文章编号：1003－1251(2024)06－0020－08

基于改进 YOLOv7-DeepSort 的红外视频多目标跟踪

宫 华[1a,2]，张众垚[1b]，胡雨桐[3]，刘 芳[1a,2]

(1. 沈阳理工大学 a. 理学院，b. 自动化与电气工程学院，沈阳 110159；2. 辽宁省兵器工业智能优化与控制重点实验室，沈阳 110159；3. 电磁空间安全国家重点实验室，天津 300308)

摘 要：针对红外图像纹理弱及多目标遮挡导致跟踪精度低的问题，构建了基于改进 YOLOv7 模型和多目标跟踪算法 DeepSort 的融合红外目标跟踪模型 MSB-YOLOv7-DeepSort。采用 SE(squeeze and excitation)通道注意力机制和双向特征金字塔网络提高红外目标的特征提取质量；利用轻量化网络 MobileNetV3 替换 YOLOv7 骨干网络，提升融合模型的推理速度。实验结果表明，MSB-YOLOv7-DeepSort 模型在跟踪准确度、跟踪精确度、正确目标跟踪比例和帧率等方面均具有较好的性能。

关 键 词：红外目标跟踪；YOLOv7；轻量化；SE 注意力机制；MobileNetV3；双向特征金字塔网络

中图分类号：TP391.4 **文献标志码**：A **DOI**：10.3969/j.issn.1003－1251.2024.06.004

Infrared Video Multi Target Tracking Based on Improved YOLOv7-DeepSort

GONG Hua[1,2]，ZHANG Zhongyao[1]，HU Yutong[3]，LIU Fang[1,2]

(1. Shenyang Ligong University，Shenyang 110159，China；
2. Liaoning Key Laboratory of Intelligent Optimization and Control for Ordnance Industry，shenyang 110159，China；
3. National Key Laboratory of Electromagnetic Space Security，Tianjin 300308，China)

Abstract：To solve the problem of low tracking accuracy caused by weak texture and multi target occlusion，a fusion infrared tracking model MSB-YOLOv7-DeepSort is constructed，based on the improved YOLOv7 and the multi-target tracking algorithm DeepSort. The SE(squeeze and excitation)channel attention mechanism and the bidirectional feature pyramid network are utilized to improve the quality of feature extraction for the infrared targets. The lightweight network MobileNetV3 is used to replace the YOLOv7 backbone network to enhance the inference speed of the fusion model. The experimental results indicate that the MSB-YOLOv7-DeepSort model has good performance in tracking accuracy and frame rate.

Key words：infrared target tracking；YOLOv7；lightweight；SE attention mechanism；MobileNetV3；bidirectional feature pyramid network

红外视频目标跟踪是指通过计算机视觉和图像处理技术，对红外视频序列中的特定目标进行持续定位与跟踪。红外目标跟踪具有全天候、抗干扰的特点，能够有效区分复杂场景的目标信息与背景信息，在制导目标检测、军事侦察、海上救援等领域具有广泛应用[1－2]。相比于可见光视频，红外图像分辨率低、纹理弱，在场景复杂、目标数量不确定、存在遮挡和相似物干扰等情况下实

收稿日期：2023－12－22

基金项目：军委科技委重点实验室预研基金项目(2022JCJQLB055008)

作者简介：宫华(1976—)，女，教授，博士；胡雨桐(1994—)，通信作者，女，工程师。

现快速而精确的红外视频目标跟踪是军事领域和计算机视觉研究的一个热点问题。

视频多目标跟踪算法主要分为相关滤波和深度学习两类。相关滤波跟踪算法将信号处理中“相关性”概念引入跟踪领域,通过设计的滤波器与目标候选区域进行相关性运算,获得当前帧目标位置。常用的滤波器包括卡尔曼滤波器、粒子滤波器等[3]。相关滤波跟踪算法对目标比例变化具有较好的敏感性,但在多目标高速运动场景中目标匹配效果表现不佳。基于深度学习的跟踪算法按照学习特征不同主要分为基于深度表观特征、基于相似度度量和基于高阶匹配特征的多目标跟踪,主要包括目标检测、特征提取和数据关联三个阶段。该算法通过提取前后帧检测的目标特征,建立当前帧检测目标特征与已有目标特征的关联关系与匹配度,根据建立的相似度矩阵,利用代价最小化算法进行目标分配,实现目标的精准跟踪。

随着 R-CNN 系列和 YOLO 系列等深度网络的发展,基于深度学习的目标跟踪方法能够学习目标高级语义特征和提取目标位置信息,对复杂场景和多目标情况具有较好的适应性。目前主流的深度学习多目标跟踪算法基于目标检测结果进行目标跟踪,其中最经典的是 DeepSort 算法,其将深度学习模型引入 Sort 目标跟踪,该类算法需要解决的关键问题是实时过程中提高多目标跟踪推理速度和目标特征表征质量。在提高跟踪目标表征质量方面:郑繁亭等[4]将嵌入 ECANet 注意力机制的 YOLOv4 和使用 GhoustNetV1 网络的 DeepSort 跟踪算法相结合,提高了行人重识别网络的跟踪性能;蒲玲玲等[5]在 YOLOv5 网络中添加 ReID 特征识别模块,与 DeepSort 算法相结合,有效提升了车辆目标的跟踪精度。在提高目标跟踪推理速度方面:毛昭勇等[6]将 YOLOv3 骨干网络替换为 EfficientNet 轻量级网络,与 DeepSort 跟踪算法相结合,有效提升了对监控视频中目标跟踪的推理效率;武明虎等[7]在 YOLOv3 中引入 DPIoU 损失函数,改变 ResNet 下采样区域,与 Sort 跟踪算法相融合,有效提升了行人目标跟踪速度。

综上所述,特征融合、注意力机制和网络轻量化能够有效提升目标跟踪性能。YOLO 系列模型在特征提取和目标检测方面具有优势,与多目标跟踪算法 DeepSort 结合时,能够提升跟踪速度和跟踪精度。

目前关于视频目标跟踪的相关研究主要集中在可见光模态,鲜有红外视频目标跟踪的研究。针对红外视频中图像信噪比低、目标遮挡影响跟踪效果的问题,本文提出检测和跟踪两阶段融合的多目标跟踪模型 MSB-YOLOv7-DeepSort。在 YOLOv7 中引入 SE(squeeze and excitation)注意力机制[8],自适应地学习不同通道特征表征的重要性,增强模型的特征提取能力;采用 MobileNetV3网络对 YOLOv7 主干网络进行轻量化处理,在减少网络参数的同时,解决因网络层数较深导致红外目标特征损失问题;引入双向特征金字塔网络(bidirectional feature pyramid network,BiFPN)对 YOLOv7 网络提取的高层语义和低层空间位置信息进行双向融合,提升红外目标检测效果;将改进的检测模型(以 MSB-YOLOv7 表示)与 DeepSort 跟踪算法进行结合,提高红外视频目标的跟踪效果。

1　基于 MSB-YOLOv7-DeepSort 的红外视频跟踪模型

本文设计检测和跟踪两阶段算法对红外视频目标进行跟踪。将 SE 注意力机制、网络轻量化策略 MobileNetV3、双向特征融合策略 BiFPN 引入 YOLOv7 模型,提出红外视频目标检测模型 MSB-YOLOv7;基于检测结果,结合 DeepSort 多目标跟踪算法,建立红外视频多目标跟踪模型 MSB-YOLOv7-DeepSort。

1.1　MSB-YOLOv7 红外目标检测模型

YOLOv7 是 YOLO 系列算法中检测速度较快、计算精度较高的一种实时物体目标检测算法,其网络结构包括骨干网络(Backbone)和头部网络(Head)两个模块,骨干网络用于提取目标特征,头部网络用于融合不同语义的目标信息。为提高图像特征的表征能力和特征提取速度,本文以 YOLOv7 作为基准检测模型,通过引入注意力机制、轻量化网络和双向特征融合策略进行改进,提升网络的检测性能。本文提出的 MSB-YOLOv7 红外目标检测模型网络结构如图 1 所示。图中:P3、P4、P5 分别表示骨干网络中的三个下采样层;SPPCSPC 表示下采样模块;CBS、CBH 表示卷积模块;Upsample 表示上采样模块;ELAN、ELAN-H 表示聚合网络模块;MP1、MP2 表示最大池化模块;Concat 表示维度连接模块;REP 表示推理模块;CONV 表示卷积层。

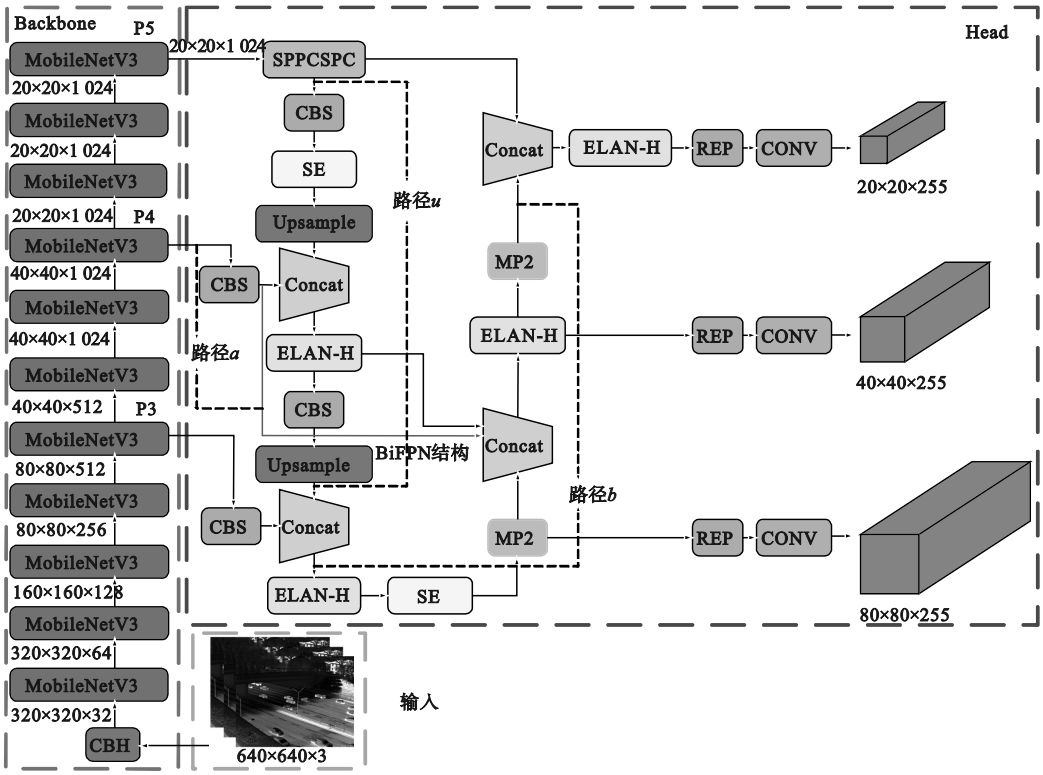


图 1 MSB-YOLOv7 模型网络结构

Fig. 1 Network architecture of MSB-YOLOv7 model

1.1.1 MobileNetV3 模型

MobileNetV3 是基于深度可分离卷积和反向残差思想建立的一种轻量级神经网络模型[9],其网络结构如图 2 所示。图中:Dwise 表示深度卷积;NL 表示非线性激活函数;FC 表示全连接层;Pool 表示池化层;hard-α 表示 NL 激活函数的“hard”形式;ReLU 表示线性修正单元激活函数。

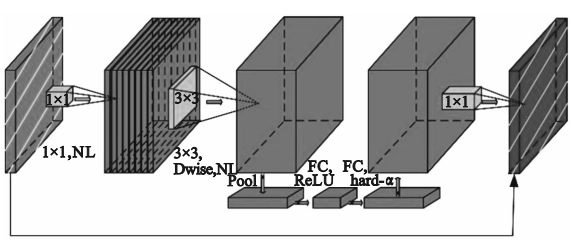


图 2 MobileNetV3 网络结构

Fig. 2 MobileNetV3 network architecture

MobileNetV3 中采用的深度可分离卷积参数量与标准卷积参数量的比值 M 为

$$M=\frac{D_{\mathrm{w}}\times D_{\mathrm{h}}\times C+C\times N}{D_{\mathrm{w}}\times D_{\mathrm{h}}\times C\times N}=\frac{1}{N}+\frac{1}{D_{\mathrm{w}}\times D_{\mathrm{h}}} \quad (1)$$

式中:D_{w}、D_{h} 和 C 分别为卷积核的宽、高和通道数;N 表示卷积核的数量。由式(1)可知,深度可分离卷积的参数量小于标准卷积。以 3×3 卷积核为例,D_{w} 和 D_{h} 分别为 3,通常 N 的数量较大,如 $N=64$,则 $1/N=1/64$,参数量比值 M 约为 1/9。本文将 YOLOv7 原始骨干网络中的 CBS 模块、ELAN 模块和 MP1 模块全部替换为 MobileNetV3 模块,在不改变输出图像分辨率和通道数的前提下降低模型运算的参数量,有效保留原始细节特征。

1.1.2 SE 注意力机制

SE 通道注意力机制是对全局空间信息进行压缩,通过通道维度的特征学习分配不同权重,增强网络对不同通道的注意力[10],其结构如图 3 所示。图中:$\boldsymbol{X}$ 表示原始输入特征;$\boldsymbol{U}$ 表示经过卷积操作后的特征;$\tilde{\boldsymbol{X}}$ 表示输出特征;H 和 H' 分别表示对应特征图的长度;W 和 W' 分别表示对应特征图的宽度;C 和 C' 分别表示对应特征图的通道数;$F_{\mathrm{sq}}(\cdot)$ 表示 Squeeze 模块;$F_{\mathrm{ex}}(\cdot,W)$ 表示 Excitation 模块;F_{tr} 表示输入;$F_{\mathrm{scale}}(\cdot,\cdot)$ 表示权重赋值。

YOLOv7 头部网络输入的特征信息经过下采样模块SPPCSPC 后的卷积操作会降低通道数,导致红外目标特征信息减少。为此,本文在 YOLOv7 算法的头部网络前部嵌入 SE 注意力模块,解决在卷积池化过程中不同通道占比程度相

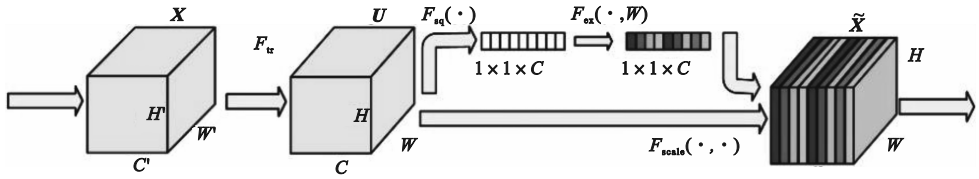


图 3　SE 注意力机制结构

Fig. 3　Architecture of SE attention mechanism

同导致红外目标特征不显著的问题;同时,在头部网络末端添加 SE 注意力模块,分配通道注意力,抑制无关信息,聚焦红外目标的有效特征信息,提升模型检测的准确率。

1.1.3　双向特征金字塔网络

针对红外视频目标存在的纹理信息弱、图像信噪比低等问题,本文采用 BiFPN 替换 YOLOv7 特征融合网络中的路径聚合网络 PANet。PANet 和BiFPN网络结构如图 4 所示。

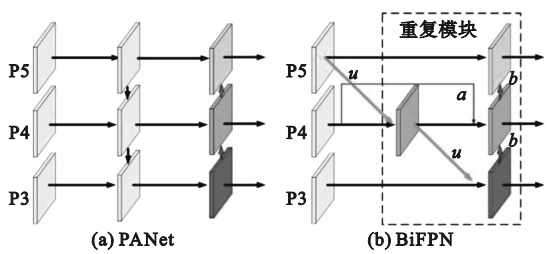


图 4　PANet 和 BiFPN 网络结构

Fig. 4　Network architecture of PANet and BiFPN

BiFPN 通过引入跨层连接和特征加权的方式,能够较好地平衡特征金字塔不同特征之间的关系,提高不同尺度目标的检测能力。针对背景复杂的红外目标检测问题,BiFPN 网络结构通过自上而下的路径 u 和自下而上的路径 b,将骨干网络输出的不同层级特征图进行多尺度特征融合,能够兼顾红外目标的局部特征和全局特征信息,增强红外图像中纹理弱和类别划分模糊的目标特征表征质量;路径 a 表示输入点和输出点位于同一层时的附加路径,能够在不增加计算成本的情况下融合更多特征,有助于在复杂背景中准确地提取红外目标的特征信息。

1.2　DeepSort 多目标跟踪算法

DeepSort 算法[11]是在 Sort 目标跟踪算法基础上引入表观特征提取网络的多目标跟踪算法,其核心思想是使用递归的卡尔曼滤波和逐帧的匈牙利算法进行数据关联,在数据关联中加入深度外观特征匹配策略,将目标检测和目标跟踪两个任务相结合。在目标跟踪过程中保存目标遮挡前的外观特征信息,减少目标 ID 跳变;当遮挡场景结束后,提取遮挡前后目标的外观特征信息进行最近邻匹配,提高遮挡情况下的跟踪准确率。

DeepSort 算法的工作原理如图 5 所示。

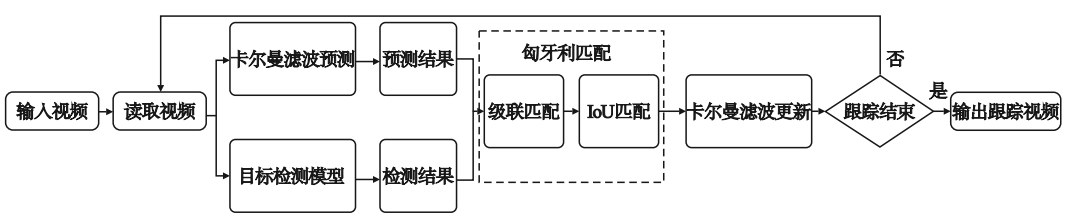


图 5　DeepSort 算法的工作原理

Fig. 5　Working principle diagram of DeepSort algorithm

首先,通过卡尔曼滤波预测目标的运动状态,获得目标的位置和状态信息,以(x,y,w,h,vx,vy,vw,vh)表示。其中:x、y 表示检测框的中心坐标;w、h 分别表示预测框的宽高比与高度;vx、vy、vw、vh 分别表征 x、y、w、h 各参量的相对变化速度。然后,结合运动信息和外观信息,运用马氏距离判断 MSB-YOLOv7 检测目标与卡尔曼滤波预测目标的关联度,以 $d^{(1)}(i,j)$ 表示第 i 个预测框和第 j 个检测框的马氏距离,计算式为

$$d^{(1)}(i,j)=(\boldsymbol{d}_j-\boldsymbol{y}_i)^{\mathrm{T}}\boldsymbol{S}_i^{-1}(\boldsymbol{d}_j-\boldsymbol{y}_i) \quad (2)$$

式中:$\boldsymbol{d}_j$ 表示第 j 个检测框的状态向量;$\boldsymbol{y}_i$ 表示第 i 个预测框的状态向量;$\boldsymbol{S}_i^{-1}$ 表示检测结果和跟踪结果的协方差逆矩阵。

当目标长时间被遮挡或视角发生位移时,需要引入外观信息,使用余弦距离解决遮挡带来的 ID 切换问题。以 $d^{(2)}(i,j)$ 表示第 i 个预测框和第 j 个检测框的余弦距离,计算式为

$$d^{(2)}(i,j)=\min\{1-\boldsymbol{r}_j^{\mathrm{T}}\boldsymbol{r}_k^{(i)}\mid\boldsymbol{r}_k^{(i)}\in\boldsymbol{R}_i\} \quad (3)$$

式中：$\boldsymbol{r}_j$ 为 $\boldsymbol{d}_j$ 的特征向量；$\boldsymbol{r}_k^{(i)}$ 表示跟踪器中保存的第 i 个特征向量；$\boldsymbol{R}_i$ 为外观特征向量库。

为实现马氏距离度量的短时位置信息和余弦距离表征的长时遮挡目标信息的优势互补，采用加权方式对上述两种距离求和，获得最终度量值 $c_{i,j}$，计算式为

$$c_{i,j}=\lambda d^{(1)}(i,j)+(1-\lambda)d^{(2)}(i,j) \quad (4)$$

式中 λ 为超参数。当且仅当 $c_{i,j}$ 介于 $d^{(1)}(i,j)$ 和 $d^{(2)}(i,j)$ 之间时认定目标关联。

DeepSort 跟踪算法对状态确定的目标通过前后两帧图片目标框之间的交并比(IoU)构建相似度矩阵，再通过匈牙利算法[12]进行匹配，完成卡尔曼滤波的更新。

1.3 MSB-YOLOv7-DeepSort 多目标跟踪模型

基于 MSB-YOLOv7-DeepSort 融合跟踪模型的跟踪流程如图 6 所示。

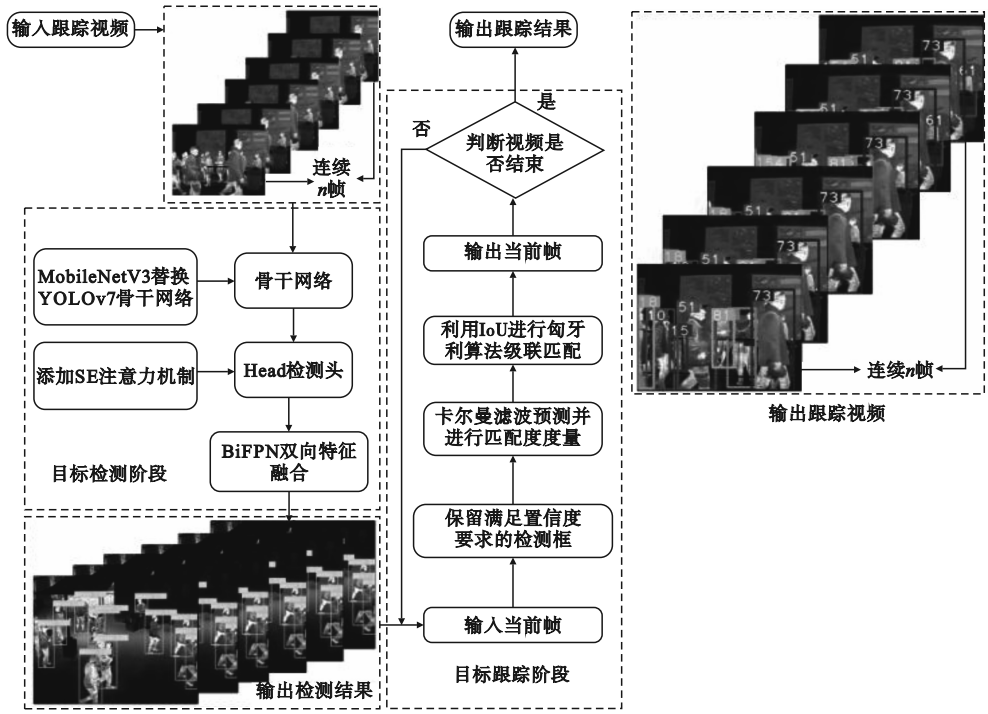


图 6 MSB-YOLOv7-DeepSort 模型跟踪流程

Fig. 6 Tracking processes of MSB-YOLOv7-DeepSort model

MSB-YOLOv7-DeepSort 模型的多目标跟踪算法步骤如下：

步骤 **1** 将多目标红外视频输入检测模型 MSB-YOLOv7；

步骤 **2** MSB-YOLOv7 模型输出视频帧的多目标检测框，并将其输入 DeepSort 跟踪算法；

步骤 **3** 根据上一帧的目标检测框信息，DeepSort 跟踪算法利用卡尔曼滤波算法预测当前帧目标的位置，输出目标的预测框；

步骤 **4** 对 MSB-YOLOv7 模型输出的目标检测框和卡尔曼滤波算法输出的目标预测框，运用匈牙利算法进行 IoU 匹配，输出满足匹配阈值的目标预测框；

步骤 **5** 输出匹配成功的当前帧跟踪结果；

步骤 **6** 判断输入视频是否结束，如未结束，循环执行步骤 2，如结束，则输出最终的跟踪结果。

2 实验结果与分析

针对本文提出的检测和跟踪两阶段红外视频目标跟踪算法，分别设计红外目标检测实验和红外目标跟踪实验，验证模型的有效性。

2.1 数据集

本文采用 CAMEL 数据集(官方下载网址为 https://camel.ece.gatech.edu/)中的红外视频序列集进行实验。数据集由检测实验数据集和跟踪实验数据集两部分构成。数据集包含 29 个红外视频序列，共 24 981 帧图片，图片分辨率为 336 × 256，其中包括行人、车辆、自行车和狗等 4 类目

标,涵盖公路、街道、室内等多种场景。检测实验数据集和跟踪实验数据集分别包含 23 个和 6 个红外视频序列,检测实验数据集共 21 895 张图片、118 003 个目标,数据集按照 8∶2 的比例划分为训练集和验证集。检测实验数据集参数如表 1 所示,跟踪实验数据集视频示例如图 7 所示。

表 1　检测实验数据集参数

Table 1　Parameters of detection dataset

	图片数量	目标数量			
		自行车	行人	车辆	狗
训练集	17 515	594	66 281	26 718	888
测试集	4 380	141	16 789	6 381	211
总计	21 895	735	83 070	33 099	1 099

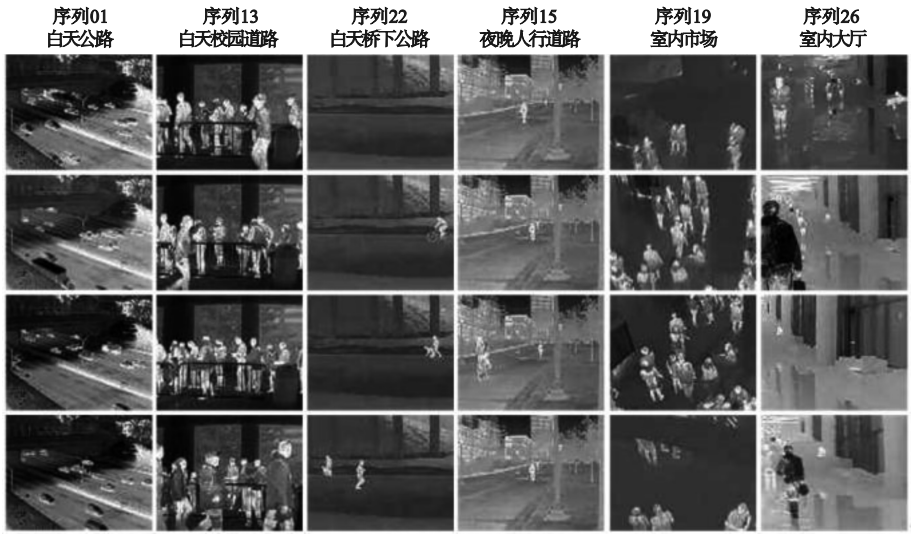


图 7　跟踪实验数据集视频示例

Fig. 7　Examples of tracking video dataset

2.2　实验环境与参数设置

实验操作系统为 Windows 11,平台配置为 Intel(R) Xeon(R) Platinum i9-13900k CPU @ 3.0 GHz,GPU 为 NVIDIA GeForce RTX 4090,显存 24 GB,采用 Pytorch 1.13.1 深度学习框架,算法基于 Python3.8 实现。

检测实验训练迭代次数设置为 300,训练的批次设置为 16,初始学习率设置为 0.01,初始图像尺寸设置为 640×640。跟踪实验训练迭代次数设置为 40,最大 IoU 匹配阈值设置为 0.7,检测目标最小置信度设置为 0.3。

2.3　评价指标

本文采用的检测实验评价指标分别为精确率、召回率、平均精度均值(mAP)、参数量和帧率(FPS)。跟踪实验评价指标分别为多目标跟踪准确度(MOTA)、多目标跟踪精确度(MOTP)、正确目标跟踪比例(IDF1)、目标 ID 切换次数(IDs)和 FPS。

2.4　MSB-YOLOv7 模型检测实验

为验证 MSB-YOLOv7 模型对红外图像中目标检测的有效性,将常用坐标注意力机制(coordinate attention,CA)、卷积模块注意力机制(convolutional block attention module,CBAM)、通道注意力机制 SE 分别嵌入 YOLOv7 模型,与采用 MobileNetV3 网络和 SE 注意力机制的 MS-YOLOv7 模型及 MSB-YOLOv7 模型进行对比实验,结果如表 2 所示。

表 2　检测模型对比实验结果

Table 2　Comparison of experiment results of detection models

模型	精确率/%	召回率/%	mAP/%	参数量/10^6	FPS
YOLOv7	98.1	96.9	98.6	36.519	27
CA-YOLOv7	98.7	95.7	98.6	36.526	16
CBAM-YOLOv7	98.9	95.1	98.6	36.524	17
SE-YOLOv7	97.1	97.1	98.8	36.521	29
MS-YOLOv7	98.4	96.0	99.1	34.563	29
MSB-YOLOv7	98.8	97.7	99.5	34.432	33

由表 2 可知,在六种对比模型的实验结果中,MSB-YOLOv7 模型的召回率、mAP 和参数量三个评价指标均获得最优值。相比原 YOLOv7 模型,MSB-YOLOv7 模型的精确率、召回率、mAP 和 FPS 分别提高了 0.7%、0.8%、0.9% 和 22.22%,参数量降低了 5.71%,模型检测的准确度和速度有效提升,模型参数量减少。各模型检测结果可视化如图 8 所示。

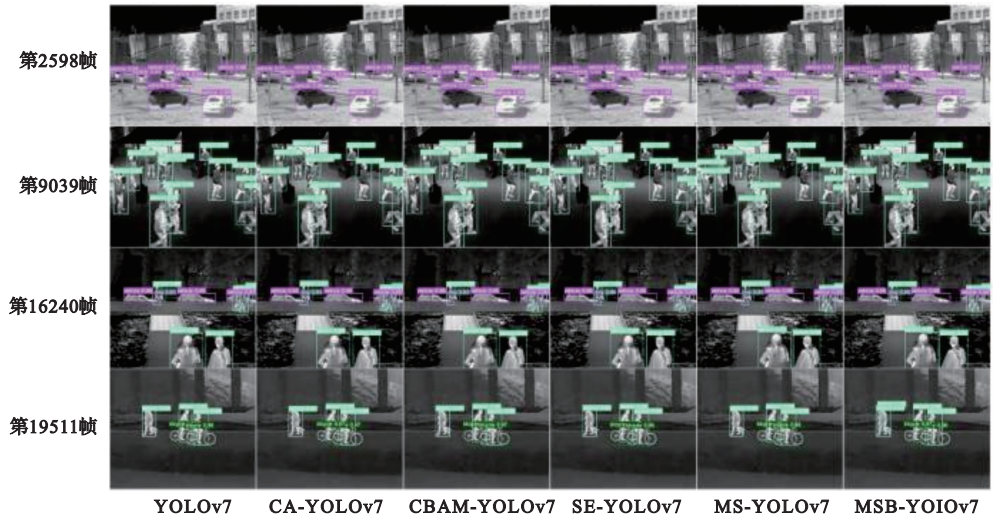


图 8 六种检测模型实验结果可视化

Fig. 8 Experiment results visualization of six detection models

2.5 MSB-YOLOv7-DeepSort 模型跟踪实验

为验证本文融合跟踪模型的跟踪有效性,将上述检测模型分别与 DeepSort 跟踪算法进行融合,进行对比实验,结果如表 3 所示。各模型跟踪结果可视化如图 9 所示。

由表 3 可知,与 YOLOv7-DeepSort 跟踪模型相比,SE-YOLOv7-DeepSort 跟踪模型的 MOTA、MOTP 、IDF1 和 FPS 分别提高了 3.62%、0.54%、3.81%和 44.62%,说明通过 SE 注意力机制对不同通道赋予不同权重的策略能够有效提升红外目标的检测效果。与SE-YOLOv7-DeepSort跟踪模

表 3 跟踪模型对比实验结果

Table 3 Comparison of experiment results of tracking models

模型	MOTA /%	MOTP /%	IDF1 /%	IDs	FPS
YOLOv7-DeepSort	36.38	75.38	35.61	176	13.85
CA-YOLOv7-DeepSort	35.27	75.24	34.21	184	14.13
CBAM-YOLOv7-DeepSort	38.58	75.30	35.31	184	20.32
SE-YOLOv7-DeepSort	40.00	75.92	39.42	200	20.03
MS-YOLOv7-DeepSort	43.81	77.39	36.24	147	20.82
MSB-YOLOv7-DeepSort	56.84	82.38	72.47	137	20.90

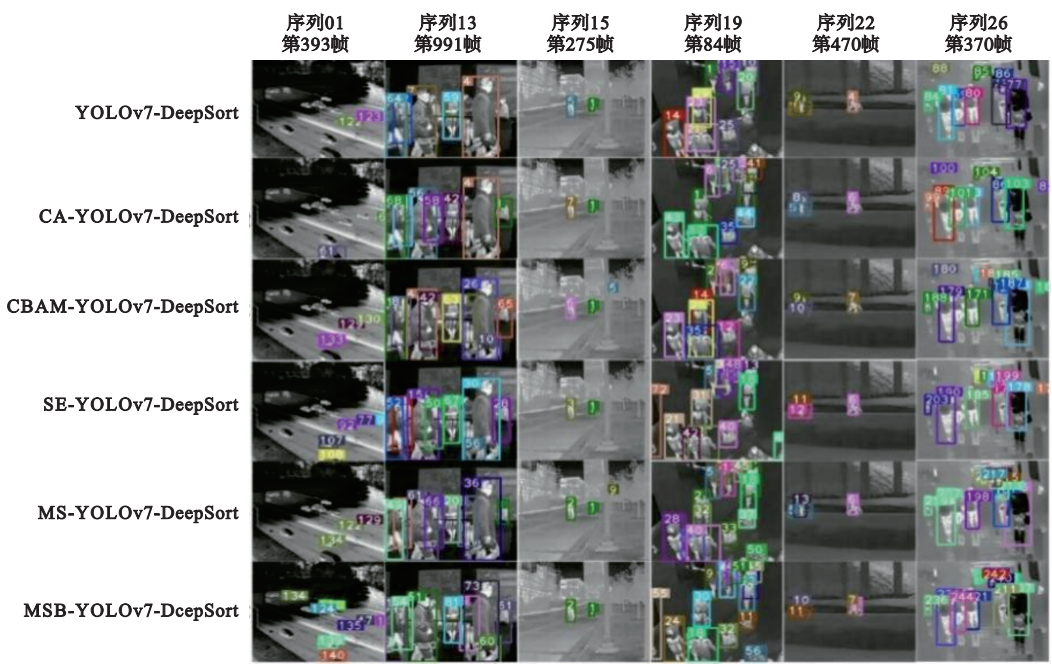


图 9 六种跟踪模型实验结果可视化

Fig. 9 Experiment results visualization of six tracking models

型相比,本文 MSB-YOLOv7-DeepSort 跟踪模型的 MOTA、MOTP、IDF1 和 FPS 分别提高了 16.84%、6.46%、33.05% 和 4.34%,IDs 降低了 31.5%,表明网络轻量化策略能够提取红外图像的细粒度信息,避免网络高深度导致红外特征信息减弱、消失的问题,提高跟踪的稳定性,同时双向特征融合策略能够进一步完善红外图像的特征信息,显著提升红外目标的跟踪效果。

由图 9 可见,MSB-YOLOv7-DeepSort 模型跟踪目标数量最多,未出现跟踪目标失配情况,验证了跟踪模型的有效性。

3　结论

针对红外视频中图像信噪比低、目标遮挡影响跟踪效果的问题,提出检测 - 跟踪两阶段的 MSB-YOLOv7-DeepSort 融合跟踪模型。通过采用通道注意力机制、轻量化网络和双向特征融合策略,有效提升了模型的特征提取质量,减少了参数量和计算量,提高了多目标检测速度与精度,实现了红外视频多目标的高质量跟踪。对比实验结果表明,MSB-YOLOv7-DeepSort 模型在 MOTA、MOTP、IDF1、IDs 和 FPS 等跟踪效果评价指标中均得到较大提升,可为红外视频目标跟踪提供有效算法和技术支持。

参考文献(References):

[1] 华宇宁,崔春娜. 一种基于粒子滤波的红外多目标跟踪算法[J]. 沈阳理工大学学报,2017,36(1):66 - 70.
HUA Y N,CUI C N. Infrared multi-target tracking algorithm based on particle filter[J]. Journal of Shenyang Ligong University,2017,36(1):66 - 70. (in Chinese)

[2] SOUSA E,VARDASCA R,TEIXEIRA S,et al. A review on the application of medical infrared thermal imaging in hands [J]. Infrared Physics & Technology,2017,85:315 - 323.

[3] BASATI A,FAGHIH M M. PDAE:efficient network intrusion detection in IoT using parallel deep auto-encoders[J]. Information Sciences,2022,598:57 - 74.

[4] 郑繁亭,邢关生. 基于改进 DeepSort 的行人多目标跟踪算法[J]. 现代电子技术,2023,46(5):40 - 46.
ZHENG F T,XING G S. Pedestrian multi-target tracking algorithm based on improved DeepSort[J]. Modern Electronics Technique,2023,46(5):40 - 46. (in Chinese)

[5] 蒲玲玲,杨柳. 改进 YOLOv5 的多车辆目标实时检测及跟踪算法[J]. 科学技术与工程,2023,23(28):12159 - 12167.
PU L L,YANG L. Improved real-time detection and tracking algorithm for multi vehicle targets in YOLOv5[J]. Science Technology and Engineering,2023,23(28):12159 - 12167. (in Chinese)

[6] 毛昭勇,王亦晨,王鑫,等. 面向高速公路的车辆视频监控分析系统[J]. 西安电子科技大学学报,2021,48(5):178 - 189.
MAO Z Y,WANG Y C,WANG X,et al. Vehicle video surveillance and analysis system for the expressway[J]. Journal of Xidian University,2021,48(5):178 - 189. (in Chinese)

[7] 武明虎,黄咏曦,王娟. 基于改进 YOLOv3 的街道行人检测与跟踪方法[J]. 科学技术与工程,2021,21(17):7230 - 7236.
WU M H,HUANG Y X,WANG J. Pedestrian detection and tracking method on street based on improved YOLOv3[J]. Science Technology and Engineering,2021,21(17):7230 - 7236. (in Chinese)

[8] SHI H W,YANG W Z,CHEN D N,et al. CPA-YOLOv7:contextual and pyramid attention-based improvement of YOLOv7 for drones scene target detection[J]. Journal of Visual Communication and Image Representation,2023,97(3):103965.

[9] 易子娟,贾渊. 基于改进 MobileNetV3 的色环电阻识别[J]. 计算机系统应用,2023,32(4):361 - 367.
YI Z J,JIA Y. Recognition of color-ring resistors based on improved MobileNetV3[J]. Computer Systems and Applications,2023,32(4):361 - 367. (in Chinese)

[10] LIU S,QI L,QIN H F,et al. Path aggregation network for instance segmentation[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City,UT,USA:IEEE,2018:8759 - 8768.

[11] VEERAMANI B,RAYMOND J W,CHANDA P. DeepSort:deep convolutional networks for sorting haploid maize seeds [J]. BMC Bioinformatics,2018,19(9):1 - 9.

[12] MA H W,WEI X R,WANG P,et al. Multi-arm global cooperative coal gangue sorting method based on improved Hungarian algorithm[J]. Sensors,2022,22(20):7987.

(责任编辑:宋颖韬)