DOI:10.16136/j.joel. 2023.11.0500

一种新的多装甲目标在线跟踪方法

王慧敏*

(陆军装甲兵学院 兵器与控制系,北京 100072)

摘要:实现多装甲目标跟踪在协同跟踪和打击中发挥着至关重要的作用,而实现多装甲目标跟踪 需要解决目标之间遮挡、轨迹交叉以及目标尺度不断变化等问题。因此,提出了一种基于视觉注 意力的 Gabor 在线多装甲目标跟踪方法,实现对地面战场中多装甲目标的跟踪。通过模拟视网膜 结构,构造了一种能够增强检测的视觉注意力 Gabor 分支,并引入时间信息,采用在线学习的目 标特性卷积神经网络解决目标遮挡的问题,更重要的是,通过实际拍摄和互联网下载等手段构建 了多装甲目标跟踪数据集,并且通过对目前成熟的多目标跟踪方法和本文提出的方法进行对比 实验。实验证明了本文提出的方法不仅具有优异的跟踪性能,而且能够满足实际的应用需求。 关键词:多装甲目标;在线跟踪;视觉注意力;Gabor;卷积神经网络 中图分类号:E923;TJ811 文献标识码:A 文章编号:1005-0086(2023)11-1178-09

A new method for online tracking of multi-armored targets

WANG Huimin*

(Weapons and Control Department, Army Academy of Armored Forces, Beijing 100072, China)

Abstract: The realization of multi-armored target tracking plays a vital role in cooperative tracking and strike, and the realization of multi-armored target tracking needs to solve the problems of occlusion, interspersed and constantly changing target scales between targets. Therefore, an online multi-armored target tracking method based on visual-attention Gabor filter is proposed to achieve the tracking of multi-armored targets in the ground battlefield. A visual-attention Gabor filter branch is constructed to enhance detection by simulating the retinal structure. By introducing temporal information, the problem of target occlusion is solved by using an online learned target-specific convolutional neural network. What is more important, a multi-armored target tracking dataset is constructed by means of actual shooting and downloading from the internet, and the current mature multi-target tracking methods are compared with the method proposed in this paper through experiment. The experiments show that the method in this paper not only has excellent tracking performance, but also can meet the actual application requirements. **Key words**; multi-armored targets; online tracking; visual-attention; Gabor filter; convolutional neural network

0 引 言

多装甲目标跟踪是全自动观测和攻击型装备 (如无人机、装甲侦察车等)研究的热点,对掌握地 面战场动态至关重要。

与其他常见的多目标跟踪不同,多装甲目标 跟踪不仅面临着装甲目标更微小、地面战场环境 更复杂等难题,还要满足视觉跟踪系统的跟踪距 离远大于装甲目标攻击距离的要求。通常多装甲 目标的作战环境包括丛林、沙漠、草原以及诸如灰 尘、炮口火光和烟雾等干扰,因此在目标跟踪过程 中常遇到遮挡和轨迹交叉等问题,并缺乏成熟的 多装甲目标跟踪数据集。

基于以上问题,本文提出了一种基于视觉注意力的 Gabor 在线多装甲目标跟踪方法,具体内容包括:1)构建了离线候选区域推荐模块;2)提出了一种空间注意力分支;3)提出了一种在线候选区域评估模块;4)提出了一种新的使用时间信

^{*} E-mail:h_m_wang@126.com 收稿日期:2022-07-04 修订日期:2022-10-08

息的运动模型;5)构建了多装甲目标跟踪数据集。

1 研究现状

1.1 装甲目标检测

装甲目标检测作为多装甲目标跟踪算法的基 石,关键是提取出有效的特征,一般采用基于红外 图像的方法和基于可见光图像的方法。在早期阶 段,装甲目标的特征采用人工提取方式,具体包 括:梯度方向直方图法、尺度不变特征转换法等 等,同时,基于人体工程学提取的特征通常与支持 向量机(基于浅层学习的判别分类器)结合在一起 使用。然而,基于人工提取特征的方法对遇到的 特殊问题通常是有限的,且当装甲目标的外观为 小尺度或变化很大时,采用人体工程学提取特征 有可能是失败的。随着基于候选区域的卷积神经 网络的成功应用,一系列基于卷积神经网络及其 变体的装甲目标检测方法得到了发展。例如: SUN 等^[1]提出了一种 TPA (top-down aggregation)装甲目标检测方法。ZHAO 等^[2]使用一种更 快速的区域卷积神经网络和图像金字塔来检测装 甲目标。基于区域卷积神经网络的方法显著提高 了装甲目标检测的准确性。然而,在上述方法中, 由滑动窗口算法或参考包围框算法导致的计算耗 损问题以及小型装甲目标检测准确度不足问题仍 未得到解决。

1.2 多目标跟踪

目标跟踪是计算机视觉研究的热点,也是视 频处理过程中的一项基本任务,其主要对感兴趣 的目标进行某种程度的推理,尤其是处理多目标 跟踪中出现的遮挡、尺度变化和轨迹交叉等问题。 在国外,目标跟踪主要研究基于判别分类的跟踪 方法,而国内大多还是使用基于模型匹配的跟踪 方法,但对于复杂场景中的图像特征分析和定量 (定性)评价,还没有较好的解决方案。受卷积神 经网络成功应用的启发,已经提出了许多利用分 层卷积神经网络实现视觉跟踪的方法。齐元凯 等[3]提出了一种自适应技术,通过利用水平集分 割技术和包围框定位以获得紧密的封闭框,并且 设计一种卷积神经网络来识别目标是否被遮挡。 齐元凯等[4] 提出了一种基于卷积神经网络的跟踪 算法,其主要是将不同卷积神经网络层的深层特 征进行组合,以更好地区分目标对象和背景杂波。 齐元凯等^[5]提出将跟踪中遇到的遮挡、形变之类 的问题作为每一帧图像的属性信息,进而训练出 一种有效的视觉跟踪分类器。目前,基于检测结 果的跟踪策略通常用于大多数成熟的多目标跟踪 方法,主要是由检测模块和跟踪模块组成。张盛平^[6]提出一种新的基于基础匹配的鲁棒跟踪框架。储琪等^[7]使用时空注意力机制实现多目标跟踪,他们使用传统的候选区域选择网络对深度卷积特征进行操作以生成目标候选区域,并采用多目标特性卷积神经网络分支来评估目标候选区域。HENSCHEL等^[8]演示了如何将检测器应用于跟踪系统并解决了跟踪过程中的加权图标记问题。张元平等^[9]提出了一种用于多目标跟踪的混合可变形卷积神经网络。

2 多装甲目标在线跟踪方法

考虑到多装甲目标跟踪中出现的所有难题,本 文提出了一种基于视觉注意力的 Gabor 在线多装甲 目标跟踪方法。

聚焦于地面战场的多装甲目标跟踪问题,本文提出了一种基于 SSD(single shot multi-box detector)^[10]框架和视觉注意力的 Gabor 关联滤波器跟踪方法,如图 1 所示。



图 1 基于视觉注意力的 Gabor 在线多装甲目标跟踪方法 Fig. 1 Visual-attention Gabor filter based online multi-amored target tracking method

多装甲目标跟踪方法主要包括两部分:

1) 离线候选区域推荐模块

离线候选区域推荐模块具备常规检测和精细化 检测两种模式,主要用于生成感兴趣特征(亦称目标 候选区域),该特征由多层候选区域推荐分支和视觉 注意力的 Gabor 分支共同生成,其中视觉注意力的 Gabor 分支采用了模拟视网膜技术,通过增强浅层特 征,实现对小型装甲目标检测的目的。

2) 在线候选区域评估模块

在线候选区域评估模块包括运动模型和目标特 性评估卷积神经网络两部分,分别用于评估遮挡问 题和预测装甲目标的精确位置。

通过离线和在线模块,多装甲目标的位置和类 别得到评估并实现不同帧图像之间的全连接。

2.1 离线候选区域推荐模块

基于深度卷积神经网络的离线候选区域推荐模 块已被广泛应用于多目标跟踪的检测阶段,主要采 用了基于固定比例和纵横比的参考包围框策略生成 候选区域。为了提高检测的准确度,大量的包围框 用于提取卷积深度特征。显然,在要求高度实时性 的广阔地面战场环境中,采取穷举策略进行多装甲 目标跟踪计算的将是极其耗时的,且检测的浅层特 征中仅包含了表层信息,易引起小目标检测错误。 为解决上述问题,使用了多层候选区域推荐分支和 视觉注意力的 Gabor 分支共同生成目标候选区域, 从而加强了小目标浅层特征候选区域的生成。离线 候选区域推荐模块如图 2 所示,图中包含两种模型, 模型1用于视频初始化输入和新目标轨迹交叉的精 准检测;模型2常用在没有新目标轨迹交叉、非初始 化视频输入等情况,此外,模型2属于模型1的一部 分,主要用于增强小目标浅层特征的检测。



Fig. 2 Offine candidates recommendation module

2.2 视觉注意力的 Gabor 分支

视网膜作为眼睛视觉的核心功能部分,其包含 了 5 个类型的神经细胞:视感受细胞(视杆细胞和视 锥细胞)、水平细胞、双极细胞、无足细胞和神经节细 胞,由上述细胞可形成自中心至外围的感受野(receptive field, RF),如图 3 所示。RF 类型可分为中 心型(图 3(a))和偏心型(图 3(c))两类,其中,中心型 RF 对中心区域的兴奋度有积极的响应。与之相反, 偏心型 RF 对中心区域的兴奋度有消极的响应。图 3(b)和图 3(d)分别显示了中心型和偏心型响应频率 与视网膜 RF 中心距离的变化关系。很明显,当 RF 中心和边缘之间的边界与兴奋度边界重合时,响应 频率是最强的。





中心至外围的 RF 能够增强视网膜细胞检测边 界的能力。RIAZ^[11]等提出了用 Gabor 来模拟视网 膜 RF 的特征和响应,该项技术广泛应用于提取空间 局部光谱特征和多模式识别。

传统的二维 Gabor 可表示为: $G_{\sigma,\theta,\lambda,\gamma,\Psi}(x,y) = g_{\sigma,\gamma}(x,y) \cdot \exp\left[i\left(2\pi \frac{x\cos\theta + y\sin\theta}{\lambda} + \Psi\right)\right],$ (1)

式中, $i^2 = -1$,x、y 表示每个像素的坐标, $g_{\sigma,\gamma}(x,y)$ 是高斯包络,其余 exp 指数部分代表了正弦光栅, σ 表示高斯包络的标准偏差,其控制了 Gabor 的 RF, γ 代表了空间的纵横比, θ 、 λ 、 Ψ 分别表示正弦光栅的方 向、波长和相移。高斯包络定义为:

$$G_{\sigma,\gamma}(x,y) = \frac{1}{2\pi\sigma^2} \cdot \exp\left[\frac{-(x^2 + \gamma y^2)}{2\sigma^2}\right]_{\circ} \quad (2)$$

图 4 中(a)—(c)显示了传统二维 Gabor 相关的 仿真图,分别表示传统的高斯函数、中心型定向复杂 正弦光栅、传统 Gabor。通过使用高斯函数,传统的 Gabor 仅模拟了中心型 RF。大量实验表明,中心型 RF 对提高目标的内部特征具有明显的作用,而偏心 型 RF 对提高目标的边缘特征具有明显的作用。为 了模拟视网膜的 RF,对传统的高斯包络进行改进, 表示为:

$$g'(x,y) = \begin{cases} g_{\text{On}} = \frac{1}{2\pi\sigma^2} \cdot \exp\left[\frac{-(x^2 + \gamma y^2)}{2\sigma^2}\right] \\ g_{\text{Off}} = \max(g_{\text{On}}) + \min(g_{\text{On}}) - g_{\text{On}} \end{cases}$$
(3)

式中,gon和 goff用于模拟中心型和偏心型高斯函数。 图 4 中(d)—(f)分别显示了偏心型高斯函数、偏心型 定向复杂正弦光栅和偏心型 Gabor。



在深度卷积神经网络中,浅层能够反映角点、纹 理和其他边缘特征。因此,采用了基于视觉注意力 的 Gabor 分支,该分支具体包括增强特征型 Gabor 组、目标位置预测分支和目标形状预测分支,用于增 强浅层的小目标候选区域生成。

此外,包含所有目标所用的参考包围框数量越 少,意味着定位和分类的计算量越少。为了弥补滑 动窗口算法和参考包围框算法的不足并保留目标候 选区域评估的计算,在 Gabor 组末端利用时空语义 特征来预测目标的位置和形状,该特征不同于传统 的语义事件,首先根据运动目标的跟踪结果和运动 方向建立目标之间的空间关系(拓扑关系和方向关 系)模型,进而得到目标之间的空间语义(前面、后 面、背对、面对以及左右)特征;然后在每个时间间隔 内捕捉目标运动对空间和外部因素的依赖性,突出 时间与空间之间的紧密关系,得到相对空间信息的 时间语义特征,具体体现在 Gabor 组末端的两个分 支 $N_{\rm P}$ 和 $N_{\rm s}$,分别用于预测目标的位置和形状,其中 分支 N_P 的位置预测图像 M_P 可表示为:

$$M_{\rm P} = f_{\rm P}(F_{\rm E}, \omega_{\rm P}), M_{\rm P} \in R^{\omega \times h}, \qquad (4)$$

式中, $F_{\rm E}$ 代表 Gabor 组增强型特征, $\omega_{\rm P}$ 是分支 $N_{\rm P}$ 中的参数集, $f_{\rm P}$ 为 1×1 卷积的 S 型激活函数。每个 $M_{\rm P}(i,j)$ 对应坐标的位置[$(i+\frac{1}{2})S, (j+\frac{1}{2})S$],其 中 S 是特征图的步长。同样,在分支 $N_{\rm S}$ 中,两个目 标形状预测图 $M_{\rm S}^{(1)}$ 和 $M_{\rm S}^{(2)}$ 用于预测目标的面积和 纵横比,表示为:

 $\left[M_{\mathrm{S}}^{\scriptscriptstyle(1)},M_{\mathrm{S}}^{\scriptscriptstyle(2)}\right]=f_{\mathrm{S}}(F_{\mathrm{E}},\omega_{\mathrm{S}}),$

$$M_{\rm S}^{(1)}, M_{\rm S}^{(2)} \in R^{w \times h},$$
 (5)

式中, f_s 为一个能产生双通道地图的 1×1 卷积层模型函数。为了获得固定的目标宽度 w 和高度 h,采用如下变换:

$$w = s \cdot e^{dw}, h = s \cdot e^{dh}, \qquad (6)$$

式中,s是步幅,dw和 dh 代表了双通道图像数值。

在分支 N_P中,使用全局阈值 ε_P 过滤掉装甲目 标不可能存在的区域,在确定目标中心后,包围框的 形状由式(6)实现预测。

2.3 在线候选区域评估模块

在检测阶段,边界附近新目标的轨迹交叉和场 景中目标的消失可以通过固定的目标包围框得以解 决。然而,当装甲目标的外观特征被污染,以及被其 他目标、建筑物、炮口火光和烟雾遮挡时,离线候选 区域推荐模块检测的目标位置可能是不准确的,因 此,提出了一种在线候选区域评估模块,以此来解决 上述问题,该模块包括目标特性评估卷积神经网络 和运动模型。

2.3.1 目标特性评估卷积神经网络

由于离线训练型检测器经常出现目标检测遗漏 或着仅检测到被遮挡目标的部分区域,因此,检测的 结果无法作为目标的实际位置。同时,由离线模块 推荐的深度候选区域特征(共享卷积神经网络得到 的)往往忽视遮挡问题,所以,本文提出了在线目标 特性卷积神经网络来评估每个候选区域。令感兴趣 特征池中的第k个候选区域 $C^{(k)}$ 作为第t帧图像中 的 $\Phi_{roi}[X_{t}^{(k)}], X_{t}^{(k)}$ 代表了候选区域 $C^{(k)}$ 的状态,表 示为:

 $X_{t}^{(k)} = [x_{t}^{(k)}, y_{t}^{(k)}, w_{t}^{(k)}, h_{t}^{(k)}],$ (7) 式中, $x_{t}^{(k)}$ 和 $y_{t}^{(k)}$ 代表了候选区域的中心坐标, $w_{t}^{(k)}$ 和 $h_{t}^{(k)}$ 表示候选区域的宽度和高度。对于每一个候选 区域 $C^{(k)}$, 遮挡分数 $p_{t}^{(k)}$ 用于评估遮挡的程度:

 $p_{t}^{(k)} = f_{est} \left[\Phi_{roi}(X_{t}^{(k)}), \boldsymbol{\omega}_{est}^{(k)} \right], p_{t}^{(k)} \in [0, 1], (8)$ 式中, $\boldsymbol{\omega}_{est}^{k}$ 是第 k 个目标特性卷积神经网络的参数 集, $f_{et}(*)$ 为3个卷积层(内核大小=3×7×32,1× 1×32,3×7×32)和一个全连接层(输出大小=1)的 模型函数。 $p_t^{(k)}=0$ 表示装甲目标完全被遮挡,采用 分类阈值 ϵ_0 用于确定是否被遮挡,当遮挡分数小于 分类阈值 ϵ_0 时,认为装甲目标被遮挡。

2.3.2 运动模型

运动模型通过分析历史帧图像中目标的运动曲 线,实现当前帧图像中目标位置的预测。在大多数 多目标跟踪应用中,采用了简单的线性运动模型,但 是目标不可能总是做直线运动,比如快速转弯、急 停、倒车或者目标形状发生变化等情形,而上述情形 均可用遮挡分数的变化进行表征。为了准确地预测 位置并实现在不同帧图像中对多装甲目标打标签, 本文提出了一种新的使用时间信息的运动模型,该 模型包括目标位置和形状预测模型。通过选取各装 甲目标的检测包围框中心为参考点,设定第k 个装 甲目标候选区域在t 时刻的状态:位置 $O_t^{(k)}$ 和形状 $S_t^{(k)}$,并给定第k - 1、k 个候选区域的运动速度 $v_{t-1}^{(k)}$ 、和第k 个候选区域在t - 1 时刻的形状 $S_{t+1}^{(k)}$,然后引入遮挡分数 α ,建立中心位置 $O_{t+1}^{(k)}$ 和形 $\widetilde{O_{t+1}^{(k)}} = O_t^{(k)} + \alpha v_t^{(k)} + (1-\alpha) v_{t-1}^{(k-1)}, \quad (9)$ $\exists t + \mathbf{O}_t = [\mathbf{x}_t, y_t], \alpha \ \text{a} \ \text{k} \ \text{f} \$

$$v_{t}^{(k)} = \frac{1}{M_{t} - M_{t-1}} \left\{ \left[x_{t}^{(k)}, y_{t}^{(k)} \right]^{\mathrm{T}} - \left[x_{t-1}^{(k)}, y_{t-1}^{(k)} \right]^{\mathrm{T}} \right\},$$
(10)

式中, *M*_i 表示运算速度的时间差,因此,预测的候选 区域形状可以表示为:

$$S_{t+1}^{(k)} = \alpha S_t^{(k)} + (1 - \alpha) S_{t-1}^{(k)}, \qquad (11)$$

式中,*S_t*=[*w_t*,*h_t*]。图 5 显示了离线候选区域的推荐模块和运动模型(含遮挡分数)基于 MATLAB 软件平台仿真的预测结果,短虚线包围框代表了装甲目标的实际大小区域,粗实线包围框代表了装甲目标未受遮挡的区域,由多装甲目标跟踪数据集提供,当装甲目标实际大小区域与未受遮挡区域重合时,以粗实线包围框显示为主,细实线包围框表示运动模型预测的结果,长虚线包围框表示离线候选区域推荐模块预测的结果,图中也标注了遮挡分数。随着装甲目标逐渐被遮挡,在线评估模块预测的遮挡分数也会随之减少。



图 5 离线候选区域推荐模块和运动模型(含遮挡分数)的预测结果

Fig. 5 The prediction results of the offline candidate recommendation module and the motion model with occlusion scores

2.4 在线跟踪方法的训练

2.4.1 目标预测分支的训练

在视觉注意力的 Gabor 分支中,两个分支 N_P 和 N_s 分别用于预测装甲目标的位置和形状。为了训 练两个分支,需要用到如下 3 种类型损失函数来描 述多目标跟踪任务中预测值与真实值之间的偏差:

 $L = L_{pos} + L_{cls} + L_{reg}$, (12) 式中, L_{pos} 表示目标位置损失函数,用来表示包围框 彼此中心之间的距离偏差, L_{cls} 表示目标分类损失函 数(采用交叉熵损失函数),用来表示正样本和负样 本的损失,输出为有限个离散型变量,L_{reg}表示定位 损失函数,用来表示包围框的高度和宽度损失,输出 为连续型变量。3类损失函数往往用于检测模型的 最后,根据模型输出(位置、类别和定位)和实际包围 框分别计算位置损失、类别损失和定位损失。

令 $(x_g, y_g, w_g, h_g)_n$ 表示第n个绘制的真实包围框,则二元标记图像的中心区域可表示为:

$$\mathbf{R}_{C} = (x_{g}, y_{g}, 0.1w_{g}, 0.1h_{g})_{\circ}$$
(13)

负区域是指不包括绘制真实包围框的特征图, 可表示为: $\mathbf{R}_n = F/(x_g, y_g, 0.1w_g, 0.1h_g),$ (14)

式中,F表示增强型特征图,在装甲目标形状预测分支中,采用了定位损失函数。

2.4.2 在线候选区域评估模块的训练

在线候选区域评估模块用于评估装甲目标的遮 挡和确切位置。在跟踪的初始阶段,目标特性卷积 神经网络的参数是随机的,且该网络缺乏评估能力。 为了能够得到很好的鲁棒性能,在线目标特性卷积 神经网络需要足够的样本进行训练,运动模型也需 要足够的帧图像来分析装甲目标的运动曲线。通过 实验,使用 $N_{init} = 0.1N_V$ 来训练初始在线模块,其中 N_V 代表了视频中所有帧图像的数量。对于不足 100 帧图像的视频,可采用前 10 帧图像完成初始在线模 块的训练,并利用误差反向传播算法实现在线模块 的更新。

为了得到充足的训练样本,采用了离线候选区 域推荐模块的检测结果作为初始阶段的正样本。令 $Sample_{P}^{n} = (x_{d}, y_{d}, w_{d}, h_{d})_{n}$ 表示某一种检测结果和 正样本。负样本是由对应的正样本和位置偏移构 成,可表示为:

 $Sample_{N}^{n} = (x_{d}, y_{d}, \sigma_{1}w_{d}, \sigma_{2}h_{d})_{n},$ (15) 式中,参数 σ_{1} 和 σ_{2} 在区间[0.7,0.9]和[-0.9, -0.7]内随机选择。

初始阶段过后,在线候选区域评估模块是由评估结果实现更新。在包围框内,不良样本中污染的特征会降低评估模型对目标和背景进行区分的能力,直至候选区域无法进行评估。为解决冲突,引入了一个时间注意力参数来平衡当前和历史中的样本,可表示为:

$$\lambda_t^k = \begin{cases} 0, p_t^k \le \varepsilon_0 \\ 0, 8, p_t^k > \varepsilon_0 \end{cases}$$
(16)

对于第 k 个候选区域 C^k, t 帧图像中目标特性的 损失函数可表示为:

$$L = L_{t}^{k-} + \lambda L_{t}^{k+} + (1-\lambda)L_{h}^{k+}$$
, (17)
式中, L_{t}^{k-} 表示当前帧图像中负样本的损失, L_{t}^{k+} 表示
当前帧图像中正样本的丢失, L_{h}^{k+} 表示历史帧图像中
正样本的损失,分别表示为:

$$L_{t}^{k-} = -\frac{1}{N_{t}^{k-}} \sum_{i=1}^{N_{t}^{k}} \lg \left\{ 1 - f_{\text{est}} \left[\Phi_{\text{roi}}(X_{t}^{k-}), \boldsymbol{\omega}_{\text{est}}^{k} \right] \right\}, \qquad (18)$$

$$L_t^{k+} = -\frac{1}{N_t^{k+}} \sum_{\substack{i=1\\ N_t^{k+}}}^{N_t} \lg f_{\text{est}} [\boldsymbol{\Phi}_{\text{roi}}(X_t^{k+}), \boldsymbol{\omega}_{\text{est}}^k], \quad (19)$$

$$L_{t}^{k+} = -\frac{1}{N_{h}^{k+}} \sum_{i=1}^{m_{h}} \lg f_{\text{est}} [\Phi_{\text{roi}}(X_{t}^{k+}), \omega_{\text{est}}^{k}], \quad (20)$$

式中,N^{k-}、N^{k+} 和 N^{k+} 分别表示当前帧图像中负、正 样本和历史帧图像中正样本的数量。

3 实验测试

3.1 数据集和实施细节

目前,常见目标和车辆目标的检测数据集已有 相关研究,如 VOC(visual object classes)^[12]和 KIT-TI^[12]数据集。然而,多装甲目标跟踪数据集仍属空 白。为了解决多装甲目标跟踪数据集缺乏的问题, 采用实际拍摄和互联网下载等手段获取了 80 个视 频序列,从而建立了多装甲目标跟踪数据集。

所有视频中包含 11536 个地面战场场景帧图像 和 30132 个装甲目标,一方面,所有的视频帧图像被 归一化为 1920 pixel×770 pixel大小,并使用图形图 像标注工具 LabelImg 以 PASCAL VOC 格式进行标 记;另一方面,多装甲目标跟踪数据集中的装甲目标 尺度变化范围很广,从 10 pixel×10 pixel增加至超 过 700 pixel×700 pixel,尤其在 30132 个装甲目标 中,绝大多数是远程小型装甲目标。在实验中选取 50 个视频序列作为训练集,剩余的 30 个视频序列作 为测试集,在 VOC2007、KITTI 和多装甲目标跟踪 数据集中目标实例大小的分布如图 6 所示。与常见 目标和车辆目标数据集相比,多装甲目标跟踪数据 集的平均目标尺度明显更小,更符合地面战场的环 境特点和作战样式。



本研究采用预先训练的 SSD512 模型作为多层 候选区域推荐分支的核心且所有框架运行在具有 11 GB 内存的 NVIDIA GeForce GTX 2080Ti 图形处理 器上。

3.2 指标评估

为了实现对多装甲目标跟踪方法的评估,主要 采用了召回率(recall)对离线候选区域推荐模块进 行量化评估,可表示为:

$$recall = \frac{tp}{t\bar{p} + f\bar{n}},\tag{21}$$

式中,tp(true positive)表示正确检测到的正样本数, fn(false negative)表示错误检测到的正样本数。评 估跟 踪方法采用了 CLEAR MOT(multi object tracking)^[13]指标,该指标包括多目标跟踪精度(muti-object tracking precision, MOTP)和多目标跟踪 准确度度(muti-object tracking accuracy, MOTA)两 种评价标准。多目标跟踪精度可以表示为:

$$MOTP = \frac{\sum_{k,t} d^k_t}{\sum_t N_t},$$
(22)

式中,d^k 表示第 k 个真实边界框与对 t 帧图像跟踪 结果之间的距离,N_t 代表进行匹配的总次数。 MOTP 反映了多目标跟踪器精确评估目标位置的能 力。而多目标跟踪准确度可以表示为:

$$MOTA = 1 - (FP + FN + IDS) = 1 - \frac{\sum_{k} (fp + fn + ids)}{\sum_{i} g_{i}}, \quad (23)$$

式中,g,表示在第 t 帧图像中出现目标的数量。 MOTA 反映了所有目标的匹配错误信息,包括误报、 丢失、不匹配。此外,使用多目标跟踪率和多目标丢 失率的百分比作为本文研究内容的指标。

3.3 离线候选区域推荐模块分析

为了解决参考包围框算法运行耗时长和浅层特 征检测出现错误等难题,提出了一种新的离线候选 区域推荐模块。该离线模块包括两个分支,分别为 多层候选区域推荐分支和视觉注意力 Gabor 分支, 前者是基于 SSD 框架的通用目标候选区域,后者是 用于增强浅层的小目标候选区域生成。离线模块包 含两种模式:

模式1:多层候选区域推荐分支+视觉注意力的 Gabor分支。

模式 2:SSD 浅层+视觉注意力的 Gabor 分支。

为了演示离线模块各个部分的功能,将上述模 式与下列模式进行检测召回率的比较。

模式 3:基于 SSD 的多层候选区域推荐分支。

实验结果如表 1 所示。检测召回率的结果共划 分为 3 个不同的子集:小型装甲目标(*Size* < 32 × 32)、中型装甲目标(32 × 32 < *Size* < 96 × 96)和大 型装甲目标(Size >96×96)。与模式3相比,模式1 具有附加的视觉注意力Gabor分支,模式1和模式3 对比结果表明,视觉注意力Gabor分支可以显著提 高小、中、大型规模目标检测召回率,分别为15.7%、 4.9%和0.4%,原因是Gabor分支可以增强浅层特 征,并能有效地预测装甲目标的位置。与模式1相 比,模式2在不同大小目标上的检测召回率略有下 降,且由于模式2的平均检测时间要小于模式1的平 均检测时间,反映了模式2的检测效率高于模式1, 因此,模式1用在初始视频输入和新目标轨迹交叉 的情况下,而模式2用在非初始视频输入和没有新 目标轨迹交叉的情况下。

表 1 离线候选区域推荐模块在

不同模式下的检测召回率及平均检测时间

Tab. 1The detection recall rate and average detectiontime of different modes of the offline candidate

recommendation module

Method	Small /%	Medium /%	Large /%	Average detection time/s
Mode 1	89.2	95.4	92.5	12.5
Mode 2	88.5	94.3	90.1	7.30
Mode 3	73.5	90.5	92.1	10.4

3.4 基准评估结果

为了证明多装甲目标跟踪方法的有效性,在装 甲目标跟踪数据集中,与几种成熟的方法进行对比 实验,进行比较的方法包括离线跟踪方法(如 SiameseCNN^[14]、CNNTCM^[15]、DCO-X^[16]和LP-SS-VM^[17])和在线跟踪方法(如 NOMTHM^[18]、 STAM^[19]、SSP^[20]、MOTBeyondPixels^[21]、Deep-SORT^[22])。

通过对几种成熟的方法进行跟踪测试,得到了 如表 2 所示的跟踪测试结果,其中在 MOTA、MOTP 和多目标跟踪率(mostly tracked,MT)方面,分别为 85.65%、87.55% 和 80.52%,而在多目标丢失率 (mostly lost,ML)方面,仅比 DeepSORT 高 0.05 个 百分点。较高的 MOTA 和 MT 表明了离线模块具 有更高的检测能力,较高的 MOTP 表明了运动模型 能够准确地预测目标运动的轨迹,多目标丢失率的 下降表明本方法具有很低的误报率,跟踪时间最短 表明离线候选区域推荐模块的两种模式不仅满足召 回率,而且提高了跟踪效率,且总的运行时间能够满 足实际要求。

Tab. 2 Comparison with several mature methods on tracking testing in armored targets tracking dataset									
Method	Mode	$MOTA/\frac{0}{0}$	$MOTP/\frac{9}{0}$	$MT/\frac{9}{0}$	$ML/\sqrt[9]{0}$	Tracking time/s			
SiameseCNN	Offline	41.51	65.60	12.25	24.20	0.81			
CNNTCM	Offline	51.42	69.80	22.73	22.66	0.73			
DCO-X	Offline	62.00	71.43	32.51	23.26	0.96			
LP-SSVM	Offline	72.60	71.52	44.35	25.35	0.16			
NOMTHM	Online	64.10	63.12	33.45	22.27	0.09			
STAM	Online	73.25	65.37	26.35	20.20	0.24			
SSP	Online	79.24	77.65	66.36	10.25	0.64			
MOTBeyondPixels	Online	81.17	82.62	73.05	2.77	0.30			
DeepSORT	Online	83.22	85.45	77.12	3.22	0.25			
Ours	Online	85.65	87.55	80.52	3.27	0.16			

表 2 在装甲目标跟踪数据集中与几种成熟方法的跟踪测试比较

4 结 论

多装甲目标跟踪往往面临着环境复杂、目标尺 寸小、目标遮挡和目标之间轨迹交叉等问题,因此, 提出了一种基于视觉注意力的 Gabor 在线多装甲目 标跟踪方法。

通过对比实验,结果表明:本方法在 MOTA 和 MOTP 方面取得了显著的效果,实现了目标跟踪率 的上升和目标丢失率的下降,并且总的目标跟踪运 行时间也能够满足实际需求。

参考文献:

- [1] SUN H Z, CHANG T Q, ZHANG L, et al. Armored target detection in battlefield environment based on top-down aggregation network and hierarchical scale optimization
 [J]. International Journal Pattern Recognition and Artificial Intelligence, 2019, 33(4): 312-334.
- [2] ZHAO D W,FU H,XIAO L,et al. Multi-object tracking with correlation filter for autonomous vehicle [J]. Sensors, 2018,18(7):2004.
- [3] QI Y K,QIN L,ZHANG S P,et al. Robust visual tracking via scale-and-state-awareness[J]. Neurocomputing, 2019, 329(15):75-85.
- [4] QI Y K, ZHANG S P, QIN L, et al. Hedging deep features for visual tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 41(5):1116-1130.
- [5] QI Y K, ZHANG S P, ZHANG W G, et al. Learning attribute-specific representations for visual tracking[C]//33th AAAI Conference on Artificial Intelligence, January 27-February 1, 2019, Honolulu, Hawaii, United States. New York: AAAI, 2019, 33(1):8835-8842.
- [6] ZHANG S P,LAN X Y,QI Y K,et al. Robust visual tracking via basis matching[J]. IEEE Transactions on Circuits

System for Video Technology, 2017, 27(3): 421-430.

- [7] CHU Q, OUYANG W L, LI H S, et al. Online multi-object tracking using CNN-based single object tracker with spatial-temporal attention mechanism[C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 4846-4855.
- [8] HENSCHEL R,LAURA L T, CREMERS D, et al. Fusion of head and full-body detectors for multi-object tracking
 [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, June 18-22, 2018, Seattle,USA. New York: IEEE, 2018:1509-1518.
- [9] ZHANG Y P, TANG Y Y, FANG B, et al. Multi-object tracking using deformable convolution networks with tracklets updating[J]. International Journal of Wavelets, Multiresolution and Information Processing, 2019, 11(6):1950042.
- [10] DING M Y, CAI J, ZHOU M, et al. Research on discriminative pedestrian single target tracking algorithm with adaptive tracking state[J]. Journal of Optoelectronics • Laser, 2022, 33(9):940-947.

丁明远,蔡靖,周冕,等.跟踪状态自适应的判别式行人 单目标跟踪算法研究[J].光电子·激光,2022,33(9): 940-947.

- [11] RIAZ F,HASSAN A,REHMAN S,et al. Texture classification using rotation and scale-invariant gabor texture features[J]. IEEE Signal Processing Letters, 2019, 20(6): 607-610.
- WU D, LI T, WAN Q. Salient object region detection based on background-bias prior and center-bias prior
 J. Journal of Optoelectronics • Laser, 2022, 33 (8): 799-806.

吴迪,李婷,万琴.基于背景先验与中心先验的显著性 目标检测[J].光电子·激光,2022,33(8):799-806.

[13] LIU Q K, CHEN D D, CHU Qi, et al. Online multi-object

• 1186 •

tracking with unsupervised re-identification learning and occlusion estimation [J]. Neurocomputing, 2022, 483(4): 333-347.

- [14] NAIEMI F, GHODS V, KHALESI H. An efficient character recognition method using enhanced hog for spam image detection [J]. Soft Computing, 2019, 23 (1): 11759-11774.
- YU Q Z, ZHOU S, JIANG Y, et al. High-performance sar image matching using improved sift framework based on rolling guidance filter and roewa-powered feature [J].
 IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2019, 12(3):920-933.
- [16] LIMA M D D, COSTA N, BARBOSA R M. Improvements on least squares twin multi-class classification support vector machine [J]. Neurocomputing, 2018, 313: 196-205.
- [17] FANG M S, HUANG Y R, HAN T. Remote sensing image detection algorithm based on selective fusion of multiscale features [J]. Journal of Optoelectronics • Laser, 2022, 33(6):629-636.

方明帅,黄友锐,韩涛.基于多尺度特征选择性融合的 遥感图像检测算法[J].光电子·激光,2022,33(6): 629-636.

[18] ANG S Y, DENG D S, ZHENG Q C. Ground-glass opacity target detection in CT scans based on improved Faster R-CNN model[J]. Journal of Image and Graphics, 2021, 26 (9):2171-2180.

杨淑莹,邓东升,郑清春.改进Faster R-CNN 模型的 CT 图磨玻璃密度影目标检测[J].中国图象图形学报,

2021,26(9):2171-2180.

- [19] RAO L Z, WANG J H, ZHENG X, et al. Monocular vision based method for target tracking of USV[J]. Transducer and Microsystem Technologies, 2022, 41(7):44-47+51.
 饶六中,王建华,郑翔,等.基于单目视觉的无人水面艇 目标跟踪方法[J]. 传感器与微系统, 2022, 41(7):44-47 +51.
- [20] ZHONG Z, SUN L, HUO Q. An anchor-free region proposal network for Faster R-CNN-based text detection approaches[J]. International Journal on Document Analysis and Recognition, 2019, 22(3): 315-327.
- [21] SHEN L F, LI Y. Design of target tracking control system for inspection robot based on narrowband intenet of things[J]. Computer Measurement & Control, 2022, 30 (4):98-102.
 申丽芳,李莹.基于窄带物联网的巡检机器人目标跟踪 控制系统设计[J].计算机测量与控制,2022,30(4):98-

102.
[22] YANG B, WANG X H. Target tracking based on adaptive block extraction[J]. Computer Engineering and Design, 2022,43(6):1719-1724.
杨波,王小虎.自适应分块优化的目标跟踪算法[J].计

算机工程与设计,2022,43(6):1719-1724.

作者简介:

王慧敏 (1988-),男,硕士,讲师,主要从事武器运用工程方面的 研究.