

基于多分支和重参数化的孪生网络跟踪算法

金 铭, 唐 宇, 韩 勇, 刘 帅, 闫锋刚
(哈尔滨工业大学(威海)威海 264200)

摘要: 针对孪生网络对尺度变化目标特征表达能力不足的问题, 本文使用不同尺寸的卷积、池化分支和剪枝操作构成多分支结构, 以提高特征的鲁棒性并保证孪生网络的平移不变性。针对多分支结构带来参数量过多的问题, 本文在跟踪阶段将多分支结构重参数化为单一的卷积, 有效减少跟踪阶段时间成本。实验结果表明: 本文提出的算法相比于 SiamFC, 在 OTB100 数据集上, 其精度、成功率和跟踪速度分别提高了 5.1%、3%、30 FPS, 在 UAV123 和 Temple-Color-128 数据集上跟踪精度和成功率均有所提高。

关键词: 视觉跟踪; 孪生网络; 特征提取; 结构重参数化

中图分类号: TN911.73; TP391.4

文献标志码: A

文章编号: 2095-1000(2024)04-0022-09

DOI: 10.12347/j.ycyk.20240408001

引用格式: 金铭, 唐宇, 韩勇, 等. 基于多分支和重参数化的孪生网络跟踪算法[J]. 遥测遥控, 2024, 45(4): 22-30.

Siamese Network Tracking Algorithm Based on Diverse Branch Block and Reparameterization

JIN Ming, TANG Yu, HAN Yong, LIU Shuai, YAN Feng Gang
(Harbin Institute of Technology, Weihai 264200, China)

Abstract: Aiming at the problem that the Siamese network has insufficient ability to express the features of scale-varying targets, a multi-branch structure is constructed by using convolution, pooling branches and pruning operations of different sizes to improve the robustness of features and ensure the translation invariance of the Siamese network. Aiming at the problem that the multi-branch structure brings too many parameters, the multi-branch structure is reparameterized into a single convolution in the tracking stage, which effectively reduces the time cost in the tracking stage. The experimental results show that compared with SiamFC, the accuracy, success rate and tracking speed of the proposed algorithm on the OTB100 datasets are improved by 5.1%, 3% and 30 FPS, respectively. The tracking accuracy and success rate are improved on the UAV123 and Temple-Color-128 datasets.

Keywords: Visual tracking; Siamese network; Feature extraction; Structural reparameterization

Citation: JIN Ming, TANG Yu, HAN Yong, et al. Siamese Network Tracking Algorithm Based on Diverse Branch Block and Reparameterization[J]. Journal of Telemetry, Tracking and Command, 2024, 45(4): 22-30.

0 引言

目标跟踪作为计算机视觉领域富有挑战性的研究方向之一, 是高层计算机视觉任务的基石^[1], 在公共安防、无人驾驶、人机交互等领域有着广泛的应用。但在复杂的现实环境下, 要实现实时且准确地跟踪, 仍需要进一步深入研究。

近年来, 基于相关滤波的跟踪算法和基于孪生网络的跟踪算法是目标跟踪领域中两个主要研

究方向。前者大多采用单一特征或融合特征^[2-5], 但相较于孪生网络跟踪算法所使用的深度特征, 以上特征鲁棒性较差, 从而导致跟踪器的性能不佳。对此, 研究人员在边界效应^[6]、尺度估计^[7]等方面进行一系列的研究以提高跟踪器性能。部分研究发现, 相关滤波算法结合深度特征进行跟踪时, 跟踪速度会大幅下降, 难以满足实时性。而孪生网络跟踪算法在保证实时性的同时取得了良好的跟踪精度, 因此基于孪生网络的跟踪框架逐

渐成为研究的主流。

SiamFC是首个孪生网络跟踪算法^[8]，针对尺度变化问题，它通过搜索多个比例的搜索图像以适应跟踪目标，但带来了成倍的计算量。对此Li等人引入区域推荐方法^[9]，提高了模型对不同尺度目标的适应能力，但该方法会引入尺度的先验和超参数，影响模型的泛化性。对此GUO、PENG、CHEN等人采用分类子网络直接预测目标中心点与边界的距离^[10-12]。为减少跟踪阶段的计算负担并避免目标模板被污染，SiamFC仅采用第一帧目标信息作为跟踪模板，后续不执行任何的模型更新，但这限制了算法的性能，当目标发生形变、遮挡时，可能发生跟踪漂移。对此，David采用线性更新策略^[13]、Feng采用记忆模型^[14]、Yang^[15]和Zhang^[16]采用边框回归等方法对模型进行更新，提高了跟踪性能。SiamFC采用AlexNet作为骨干网络，该网络不能提取表达能力非常强的目标特征表示，但针对该问题的研究较少，主要集中在增加骨干网络的深度^[17]。该方法在提高跟踪精度的同时增加了计算量，影响跟踪速度，并且需要在超大规模的数据集上进行训练，所需算力和时间成本高。不同于此前研究，本文针对在扩展骨干网络宽度的同时不增加参数这一问题，从多分支结构出发，基于训练和跟踪阶段可分离性质，提出了一种不增加跟踪阶段额外时间成本的多分支结构，通过应用该结构来优化骨干网络，提高跟踪器对目标特征的表达能，在多个测试数据集上测试可知，相比于基准算法，本文提出的算法在跟踪精度、成功率、速度上均有较大幅度提升。

1 基于多分支和重参数化的孪生网络算法

1.1 孪生网络与多分支结构

孪生网络将跟踪问题当作相似性学习问题，它的骨干网络提取到的深度特征会通过度量函数 f 得到模板图像与搜索图像的相似程度，如式(1)所示，其中 z 表示模板分支的输入， x 表示搜索分支的输入， $\varphi(\cdot)$ 作为骨干网络提取图像的特征表示， $*$ 表示互相关运算， b 是一个偏置项。

$$f(z, x) = \varphi(z) * \varphi(x) + b \quad (1)$$

虽然孪生网络取得了不错的跟踪性能，但其骨干网络为单一的卷积，在宽度上仍有改进空间。前人在图像分类任务中，通过使用多分支结构搭建的骨干网络提取的特征能更好地应对尺度变化

问题^[18]，但将相同的多分支结构迁移到孪生网络中反而使跟踪性能出现下降。一方面是因为多分支结构不关注环境信息，另一方面则是因为其破坏了孪生网络的平移不变性。对此，本文提出了一种适用于孪生网络的新型多分支结构。

1.2 适用于孪生网络的新型多分支结构

新型多分支结构如图1所示，为了利用环境信息对目标进行定位，新型多分支结构中包含了平均池化分支，以提高跟踪器在目标尺度等特征发生变化时的精确度和成功率。为了使每条分支输出的特征图得以融合， $K \times K$ 卷积需要进行填充操作，而填充操作会破坏全卷积孪生网络的平移不变性，导致目标容易发生跟踪漂移现象。对此，新型多分支结构在特征融合后通过对特征进行剪枝操作，去除受填充影响的特征，需要剪枝的大小为 $K/2$ ，剪枝后的特征可以保证孪生网络的平移不变性。同时，为了提高跟踪器对不同尺度目标的特征表达能力，新型多分支结构使用不同尺寸的卷积，并且多次使用 1×1 卷积进行跨通道信息融合，使输出的特征中既关注了全局特征也关注了局部特征。但多分支会引入过多参数，因此需要对其进行重参数化以压缩参数量。

1.3 重参数化新型多分支结构

为降低跟踪阶段时孪生网络的参数量，基于训练和跟踪可分离的性质，在跟踪阶段将训练阶段的新型多分支结构重参数化为单一卷积结构，如图2所示，具体方法如下。

1.3.1 卷积连接归一化层重参数化为单一卷积

卷积核是四阶张量，具有 C 个输入通道、 D 个输出通道和核尺寸为 K 的卷积核表示为 $F \in \mathbb{R}^{D \times C \times K \times K}$ ，可选的偏置表示为 $b \in \mathbb{R}^D$ ，这样的卷积核以 C 通道的特征图 $I \in \mathbb{R}^{C \times H \times W}$ 作为输入，输出 D 通道的特征图 $O \in \mathbb{R}^{D \times H' \times W'}$ ，其中 H' 和 W' 由 K 、填充和步长配置确定。将偏置复制到 $\text{REP}(b) \in \mathbb{R}^{D \times H' \times W'}$ ，并将其加到卷积结果上。卷积公式用下式表示：

$$O = I * F + \text{REP}(b) \quad (2)$$

卷积连接归一化层的输出可以用下式表示：

$$O_{j,\dots} = ((I * F)_{j,\dots} - \mu_j) \frac{\gamma_j}{\sigma_j} + \beta_j \quad (3)$$

其中， j 为输出通道的索引， μ_j 是累积的通道平均值， σ_j 是累积的通道标准差， γ_j 是要学习的缩放因

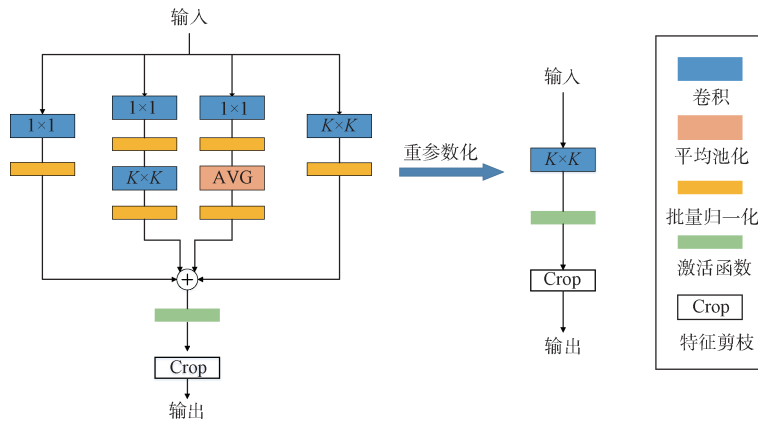


图 1 新型多分支结构

Fig. 1 New diverse branch block

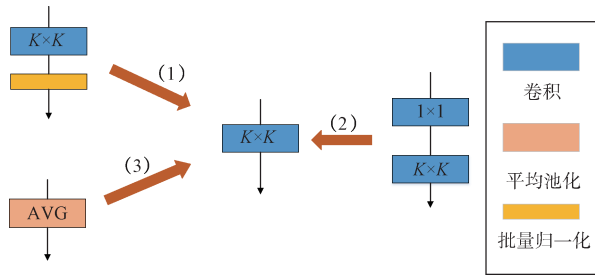


图 2 三种结构重参数化为单一卷积

Fig. 2 The three structures are reparameterized into a single convolution

子, β_j 是要学习的偏置项。由上式, 用卷积核 F' 和偏置 b' 表示重参数化后的卷积。

$$F'_{j,\dots} \leftarrow \frac{\gamma_j}{\sigma_j} F_{j,\dots}, \quad b'_j \leftarrow -\frac{\mu_j \gamma_j}{\sigma_j} + \beta_j \quad (4)$$

1.3.2 序列卷积重参数化为单一卷积

假设 1×1 卷积的输入和输出通道分别为 C 、 D , $K \times K$ 卷积的输入和输出通道分别为 D 、 E , 使用式(4)将卷积连接归一化层重参数化得到 $F^{(1)} \in \mathbb{R}^{D \times C \times 1 \times 1}$, $b^{(1)} \in \mathbb{R}^D$, $F^{(2)} \in \mathbb{R}^{E \times D \times K \times K}$, $b^{(2)} \in \mathbb{R}^E$, 输出可以表示为:

$$O' = I * F^{(1)} * F^{(2)} + \text{REP}(b^{(1)}) * F^{(2)} + \text{REP}(b^{(2)}) \quad (5)$$

由于 $I * F^{(1)}$ 是 1×1 卷积, 它只进行通道线性组合而不进行空间聚合, 因此可以通过线性重组将其参数融入进 $K \times K$ 核中, 即:

$$F' \leftarrow F^{(2)} * \text{TRANS}(F^{(1)}) \quad (6)$$

$\text{TRANS}(F^{(1)}) \in \mathbb{R}^{C \times D \times 1 \times 1}$ 是 $F^{(1)}$ 的转置。式(5)的第二项是常数矩阵上的卷积, 因此输出也是常数矩阵。

据此, 可以构建 \hat{b} :

$$\hat{b}_j \leftarrow \sum_{d=1}^D \sum_{u=1}^K \sum_{v=1}^K b_d^{(1)} F_{j,d,u,v}^{(2)}, \quad 1 \leq j \leq E \quad (7)$$

从而使

$$\text{REP}(b^{(1)}) * F^{(2)} = \text{REP}(\hat{b}) \quad (8)$$

$$b' \leftarrow \hat{b} + b^{(2)} \quad (9)$$

1.3.3 平均池化重参数化为单一卷积

对于 C 个通道, 卷积核尺寸为 K 、步长为 s 的平均池化相当于 K 和 s 相同的卷积。这样的卷积 $F' \in \mathbb{R}^{C \times C \times K \times K}$ 可以由下式构造:

$$F'_{d,c,\dots} = \begin{cases} \frac{1}{K^2} & \text{若 } d=c \\ 0 & \text{其它} \end{cases} \quad (10)$$

综上, 对于新型多分支结构中的分支 1 和分支 4, 使用式(4)即可重参数化为单一卷积, 分别表示为 $F^{(1)}$ 、 $b^{(1)}$ 和 $F^{(4)}$ 、 $b^{(4)}$ 。对于分支 2, 使用式(6)与(9)即可重参数化为单一卷积分支, 表示为 $F^{(2)}$ 、 $b^{(2)}$ 。对于分支 3, 使用式(10)将平均池化重参数化为单一卷积后与分支 2 操作一致, 分支 3 表示为 $F^{(3)}$ 、 $b^{(3)}$ 。基于卷积的可加性, 将以上四个卷积重参数化为单一卷积:

$$F' \leftarrow F^{(1)} + F^{(2)} + F^{(3)} + F^{(4)} \quad (11)$$

$$b' \leftarrow b^{(1)} + b^{(2)} + b^{(3)} + b^{(4)} \quad (12)$$

1.4 改进的骨干网络及算法整体流程

以 SiamFC 为基准, 对其骨干网络进行改进, 将它的 5×5 卷积拆分成两个 3×3 卷积来降低参数量, 并将除最后一层外的 3×3 卷积替换为新型多分支结构, 提高对尺度变化目标的特征表达能力。训练阶段的骨干网络结构如表 1 所示, 表中 DBB 表示新型多分支结构, AVG 表示平均池化。在跟

踪阶段通过重参数化压缩网络参数量, 具体结构 如表2所示。

表1 训练阶段的骨干网络结构

Table 1 The backbone network structure of the training phase

卷积层	卷积结构	步长	通道数	模板尺寸	搜索尺寸
			3	127×127	255×255
Conv1	11×11	2	96	59×59	123×123
Pool1	3×3	2	96	29×29	61×61
DBB1	1×1, 1×1-3×3, 1×1-AVG, 3×3	1	256	27×27	59×59
DBB2	1×1, 1×1-3×3, 1×1-AVG, 3×3	1	256	25×25	57×57
Pool2	3×3	2	256	12×12	28×28
DBB3	1×1, 1×1-3×3, 1×1-AVG, 3×3	1	384	10×10	26×26
DBB4	1×1, 1×1-3×3, 1×1-AVG, 3×3	1	384	8×8	24×24
Conv2	3×3	1	256	6×6	22×22

表2 跟踪阶段的骨干网络结构

Table 2 The backbone network structure of the tracking phase

卷积层	卷积结构	步长	通道数	模板尺寸	搜索尺寸
			3	127×127	255×255
Conv1	11×11	2	96	59×59	123×123
Pool1	3×3	2	96	29×29	62×61
Conv2	3×3	1	256	27×27	59×59
Conv3	3×3	1	256	25×25	57×57
Pool2	3×3	2	256	12×12	28×28
Conv4	3×3	1	384	10×10	26×26
Conv5	3×3	1	384	8×8	24×24
Conv6	3×3	1	256	6×6	22×22

算法整体流程如图3所示, 通过骨干网络来提取模板图像和搜索图像的模板特征和搜索特征, 对模板特征和搜索特征进行互相关操作得到响应图, 响应值最大处即为目标位置。该方法可以提取鲁棒性更强的特征, 从而有效地应对目标在尺度变化、遮挡等复杂环境下的跟踪问题。

2 实验结果及分析

实验在 Windows10 操作系统下采用 PyTorch 2.0.1 深度学习框架进行, 使用 GPU(NVIDIA GTX 2080Ti)进行加速。使用 GOT-10K^[19]训练集离线训练本文算法 SiamDBB, 采用 Xavier 均匀分布对骨干网络进行初始化, 学习率采用 10^{-2} 到 10^{-5} 的线性衰减策略, 共训练 50 个周期, 并采用动量为 0.9 的随机梯度下降方法作为优化器。选择 OTB100、UAV123 和 Temple-Color-128^[20-22]作为测试集。

2.1 新型多分支结构的消融实验

为了验证新型多分支结构中每条分支和特征

剪枝的有效性, 本节在 OTB100 数据集上进行消融实验。实验以 SiamFC 为基准, 在其骨干网络 AlexNet 中添加一层 $K=3$ 的新型多分支结构(DBB, Diverse Branch Block), 每次试验仅改变该层分支结构, 实验结果如表1所示。

由表3可知, 与单一卷积相比, 增加了不同分支的骨干网络使跟踪精确度和成功率有不同程度的提升。这是因为: 1×1 卷积可以实现信息的跨通道整合与交互, 且与 3×3 卷积形成了不同尺度的卷积组合, 丰富了特征空间; $1 \times 1 \sim 3 \times 3$ 序列卷积在感受野与单一卷积相同的情况下叠加了一层卷积, 可以引入更多非线性; 平均池化使提取的特征中包含了部分背景信息, 可以帮助跟踪器进行目标定位; 特征剪枝操作去除了受填充影响的特征, 恢复了孪生网络的平移不变性, 提高了跟踪性能。通过消融实验, 证明了新型多分支结构和特征剪枝的有效性。

2.2 定性分析

为了直观呈现算法的跟踪效果, 如图4所示, 在 Soccer、Motor Rolling 和 Skiing 高难度视频序列上, 将本文算法 SiamDBB 与 SiamFC、CFNet^[23]进行对比分析。

① 相似目标干扰: 在“Soccer”视频中, 目标不断变换位置和尺度的同时, 环境中充斥着相似物体的干扰。由于 SiamFC 和 CFNet 的骨干网络均采用单一卷积, 特征表达能力不足, 都出现了跟踪漂移问题。本文算法通过引入多分支结构, 将不同复杂度的分支组合, 丰富特征空间, 使提取的特征更具鲁棒性, 在面对相似目标干扰时跟踪效果更好。

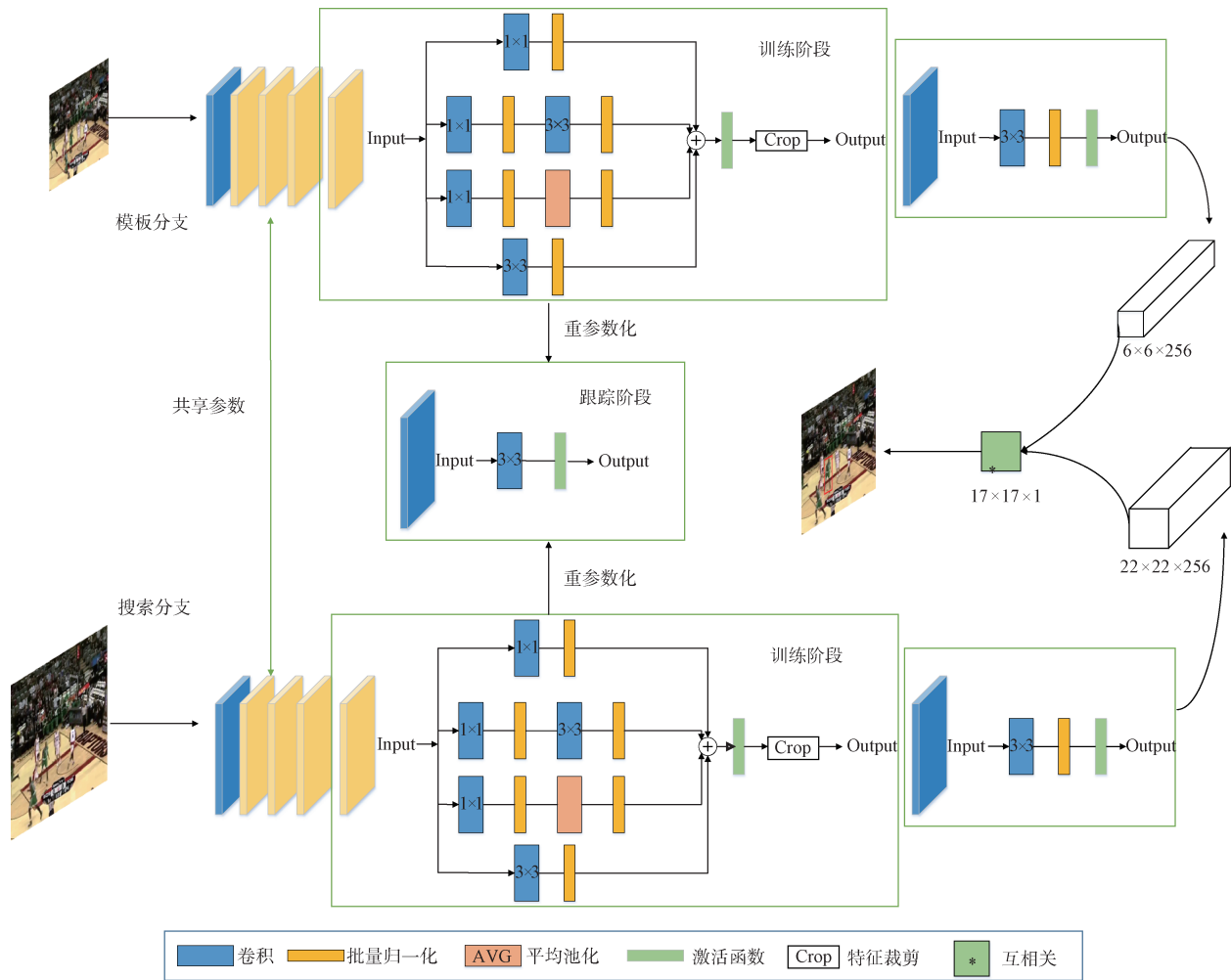


图3 算法流程示意图

Fig. 3 Algorithm flow diagram

表3 消融实验

Table 3 Ablation experiment

骨干网络	3×3	1×1	1×1-3×3	1×1-平均池化	特征剪枝	精确度	成功率
	√					0.759	0.557
	√	√				↑ 0.7%	↑ 0.9%
	√		√			↑ 2.2%	↑ 0.7%
	√			√		↑ 0.9%	↑ 0.6%
AlexNet+DBB	√	√	√			↑ 4.1%	↑ 2.3%
	√	√		√		↑ 3.1%	↑ 2.4%
	√		√	√		↑ 3.3%	↑ 2.0%
	√	√	√	√		↑ 3.8%	↑ 3.3%
	√	√	√	√	√	↑ 4.8%	↑ 4.6%

② 尺度变化：在“Motor Rolling”和“Skiing”视频中，目标旋转的同时，摄像的角度和距离也在变化，SiamFC和CFNet由于每层卷积的尺度固定，难以适应这样的复杂场景，而本文提出的多分支结构使用不同尺度的卷积，增强了特征

对尺度变化目标的适应能力。

2.3 算法整体性能

在OTB100数据集中，将本文算法与其他8个算法进行了比较，包括：CFNet、DSST^[24]、MUSTER^[25]、Staple^[26]、SRDCF^[6]、MEEEM^[27]、

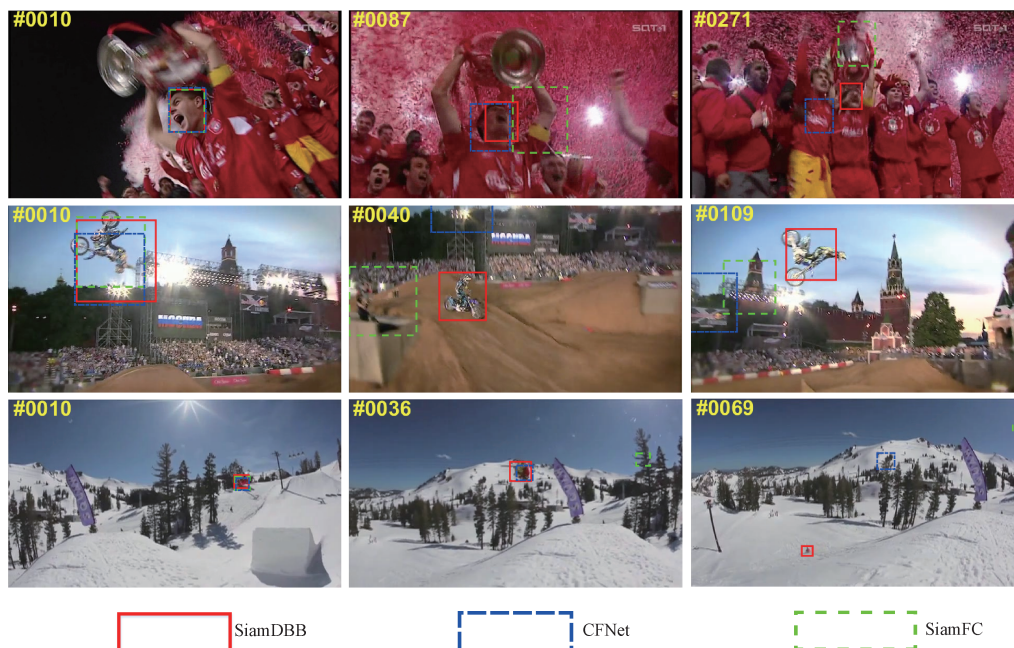


图4 算法定性分析对比图

Fig. 4 Qualitative analysis comparison diagram of algorithm

SiamFC^[8]、SiamTri^[28]。图5显示了本文算法与8个比较算法在跟踪精度和成功率两方面的对比图。得益于多分支结构的应用，本文算法相较于Sia-

mFC，跟踪精度提高了5.1%，跟踪成功率提高了3%。相比于同样基于SiamFC算法进行改进的CFNet和SiamTri，本文算法在跟踪性能上更具优势。

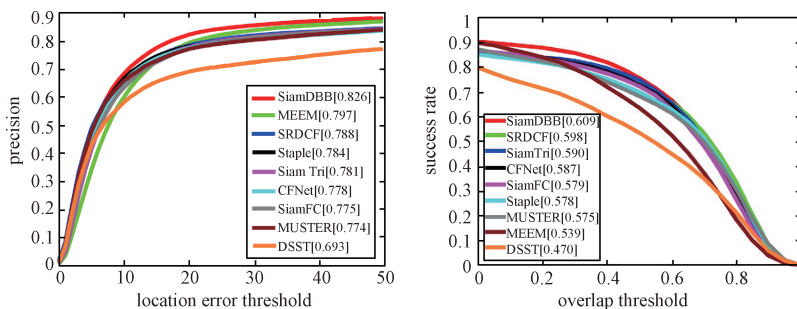


图5 算法在OTB100上的跟踪精度和跟踪成功率

Fig. 5 The tracking accuracy and tracking success rate of the algorithm on OTB100

为了证明重参数化对跟踪性能的影响，将重参数化的SiamDBB算法与未重参数化的SiamDBB算法进行对比发现：在跟踪精度和跟踪成功率上两者表现出相同的性能，而在跟踪速度方面，如表4所示，前者的跟踪速度较后者加快了47 FPS。这是因为未重参数化的SiamDBB在跟踪阶段时的骨干网络依旧是多分支结构，因此花费大量时间在特征提取上，使得跟踪速度远慢于重参数化的SiamDBB算法。而后者优化了SiamFC的骨干网络结构，进一步减少了参数量，因此跟踪速度可以达到110 FPS。

表4 跟踪速度对比

Table 4 Tracking speed comparison

算法	SiamDBB	SiamDBB(未重参数化)	SiamFC	CFNet
速度(FPS)	110	63	80	43

在UAV123数据集上，将SiamDBB与UAV123数据集提供的31个基础算法和MUSTER、SRDCF、KCF^[29]、DSST、DCF、MEEM、Struck^[30]等7个主流跟踪算法进行比较。图6展示了在UAV123数据集上9个性能最佳的跟踪算法。SiamDBB取得了最佳的跟踪效果，相较于SiamFC在跟

踪精度上提高了 1.8 个百分点, 在跟踪成功率上提高了 1.5 个百分点。

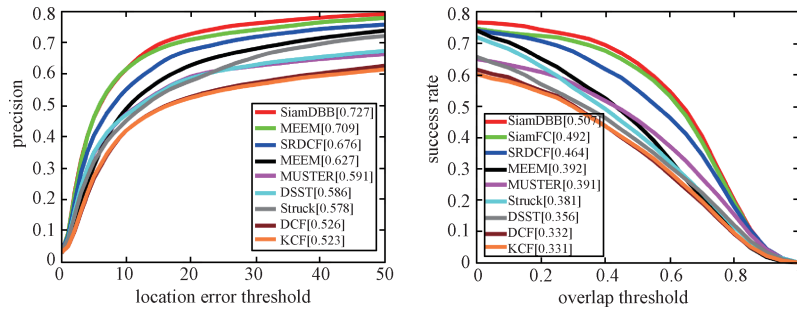


图 6 算法在 UAV123 上的跟踪精度和跟踪成功率

Fig. 6 The tracking accuracy and tracking success rate of the algorithm on UAV123

在 Temple-Color-128 数据集上, 将 SiamDBB 与 SiamFC、MEEM、Struck、KCF 等跟踪算法进行比较。图 7 展示了在 Temple-Color-128 数据集上的跟踪精度和跟踪成功率对比图。在该数据集上,

SiamDBB 在跟踪精度和跟踪成功率方面均取得了最佳性能, 跟踪精度达到了 73.18%, 跟踪成功率达到了 51.87%, 比 SiamFC 高出约 1.8 个百分点。

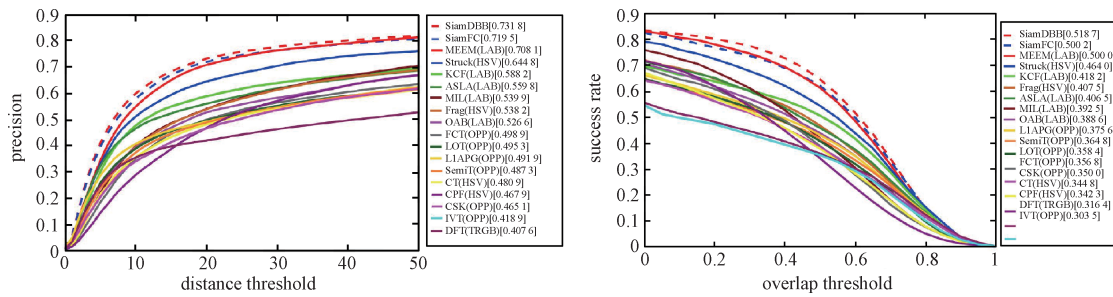


图 7 算法在 Temple-Color-128 上的跟踪精度和跟踪成功率

Fig. 7 The tracking accuracy and tracking success rate of the algorithm on Temple-Color-128

2.4 算法在各场景性能分析

本节在 OTB100 数据集下对 SV(尺度变化)、OPR(平面外旋转)、IPR(平面内旋转)、OCC(遮挡)、DEF(形变)、FM(快速运动)、IV(照明变化)、BC(背

景杂乱)、MB(运动模糊)、OV(出视野)和 LR(低分辨率)场景进行性能分析。表 5 和表 6 分别展示了在不同场景下算法的跟踪精度和跟踪成功率, 最优结果进行加粗显示, 次优结果由下划线标注。

表 5 不同场景下的跟踪精度

Table 5 Tracking accuracy in different scenarios

算法	SV	OPR	IPR	OCC	DEF	FM	IV	BC	MB	OV	LR
SiamDBB	0.809	0.816	0.822	0.780	0.785	0.808	0.814	0.752	0.818	0.741	<u>0.873</u>
SiamFC	0.732	0.778	0.759	0.739	0.748	0.739	0.712	0.706	0.728	0.655	0.762
MEEM	<u>0.751</u>	<u>0.810</u>	<u>0.814</u>	<u>0.774</u>	0.732	0.756	0.766	0.759	0.766	<u>0.739</u>	0.678
MUSTER	0.710	0.744	0.773	0.734	0.689	0.683	0.782	0.784	0.678	0.590	0.673
SRDCF	0.747	0.740	0.742	0.732	0.736	<u>0.769</u>	<u>0.786</u>	<u>0.775</u>	<u>0.767</u>	0.602	0.663
Staple	0.727	0.738	0.768	0.727	<u>0.752</u>	0.709	0.783	0.749	0.700	0.668	0.610
SiamTri	0.748	0.763	0.774	0.726	0.680	0.763	0.746	0.715	0.727	0.723	0.897
CFNet	0.732	0.760	0.786	0.700	0.714	0.707	0.707	0.756	0.681	0.604	0.750
DSST	0.658	0.665	0.711	0.609	0.568	0.575	0.726	0.704	0.570	0.480	0.602

表6 不同场景下的跟踪成功率
Table 6 Tracking success in different scenarios

算法	SV	OPR	IPR	OCC	DEF	FM	IV	BC	MB	OV	LR
SiamDBB	0.592	0.581	0.592	0.569	0.553	0.609	0.593	0.556	0.632	0.558	0.635
SiamFC	0.542	0.562	0.560	0.557	0.543	0.579	0.534	0.524	0.580	0.510	0.532
MEEM	0.475	0.534	0.539	0.526	0.482	0.543	0.531	0.536	0.575	0.515	0.376
MUSTER	0.510	0.535	0.551	0.552	0.521	0.531	<u>0.599</u>	<u>0.581</u>	0.544	0.462	0.454
SRDCF	<u>0.563</u>	0.549	0.543	<u>0.557</u>	<u>0.544</u>	<u>0.597</u>	0.609	0.583	<u>0.594</u>	0.461	0.495
Staple	0.521	0.533	0.548	0.543	0.551	0.540	0.592	0.560	0.541	0.475	0.400
SiamTri	<u>0.563</u>	<u>0.563</u>	<u>0.580</u>	0.549	0.501	0.585	0.579	0.542	0.567	<u>0.543</u>	<u>0.634</u>
CFNet	0.546	0.553	0.568	0.527	0.526	0.554	0.541	0.561	0.540	0.454	0.554
DSST	0.403	0.441	0.475	0.416	0.406	0.427	0.490	0.481	0.439	0.364	0.301

从表中可以看出, 本文算法在大部分挑战因素中取得了最好的跟踪效果。相较于 SiamFC, 本文算法的性能在所有场景下都获得了大幅度提升。跟踪精度方面, SiamDBB 在遮挡、尺度变化、运动模糊、变形、照明变化、快速运动、出视野、平面外旋转和平面内旋转场景下取得了最好的跟踪性能, 在低分辨率场景下取得次优的跟踪性能; 跟踪成功率方面, SiamDBB 在遮挡、尺度变化、运动模糊、变形、快速运动、出视野、平面外旋转、平面内旋转和低分辨率场景下取得了最好的跟踪性能。

3 结束语

本文从卷积结构出发, 提出了基于多分支和重参数化的孪生网络跟踪算法。通过构建多分支结构来替换单一卷积, 在训练阶段可以实现对不同尺度目标的适应性学习, 提高模型对尺度变化目标的处理能力, 强化对目标信息的挖掘, 在跟踪阶段对多分支结构进行重参数化, 大幅减少计算复杂度。实验结果表明, 本文算法的性能在 OTB100、UAV123 和 Temple-Color-128 数据集上, 相较于基准算法, 获得了较大幅度提升。

参考文献

- [1] MARVASTI-ZADEH S M, CHENG L, GHANEI-YAKHDAN H, et al. Deep learning for visual tracking: A comprehensive survey[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(5): 3943-3968.
- [2] ZHU X F, WU X J, XU T, et al. Robust visual object tracking via adaptive attribute aware discriminative correlation filters[J]. IEEE Transactions on Multimedia, 2022, 24: 301-312.
- [3] KUMARI L K, RAMALAKSHMI K, SRINIVASARAGHAVAN V. Robust visual object tracking in clustered environment[C]// 2022 5th International Conference on Contemporary Computing and Informatics(IC3I), Uttar Pradesh, India. 2022: 1499-1503.
- [4] REN H, XIANG L, SHI T. Research on background learning correlation filtering algorithm with multi-feature fusion[J]. IEEE Access, 2023, 11: 32895-32906.
- [5] KUMAR R, DEB A K. Pedestrian tracking in UAV images with Kalman filter motion estimator and correlation filter[J]. IEEE Aerospace and Electronic Systems Magazine, 2023, 38(7): 4-19.
- [6] WU L, FANG H, CUI X, et al. Improved part-based Kernel correlation filter for object tracking and robot grasping [C]// 2023 China Automation Congress(CAC), Chongqing, China. 2023: 505-510.
- [7] CAO Y, SHI G, ZHANG T, et al. Bayesian correlation filter learning with Gaussian scale mixture model for visual tracking[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2022, 32(5): 3085-3098.
- [8] BERTINETTO L, VALMADRE J, HENRIQUES J F, et al. Fully-convolutional Siamese networks for object tracking[C]//European Conference on Computer Vision, Amsterdam, The Netherlands. 2016: 850-865.
- [9] LI B, YAN J, WU W, et al. High performance visual tracking with Siamese region proposal network[C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA. 2018: 8971-8980.
- [10] GUO D, WANG J, CUI Y, et al. SiamCAR: Siamese fully convolutional classification and regression for visual tracking[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 6269-6277.
- [11] ZHANG Z, PENG H, FU J, et al. Ocean: Object-aware anchor-free tracking[C]// Proceedings of the European

- Conference on Computer Vision. 2020: 771-787.
- [12] CHEN Z, ZHONG B, LI G, et al. Siamese box adaptive network for visual tracking[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 6668-6677.
- [13] HELD D, THRUN S, SAVARESE S. Learning to track at 100 FPS with deep regression networks[C]// European Conference on Computer Vision, Amsterdam, The Netherlands. 2016: 749-765.
- [14] GUO Q, FENG W, ZHOU C, et al. Learning dynamic siamese network for visual object tracking[C]//2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy. 2017: 1781-1789.
- [15] HAN G, YANG R, GAO H, et al. Deep decoupling classification and regression for visual tracking[J]. IEEE Transactions on Cognitive and Developmental Systems, 2023, 15(3): 1239-1251.
- [16] LI Z, ZHANG J, LI Y, et al. Learning feature channel weighting for real-time visual tracking[J]. IEEE Transactions on Image Processing, 2022, 31: 2190-2200.
- [17] SUN N, ZHANG Z, XUE W. Contrastive-Siamese collaborative networks for visual object tracking[C]// 2022 International Conference on High Performance Big Data and Intelligent Systems(HDIS), Tianjin, China. 2022: 95-99.
- [18] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]// 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA. 2015: 1-9.
- [19] HUANG L, ZHAO X, HUANG K. GOT-10k: A large high-diversity benchmark for generic object tracking in the wild[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(5): 1562-1577.
- [20] WU Y, LIM J, YANG M H. Object tracking benchmark [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1834 - 1848.
- [21] MUELLER M, SMITH N, GHANEM B. A benchmark and simulator for UAV tracking[C]// Proceedings of the European Conference on Computer Vision. 2016: 445-461.
- [22] LIANG P, BLASCH E, LING H. Encoding color information for visual tracking: Algorithms and benchmark [J]. IEEE Transactions on Image Processing., 2015, 24(12): 5630-5644.
- [23] VALMADRE J, BERTINETTO L, HENRIQUES J, et al. End-to-end representation learning for correlation filter based tracking[C]// Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA. 2017: 5000 - 5008.
- [24] DANELLJAN M, HÄGER G, KHAN F S, et al. Discriminative scale space tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2017, 39(8): 1561-1575.
- [25] HONG Z, CHEN Z, WANG C, et al. Multi-Store Tracker (MUSTer): A cognitive psychology inspired approach to object tracking[C]// 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA. 2015: 749-758.
- [26] BERTINETTO L, VALMADRE J, GOLODETZ S, et al. Staple: Complementary learners for real-time tracking [C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA. 2016: 1401-1409.
- [27] ZHANG Jianming, MA Shugao, SCLAROFF S. MEEM: Robust tracking via multiple experts using entropy minimization[C]// Proceedings of the 13th European Conference on Computer Vision, Zurich, Switzerland. 2014: 188-203.
- [28] DONG Xingping, SHEN Jianbing. Triplet loss in Siamese network for object tracking[C]// European Conference on Computer Vision, Munich, Germany. 2018: 459-474.
- [29] HENRIQUES J F, CASEIRO R, MARTINS P, et al. High-speed tracking with kernelized correlation filters [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(3): 583-596.
- [30] HARE S, GOLODETZ S, SAFFARI A, et al. Struck: Structured output tracking with kernels[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(10): 2096-2109.

[作者简介]

金 铭 1968年生, 博士, 教授, 博士生导师。

唐 宇 2000年生, 硕士研究生。

韩 勇 1976年生, 博士, 讲师, 硕士生导师。

刘 帅 1980年生, 博士, 教授, 博士生导师。

闫锋刚 1983年生, 博士, 教授, 博士生导师, 泰山学者青年专家。

(本文编辑: 傅 杰)