滤除后,信噪比能够高达35 dB,所提算法的区域重叠率较高、跟踪精度高,且算法的实时性强,帧率达到200 fps以上,整体跟踪效果好。

关键词:多级滤波;改进孪生网络;特征融合;区域选取;红外小目标跟踪
中图分类号:TP391.41 **文献标识码**:A **国家标准学科分类代码**:520.6

Small object tracking algorithm for infrared oversampled scanning images based on deep learning

Yao Yingle[1] Zhao Juan[2]
(1. Department of Information Engineering, Institute of Zhengzhou Industrial Application Technology, Zhengzhou 451150, China; 2. Department of Mathematics and Computer Science, Wuhan Polytechnic University, Wuhan 430023, China)

Abstract: In the process of infrared small object tracking, the tracking accuracy and real-time performance are poor. Therefore, the algorithm of infrared over-sampling scanning image small object tracking based on deep learning is proposed. Firstly, the infrared oversampling scanning image model is constructed, used to filter the background and external clutter, then add the design feature fusion module and area selection module to improve the twin network, generate the fusion feature map input target area, and improve the feature representation ability and tracking accuracy through classification and regression calculation. Finally, the loss function is established to train the twin network and output the small target tracking results of infrared oversampled scanning images. The experimental results show that the proposed algorithm can be up to 35 dB, the proposed algorithm has high regional overlap rate, high tracking accuracy, strong real-time algorithm, the frame rate reaches more than 200 fps, and the overall tracking effect is good.

Keywords: multi-level filtering; improve the twin network; feature fusion; area selection; infrared small target tracking

0 引 言

红外成像系统具有抗干扰性强、穿透能力强、隐私性强等诸多优势[1-2],因而红外成像技术在军事领域、智能安防、导航等领域中得到广泛应用[3-4]。目标跟踪是红外成像系统的关键技术之一,其中弱小目标跟踪由于容易受到背景等相关因素的干扰,跟踪难度增加,因而成为红外目标跟踪领域的一项难点问题,对其进行研究具有重要的理论意义和实际意义[5-7]。

近年来,国内外学者提出了许多关于红外目标跟踪的相关研究,取得了一定研究成果。文献[8]指出深度学习的快速发展使得红外成像机器视觉更加智能化,从而大大

收稿日期:2022-11-01
* **基金项目**:河南省科技厅科技攻关支持项目(202102210361,182102210555)资助

增加了其应用范围，该文论述了深度学习在红外成像中的应用，但具体应用过程分析不明确。文献[9]基于深度学习分析红外云图像与太阳光辐射问题，开发了一种多模态聚变网络，使用一个完全连接的神经网络提取和融合空间、时间信息，取得了一定的研究成果，但忽略了对红外图像背景的有效处理，导致图像质量不佳，影响了目标跟踪效果。文献[10]提出了一种基于可见光－红外图像对的多模态可见性深度学习模型，构造了一个新的卷积神经网络结构，其中包括3个并行连接的卷积流，通过传播的方法提取各流的特征图，并从底层到深层进行融合，最后根据全连接层的输出特征描述符，对可见性范围级别进行分类，跟踪精度高，但该文构建的模型结构复杂，实际运行效率欠佳，导致目标跟踪过程的整体实时性差。文献[11]将全局感知孪生网络应用于红外目标跟踪过程分析，加入了空间感知模块得到全局有效信息，最后通过响应融合完成红外目标跟踪，对于全局信息的获取效果较好，但在细节处比如红外背景滤波、图像边缘特征等方面考虑不足，使得实际应用中的跟踪精度还有待提高。

为改善上述算法存在的不足，本文引入深度学习这一热门技术，设计了一种新的基于深度学习的红外过采样扫描图像小目标跟踪算法，期望借助深度学习高效的图像和数据解释能力，有效提高红外小目标跟踪性能。同时设计了仿真实验来验证所提算法性能。

本文提出基于深度学习的红外过采样扫描图像小目标跟踪算法；对红外过采样扫描图像进行多级滤波处理，降低过采样扫描中背景及杂波的干扰，增强红外图像质量；设计特征融合模块和区域选取模块改进孪生网络，将其应用于红外小目标跟踪研究中，提高了模型的图像特征表征能力和跟踪精度。

1 基于过采样扫描的红外图像多级滤波

1.1 红外过采样扫描图像模型建立

为提高红外成像质量，逐渐出现了过采样扫描成像新体制概念，过采样扫描图像过程中，远距离目标的点状、线状特征明显，呈现出一定的规则性，背景成像更为平滑，因此红外过采样扫描图像质量更好，有利于提高小目标跟踪效果。

本文研究的红外过采样扫描图像主要包括跟踪目标、红外背景和外界杂波3个部分，根据构成因素建立红外过采样扫描图像模型：

$$I(x,y)=I_e(x,y)+I_b(x,y)+I_n(x,y) \tag{1}$$

式中：$I_e(x,y)$、$I_b(x,y)$、$I_n(x,y)$ 分别表示跟踪目标、红外背景和外界杂波；$I(x,y)$ 表示红外图像的灰度值。

1.2 红外图像多级滤波

分析上述模型可知，红外过采样扫描图像容易受到背景以及外界杂波的干扰，为此需要对其进行滤波处理[12]。首先是对原始红外图像进行形态学滤波，构建线性元素集 $A\{a_1,a_2,\cdots,a_n\}$ 用于背景估计，从而得到背景估计红外图像，公式如下：

$$E(x,y)=\max\{I\circ a_1(x,y),I\circ a_2(x,y),\cdots,I\circ a_3(x,y)\} \tag{2}$$

式中：$a_i(x,y)$ 表示不同的线性元素。

然后根据原始图像和背景估计图像，可分离得出背景抑制图，弱化背景强度，使得红外小目标更为清晰，信噪比更高，公式如下：

$$U(x,y)=I(x,y)-E(x,y) \tag{3}$$

最后对红外图像进行形态学开运算，寻求最优的结构元素[13]，进一步滤除背景干扰和外界杂波，完成红外图像小目标多级滤波，提高红外图像质量。公式如下：

$$M(x,y)=U(x,y)\circ s(x,y) \tag{4}$$

式中：$s(x,y)$ 表示形态学中的矩形结构元素。

2 基于深度学习的红外小目标跟踪算法

近年来深度学习在计算机视觉领域发展非常迅速，且应用性能良好，本文利用深度学习进行红外小目标跟踪研究。孪生网络是深度学习的一种，主要通过训练大量的视频图像，迁移到另一层网络结构中，完成小目标跟踪，孪生网络能够较大程度提高对红外应用场景的使用性能，对于背景遮挡、目标形变等具有较好的处理效果。因此本文将引入孪生网络来研究红外小目标跟踪问题。

2.1 孪生网络改进分析

传统的孪生网络多使用特征输出层结束对图像特征的分析，然而红外图像目标相对较为复杂，同时具有深层特征和浅层特征，感受野不同且分辨率不一致，仅仅使用传统的特征输出层对其进行特征分析，无法达到理想效果[14]。本文对孪生网络进行了改进分析，在特征分析部分设计了多特征融合模块，使获取得到的图像特征能够表征更多信息。

此外，在小目标相似物条件下或需精准计算目标尺寸时，传统的孪生网络难以实现，为进一步增强跟踪效果，本文在多特征融合模块设计的基础上，添加了区域选取模块，通过分类和回归计算提高跟踪结果的精确性。改进后的孪生网络结构如图1所示。

本文主要通过设计特征融合模块和区域选取模块改进孪生网络。

1)特征融合模块

该模块通过多个卷积层处理红外图像特征信息，提取得到图像特征并进行级联分析，完成特征融合，增加图像通道数量和表征信息。

考虑到红外过采样扫描图像相较一般图像较为复杂，细节特征多且空间信息丰富，为保证图像包含充足的空间信息和语义信息，同时能够挖掘得到更多的小目标细节信息，设计了特征融合模块，对红外图像进行卷积特征融合处理。融合操作过程中，将图像特征划分为不同尺度，对浅层特征进行最大池化操作，使其与深层特征在分辨率方

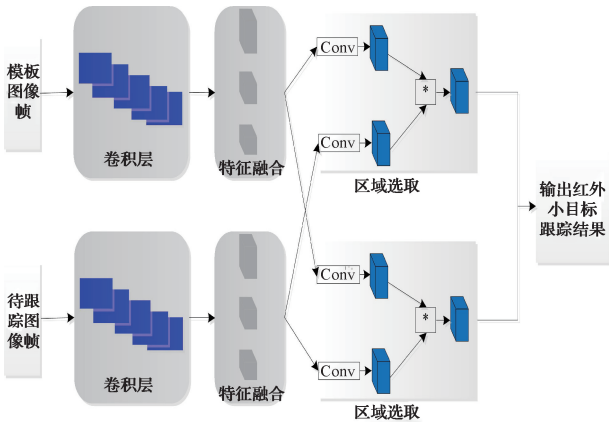

图1 孪生网络改进结构

面保持一致[15]。通常卷积层的前两层包含较大尺度的特征信息，因此此处认为后3层为浅层特征，那么浅层特征最大池化操作的表达式为：

$$p(f)=\{p(f_{conv3}),p(f_{conv4}),p(f_{conv5})\} \tag{5}$$

式中：$p(f_{conv3})$、$p(f_{conv4})$、$p(f_{conv5})$ 分别表示后3层卷积层特征的最大池化结果；$p(f)$ 表示浅层特征最大池化操作输出结果。

根据浅层特征最大池化结果进行红外图像特征融合处理，并对融合后的特征进行归一化操作，生成融合特征图。表达式为：

$$F_{fusion}=concat[bn(p(f))] \tag{6}$$

$$F'=conv(F_{fusion}) \tag{7}$$

式中：$bn()$ 表示归一化操作；$concat[]$ 表示连接函数；F_{fusion} 表示特征融合结果；F' 表示输出的特征图。

2)区域选取模块

本文设计的区域选取模块具有分类和回归计算两个功能，对背景和小目标的分类和对目标区域的回归。添加该模块改进后，孪生网络以红外图像特征融合结果为输入信息，经过分类和回归计算，输出红外图像小目标跟踪结果，改进后的网络具备更高的精确性。

红外图像特征融合后，将某一层卷积层在图1模板图像帧和待跟踪图像帧中的特征映射分别表示为 $\gamma(x)$、$\gamma(y)$。假设区域选取模块中包含 m 个锚点，对模板图像帧中该卷积层进行通道扩展，扩展为 $2m$ 和 $4m$ 个通道数量，那么 $\gamma(x)$ 的特征映射分支变化为 $\gamma_1(x)$、$\gamma_2(x)$，分别对应为分类分支和回归分支。

然后将待跟踪图像帧中的特征映射 $\gamma(y)$ 的变为 $\gamma_1(y)$、$\gamma_2(y)$ 两个分支，但通道数量不变，在这一状态下对 $\gamma(x)$ 和 $\gamma(y)$ 的两个分支进行卷积运算，公式如下：

$$R_{2m}=[\gamma_1(x)]\cdot[\gamma_1(y)] \tag{8}$$

$$R_{4m}=[\gamma_2(x)]\cdot[\gamma_2(y)] \tag{9}$$

式中：R_{2m} 和 R_{4m} 分别表示在分类分支和回归分支上图像特征输出的相关性。

由于分类分支包含 $2m$ 个输出通道，回归分支包含 $4m$ 个输出通道，那么分类网络输出的红外图像小目标位置处为两个激活函数，回归网络输出的锚点与标注框之间的距离为 da、db、dc、dg，由此选取得到目标区域，生成分类置信图和回归置信图[16]，输出该卷积层的特征提议集合 Q，表示为：

$$Q=\{a_i,b_j,c_k,g_l\} \tag{10}$$

式中：i、j、k、l 分别表示锚点位置。从特征提议集合 Q 中选取得出最佳提议，完成对红外图像小目标的跟踪处理。

2.2 基于改进孪生网络的红外小目标跟踪

利用2.1节改进后的孪生网络进行红外过采样扫描图像小目标跟踪研究，过程如下。

输入：红外过采样扫描图像模型；

输出：红外小目标跟踪结果。

对孪生网络的各项参数值进行初始化处理，并对其改进应用于跟踪过程分析，步骤如下。

1)建立红外过采样扫描图像模型，对图像进行背景估计、形态学开运算，完成对图像的多级滤波处理。

2)改进孪生网络。增加多特征融合模块和区域选取模块，对卷积层输出特征进行融合处理，然后将融合后的特征信息输入区域选取模块，开展分类和回归计算，获取最佳特征提议。

3)训练改进的孪生网络。建立损失函数，对离线训练的图像样本进行判别，计算公式为：

$$L(q,x)=\frac{1}{F'}\cdot\sum_{v}^{F'}\lg(1+\exp(-q(v)z(v))) \tag{11}$$

式中：v 表示特征图中的索引位置；$z(v)$ 表示索引位置 v 处的特征值；$q(v)$ 表示索引位置 v 处的样本标签；$L(q,x)$ 表示损失函数。

采用随机梯度下降算法训练整个改进孪生网络，优化网络参数，公式表示为：

$$\underset{\alpha}{\operatorname{argmin}}=L(q,h(x,\alpha)) \tag{12}$$

式中：α 表示改进孪生网络参数；$h(x,\alpha)$ 表示网络迭代函数。

4)对训练完成的孪生网络进行迭代训练，计算得到最优参数，使损失函数降到最低值。

5)迭代训练结束，输出红外过采样扫描图像小目标跟踪结果。

综合上述分析，跟踪流程如图2所示。

3 实验分析

3.1 实验环境

为证明本文算法的有效性，设计仿真实验对本文跟踪算法进行验证分析。实验采用Windows10系统运行程序，使用Python 3.6+Pytorch 1开源深度学习框架构建改进孪生网络。硬件环境为：Intel i7-8700kCPU，NVIDIA GTX980 M显卡，32 G内存。

3.2 实验数据及细节

实验以ILSVRC2015_VID数据集、KAIST行人数据

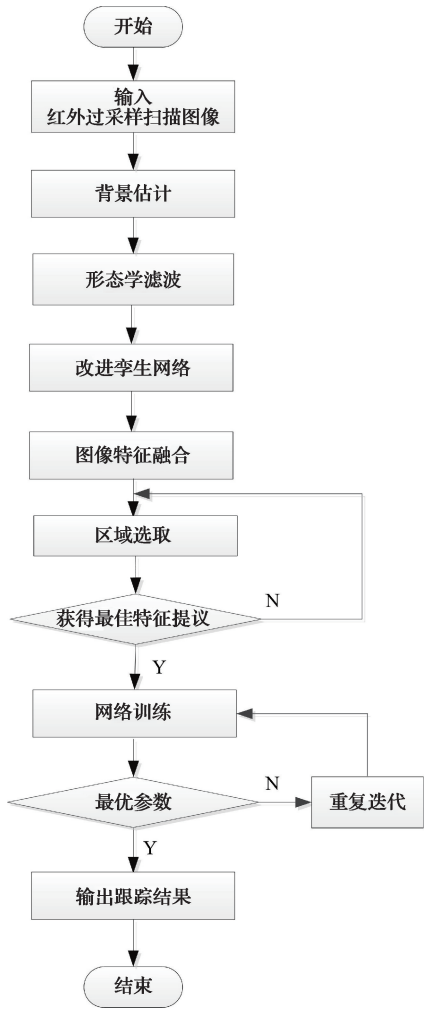

图 2　基于改进孪生网络的红外过采样扫描图像小目标跟踪流程

集和 OTCBVS 数据集为数据来源，对算法进行训练和测试分析。

ILSVRC2015_VID 数据集，该数据集为可见光数据集，包含天空、地空等多种场景，数据集中的图像序列有 22 个，图像帧数量为 16 177，实验主要以其中的红外图像弱小飞机目标为选取对象。

KAIST 行人数据集，总共包括 95 328 张图片，涵盖校园、街道以及乡下的各种场景，每种类型的图像都包含红外版本，图片大小为 640×480 pixels。

OTCBVS 数据集，该数据集含有红外行人图像、运动红外图像、红外车辆图像等多种类型，共计 13 个子数据集，4 228 对热图像和可见图像，数据量大、数据序列多，图像尺寸为 320×240 pixels。

从以上 3 个数据集中分别选取 10 个数据序列，按照 2∶8的比例随机划分为测试集和训练集，那么测试集中包含 6 个数据序列，6 785 帧图像，训练集中包含 24 个数据序列，27 140 帧图像。考虑到红外图像数量有限，采用交叉验证法开展实验分析。

数据集选取完成后，对红外图像统一进行预处理，统一像素大小为 640×480 pixels，同时滤除图像背景干扰和外界杂波干扰，然后将预处理完成的图像输入改进孪生网络中进行处理。网络训练时学习率选取范围为 0.05～0.000 1，设置网络迭代最大次数为 1 000 次。

3.3　实验指标

选取如下实验指标，将本文算法与文献[8-11]算法进行对比。

1)图像信噪比

本文首先对过采样扫描的红外图像进行了多级滤波处理，降低红外背景和外界杂波的干扰，为验证这一步骤的有效性，以图像信噪比(SNR)为指标验证本文算法。图像信噪比的计算公式为：

$$SNR = \frac{I(x,y)}{\sigma} \tag{13}$$

式中：σ 表示红外图像模型中噪声背景的标准差。

2)区域重叠率

改进孪生网络过程中，本文设计了区域选取模块，该模块运行效果的好坏对算法跟踪性能具有直接影响，为此，选取区域重叠率这一指标验证算法，计算公式为：

$$AOR = \frac{D \cap D_{real}}{D \cup D_{real}} \tag{14}$$

式中：D 表示算法计算得到的跟踪区域；D_{real} 表示红外小目标实际运动区域。

3)跟踪精度

精度是验证算法跟踪性能非

常直接且高效的指标，为此，计算本文算法与其他文献算法的跟踪精度，验证算法的性能。跟踪精度计算公式为：

$$Precision = \frac{P_{sim}}{P_{rea}} \times 100\% \tag{15}$$

式中：P_{sim} 表示利用算法计算得到的跟踪结果输出值；P_{rea} 表示红外小目标的实际跟踪结果值。

4)实时性

红外过采样扫描图像小目标跟踪多应用于空中物体，因此实时性是非常重要的一个评估指标，实验以帧率作为算法实时性评价指标，计算公式为：

$$FPS = \frac{\sum_{j=1}^{m} l_j}{\sum_{j=1}^{m} t_j} \tag{16}$$

式中：l_j 表示图像序列的长度；t_j 表示图像序列传输耗时。

3.4　实验结果分析

1)实验 1　图像信噪比

图像信噪比越大，表示噪声抑制效果越好，目标相对于背景更为明显，同时表明算法的图像处理效果更优。本文算法与文献[8-11]算法的图像信噪比对比如图 3 所示。

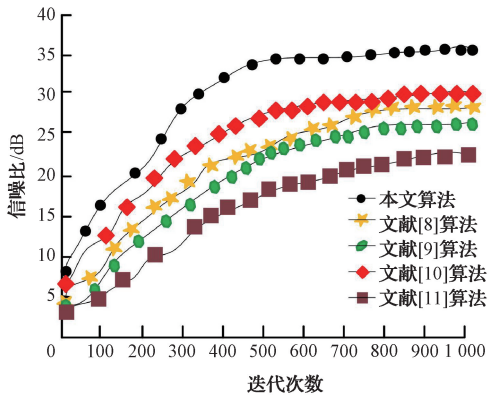

图 3　图像信噪比对比

从图 3 可以看出，采用本文算法对图像处理后，信噪比最高能够达到 35 dB。其他几种对比文献算法的信噪比较本文算法较低，文献[8-9]算法的最高信噪比不超过 30 dB，文献[10-11]算法的最高信噪比不超过 25 dB。对比可知，本文算法对图像的去噪效果较好，这是因为在构建模型的基础上，本文对红外图像进行了多级滤波，较好的抑制了背景和外界杂波干扰。

2)实验 2　区域重叠率

不同算法的区域重叠率对比结果如图 4 所示。

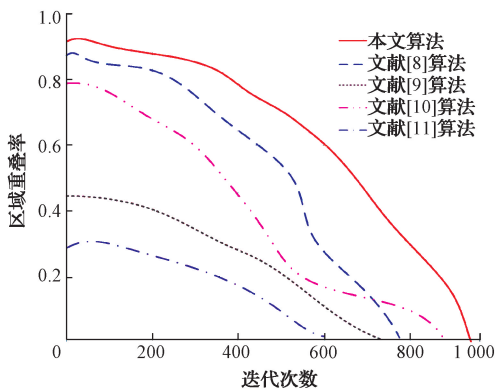

图 4　区域重叠率对比

分析图 4 可知，5 种算法随着迭代次数的增加，区域重叠率均呈现出下降趋势，但从图 4 可以明显发现，本文的重叠率曲线处于最上方位置，始终高于其他文献算法，重叠率高达 0.9 以上，相较于其他文献算法优势明显，由此验证了本文在孪生网络改进中设计区域选取模块的正确性，该模块有助于提高红外图像小目标跟踪效果。

3)实验 3　跟踪精度

为从更加直观的角度体现本文算法的跟踪性能，从实验数据集中随机选取 3 组原始红外图像，利用本文算法以及其他对比文献算法进行小目标跟踪，可视化呈现跟踪效果，如图 5 所示。

计算不同算法的平均跟踪精度，对比如图 6 所示。

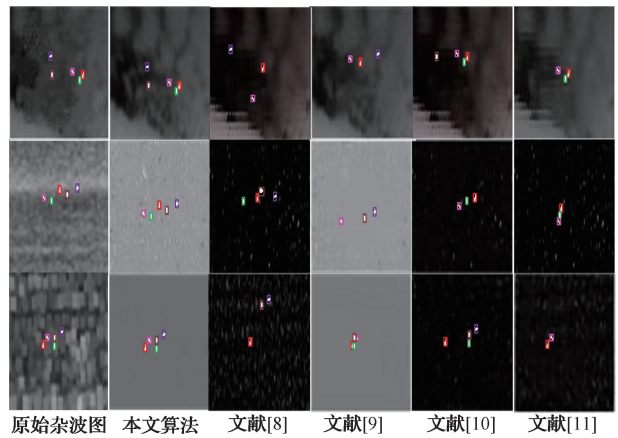

图 5　红外小目标跟踪示意图

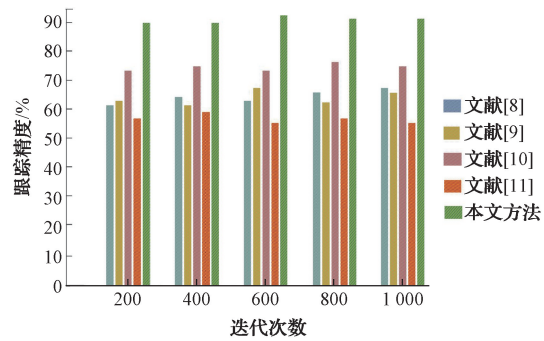

图 6　不同算法跟踪精度对比

从图 5 可以看出，文献[8-11]算法的红外小目标跟踪效果相对较差，均在不同程度上出现了小目标丢失的现象，跟踪区域漂移，且背景滤波效果不佳。可以发现本文算法对 3 幅图像中小目标的跟踪效果均较好，可以取得准确的跟踪结果，且背景滤波效果好，这进一步验证了本文对红外图像进行多级滤波的有效性。

分析图 6 可知，本文算法的跟踪精度在多次迭代下均能达到 90%左右，精度较高。其次是文献[10]算法，平均跟踪精度在 70%以上，文献[8-9]算法的平均跟踪精度在 60%左右，文献[11]精度最低，未达到 60%。根据数据对比可知，本文算法具有明显优势，能够高精度实现红外小目标跟踪，说明本文改进孪生网络的设计充分提高了算法的跟踪性能。

4)实验 4　实时性

在 3 种不同的数据集上，分别计算本文算法与其他几种文献算法的帧率，取平均值对比，测试不同算法的计算效率，结果如表 1 所示。

帧率越大，表明算法的跟踪速度越快，实时性越强。从表 1 数据可知，本文算法在 3 个不同数据集下得到的帧率分别为 201、202、205 fps，均超过了 200 fps。而文献[9,11]算法在 3 个数据集下的帧率均在 170 fps 以下，文献[8]更低，文献[10]算法的帧率相对其他对比文献较

表1 不同算法帧率对比 (fps)

算法	ILSVRC2015_VID 数据集	KAIST 行人 数据集	OTCBVS 数据集
本文	201	202	205
文献[8]	142	152	154
文献[9]	165	167	160
文献[10]	182	177	180
文献[11]	166	164	169

高,但最高不超过182 fps,与本文跟踪算法相比仍有较大差距。其原因是本文改进了孪生网络,图像特征融合和区域选取模块的设计大大提高了算法跟踪性能。

4 结 论

红外小目标跟踪对于推动航空航天智能技术等相关技术的发展具有非常重要的地位,因此本文提出基于深度学习的红外过采样扫描图像小目标跟踪算法。通过红外过采样扫描图像的多级滤波增强图像质量,滤除背景和外界杂波干扰。通过增加设计特征融合模块和区域选取模块改进孪生网络,并应用于跟踪算法,取得了较好的跟踪效果。仿真实验表明,本文所提算法在图像信噪比、区域重叠率、跟踪精度和实时性方面均展现出了优越的性能,对于红外图像在相关领域的应用具有良好的促进作用。

然而实际情况中红外小目标可能发生外观变化,甚至短时间内产生较大形变,本文对于这方面考虑不足,有可能导致目标形态差异较大,影响滤波及图像特征分析效果。因此在后续研究中,将会对红外小目标形态变化相关问题进行分析,在保证输入更准确形态信息的条件下不降低算法计算精度,进一步优化目标跟踪问题。

参 考 文 献

[1] 陈国洋,吴小俊,徐天阳. 基于深度学习的无监督红外图像与可见光图像融合算法[J]. 激光与光电子学进展,2022,59(4):1-10.

[2] 蔡军,谭静,邱会然. 基于局部积加权对比的红外弱小目标检测[J]. 电子测量与仪器学报,2021,35(12):133-141.

[3] 余仁伟,朱浩,蔡昌恺. 基于薄板样条函数的无人机多目标跟踪算法[J]. 仪器仪表学报,2021,42(3):168-176.

[4] 汪嘉鑫,徐贵川,于婷洋,等. 复杂红外背景中运动小目标快速跟踪技术[J]. 应用光学,2021,42(3):443-453.

[5] 徐亚杰,王海星. 基于新特征算子的复杂背景红外弱目标检测算法[J]. 国外电子测量技术,2021,40(12):7-11.

[6] 梁霄,李家炜,赵小龙,等. 基于深度学习的红外目标成像液位检测方法[J]. 光学学报,2021,41(21):104-112.

[7] 蔺素珍,武杰,禄晓飞,等. 基于多特征融合的红外弱小目标检测方法[J]. 电子测量技术,2022,45(8):108-115.

[8] HE Y, DENG B, WANG H, et al. Infrared machine vision and infrared thermography with deep learning: A review[J]. Infrared Physics & Technology, 2021, 116(2):103754.

[9] AJITH M, YAN J. Deep learning based solar radiation micro forecast by fusion of infrared cloud images and radiation data[J]. Applied Energy, 2021, 294(1):117014.

[10] SHEN K C, SHI Q, WANG H. Multimodal visibility deep learning model based on visible-infrared image pair[J]. Journal of Computer-Aided Design & Computer Graphic, 2021, 33(6):939-946.

[11] 李畅,杨德东,宋鹏,等. 基于全局感知孪生网络的红外目标跟踪[J]. 光学学报,2021,41(6):172-182.

[12] KIM C, KO H. Weighted Kernel filter based anti-air object tracking for thermal infrared systems[J]. Sensors,2020,20 (15): 4081-4086.

[13] 王睿男,武穆清,陈铁英,等. 基于形态学的红外图像边缘检测[J]. 北京邮电大学学报,2021,44(1):66-71.

[14] XU Y D, WANG Z Y, LI Z X, et al. SiamFC++: Towards robust and accurate visual tracking with target estimation guidelines[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12549-12556.

[15] 齐天卉,张辉,李嘉锋,等. 基于多注意力图的孪生网络视觉目标跟踪[J]. 信号处理,2020,36(9):1557-1566.

[16] 范东嘉,林名强,戴厚德,等. 用于目标跟踪的特征融合孪生网络算法研究[J]. 厦门大学学报:自然科学版,2022,61(4):714-722.

作 者 简 介

姚迎乐,硕士,副教授,主要研究方向为人工智能、深度学习、并行编译。
E-mail:yaoyingle2022@163.com

赵娟,博士,讲师,主要研究方向为人工智能、信息隐藏。