



计算机科学

COMPUTER SCIENCE

基于注意力机制最大化重叠的单目标跟踪算法

孙开伟, 王支浩, 刘虎, 冉雪

引用本文

孙开伟, 王支浩, 刘虎, 冉雪. 基于注意力机制最大化重叠的单目标跟踪算法[J]. 计算机科学, 2023, 50(6A): 220400023-5.

SUN Kaiwei, WANG Zhihao, LIU Hu, RAN Xue. [Maximum Overlap Single Target Tracking Algorithm Based on Attention Mechanism](#) [J]. Computer Science, 2023, 50(6A): 220400023-5.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于动态时空神经网络的城市交通流量预测方法](#)

City Traffic Flow Prediction Method Based on Dynamic Spatio-Temporal Neural Network

计算机科学, 2023, 50(6A): 220600266-7. <https://doi.org/10.11896/jsjcx.220600266>

[基于主动重心的青年高血压患者心肺运动时序数据增强](#)

Data Augmentation for Cardiopulmonary Exercise Time Series of Young Hypertensive Patients Based on Active Barycenter

计算机科学, 2023, 50(6A): 211200233-11. <https://doi.org/10.11896/jsjcx.211200233>

[面向交通流量预测的时空Graph-CoordAttention网络](#)

Spatial-Temporal Graph-CoordAttention Network for Traffic Forecasting

计算机科学, 2023, 50(6A): 220200042-7. <https://doi.org/10.11896/jsjcx.220200042>

[基于多模态特征融合的时间序列异常检测](#)

Anomaly Detection of Time-series Based on Multi-modal Feature Fusion

计算机科学, 2023, 50(6A): 220700094-7. <https://doi.org/10.11896/jsjcx.220700094>

[联合人体姿态估计和多目标跟踪的跨数据集学习](#)

Cross-dataset Learning Combining Multi-object Tracking and Human Pose Estimation

计算机科学, 2023, 50(6A): 220400199-7. <https://doi.org/10.11896/jsjcx.220400199>

基于注意力机制最大化重叠的单目标跟踪算法

孙开伟 王支浩 刘虎 冉雪

重庆邮电大学数据工程与可视计算重点实验室 重庆 400065

(sunkw@cqupt.edu.cn)

摘要 随着人工智能的发展,深度学习在计算机视觉研究中引起了广泛关注,在单目标跟踪领域开始对基于深度学习的单目标跟踪算法加以研究。深度学习算法的算法复杂度相对较高,将目标分类和目标状态估计完整的分割出来,有利于对每一个任务的深层探讨。但现阶段单目标跟踪算法不能很好地应对复杂的跟踪环境,模型遇到复杂跟踪环境时,经常会跟踪到背景的某一块区域或者跟踪到周围的相似目标。为了解决以上问题,文中提出了一种方法,在目标分类和目标状态估计任务中分别加入了不同的注意力机制,使得模型能够更好地处理背景混乱和相似目标遮挡的情况。为了验证上述方法的有效性,文中在多个数据集上做了大量的对比实验,并且和之前的基于深度学习的单目标跟踪算法进行比较,所提算法在EAO指标上有了3.1%的提升,在Robustness指标上有了2.3%的提升,表明了其有效性和先进性。

关键词: 计算机视觉;单目标跟踪;注意力机制;权重分配;异常检测

中图分类号 TP391

Maximum Overlap Single Target Tracking Algorithm Based on Attention Mechanism

SUN Kaiwei, WANG Zhihao, LIU Hu and RAN Xue

Key Laboratory of Data Engineering and Visual Computing, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

Abstract With the development of artificial intelligence, deep learning has attracted extensive attention in the research of computer vision. In the field of single target tracking, the single target tracking algorithm based on deep learning has been studied. The algorithm complexity of deep learning algorithm is relatively high. The complete segmentation of target classification and target state estimation is conducive to the in-depth discussion of each task. However, the current single target tracking algorithm can not deal with the complex tracking environment well. When the model encounters the complex tracking environment, it often tracks a certain area of the background or tracks the surrounding similar targets. In order to solve the above problems. In this paper, a method is proposed; different attention mechanisms are added to the task of target classification and target state estimation respectively, so that the model can better deal with background confusion and occlusion of similar targets. In order to verify the effectiveness of the above methods, this paper has done a lot of comparative experiments on multiple datasets, and compared with the previous single target tracking algorithm based on deep learning. The proposed algorithm improves 3.1% in the EAO index and 2.3% in the Robustness index. It shows the effectiveness and progressiveness of this method.

Keywords Computer vision, Single target tracking, Attention mechanism, Weight distribution, anomaly detection

1 引言

随着计算机视觉的发展,深度神经网络在大量的图像处理任务上表现优异,早期的人们通过改变网络的架构,力求找到某种对于深度学习任务通用的一种网络架构,追求网络的普适性,越来越多的大型网络不断被挖掘出来,针对这些大型网络的优化方法也层出不穷,但仅仅有个别的网络脱颖而出,这些网络能够满足大部分的计算机视觉任务,但要求的参数量也十分巨大,对于设备的要求也越来越高,对于网络更是非常大的压力,基于这一原因,研究者们逐步将全局的网络思想转向局部的优化思想,受到人的视觉感知启发,注意力机制开始进入研究者的视野。注意力机制反映了人们在从认识到

熟悉目标的过程中,并不总是关注目标的整体,而是根据任务的需求来注意目标的核心部分以达到目的。目前,注意力机制在单目标跟踪中也有了突破性的进展,其核心思想是对数据的加权处理,从而使模型关注到主要的信息,忽略无关或者被干扰的信息。目前为止,单目标跟踪任务还有两方面待解决的问题。(1)不能应对复杂跟踪环境。对于目标状态变化不明显的简单跟踪环境,目前的算法能够达到显著的跟踪效果。但当跟踪环境较复杂时,例如面对物体大幅度形变、恶劣天气、光照变化等因素,算法的准确率会出现大幅度下降。其原因则是在训练时,模型已经适应简单的跟踪环境,当跟踪环境变复杂时,模型不能及时做出反馈,或者反馈一个相似的目标,如果在跟踪中某一帧出现严重状态估计错误(框住相似目

基金项目:重庆市自然科学基金面上项目(cstc2019jcyj-msxmX0021);重庆市教委项目(KJCXZD2020027);国家自然科学基金(61806033)

This work was supported by the Natural Science Foundation of Chongqing(cstc2019jcyj-msxmX0021); Science and Technology Research Program of Chongqing Municipal Education Commission(KJCXZD2020027) and National Natural Science Foundation of China(61806033).

通信作者:王支浩(870700802@qq.com)

标),则会对整个视频跟踪过程产生很大的影响。(2)不能处理目标的不确定信息。目前的算法中,对于模板帧的处理,没有使用和搜索帧同样的处理方法,只是把目标信息分离出来进行处理,这样在损失背景信息的基础上,还忽略了目标本身的不确定信息。例如,对于狗这一类目标的训练过程,在模板帧中狗的尾巴出现,而在搜索帧中却没有出现狗的尾巴,那么模型在训练时则可能会把狗的尾巴当作一种目标信息,当在真实的追踪过程中遇到相似的目标时,会因为尾巴这一特征的若隐若现而失去对追踪目标的判断。

2 相关工作

近年来,优秀的机器学习和深度学习算法相继被人提出,在单目标跟踪领域的算法也是进步显著,例如基于机器学习的均值漂移算法(MeanShift)、算法核相关滤波(Kernel Correlation Filter, KCF)等。MeanShift算法起初并不受重视,因为当时机器学习在单目标跟踪领域并没有较大发展,直到有研究者重新定位这篇文章后,才引起了广大单目标跟踪领域学者们的关注,之后的KCF算法对相关滤波框架做了详细的描述,KCF采用循环矩阵对图像所有区域进行采样,为了满足跟踪的实时性,使用传统的脊回归算法训练目标检测器,提高了模型的运算速度。当深度学习进入单目标跟踪领域后,使用深度神经网络训练出来的跟踪器可以更加准确地跟踪目标,并且可以保持与相关滤波算法同等的速度,使跟踪器的鲁棒性大大提高。基于孪生网络系列的单目标跟踪算法是一个良好的开端,网络将目标定位和目标估计任务分开,网络中的特征之间既相互独立又相互联系,浅层特征主要提取图片的时空信息,深层特征主要提取图片的语义信息,两者互相促进,使得模型获得一个更快的跟踪速度。但基于孪生网络的单目标跟踪算法对于快速移动的目标泛化能力较差,容易跟丢目标,之后的基于孪生网络的算法主要在增加网络的深度方面进行深入研究,获取更具有鲁棒性的特征和更加具有泛化能力的模型。当获取深度特征的方法达到一定的瓶颈期时,研究者把研究目标转向尺度估计,使得跟踪器能够有更加灵活的尺度适应方法,提出了一种使候选框和原标记框交并比(Intersection-over-Union, IoU)最大化的训练思路,即使得跟踪算法的预测框和原标记框之间的重叠率能够最大。ATOM算法利用IoU进行尺度搜索,通过模板帧图像和搜索帧图像的端到端的训练,训练出一个准确的IoU预测器,来预测每个候选框的IoU值,从而确定目标的状态。ATOM算法在训练IoU预测器时使用了传统的离线训练方法,学习通用的特征表示,将目标定位任务放到了能够在线更新的网络中。ATOM算法还指出了离线训练和在线更新应分开,对于在线更新模型,又提出了一种高效的优化策略,设计了以共轭梯度算法为基础的在线更新算法,以此得到了一个更加健壮的跟踪器。

3 基于注意力机制的重叠最大化单目标跟踪算法

3.1 基于注意力的权重分配机制

在目标估计网络中,通过池化层的特征已经拥有目标和有益的背景信息,这些信息可以直接经过全连接层进行组合,但直接组合的新的特征可能会影响模型的判断,因为对于一个目标,目标的所有信息并不都是模型所需要的,并且目标也

有可能出现遮挡、模糊等情况,这些因素也会影响模型的精度,模型需要从大量的信息中提取出具有表达能力的信息。基于注意力机制^[4-5]可以给数据分配特定的权重,将采用以下方法赋予某些区域一定的权重:在模板帧通过ResNet-18之后会提取到block3和block4两部分的特征,这两部分同时进入RoI Align^[6](见图1)池化层进行特征降维和压缩,当经过RoI Align池化的特征出来后,对于此部分特征图进行区域划分,分为左上、右上、左下、右下、中心5个候选区域,这5个区域能够从各个方面分析图片特征,之后进行区域特征分析,并将5个区域赋予固定的权重参数,这些权重参数在一定程度上代表了此区域的重要程度,且和特征图中的参数进行矩阵运算,使得所有的参数都带有本区域重要程度这一特征,接着将这两部分的特征进行加权融合,此时组合的特征、目标信息就由原来散乱分布变成了有规律分布,从某种意义上说,模型能够自动聚焦于信息量大的地方,而自动忽略那些被遮挡或者模糊的目标局部信息。详细的注意力机制方法如图2所示。

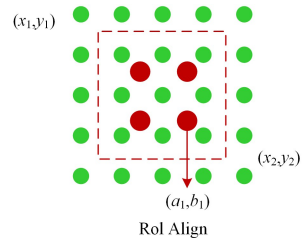


图1 RoI Align 池化方法

Fig. 1 RoI Align pooling method

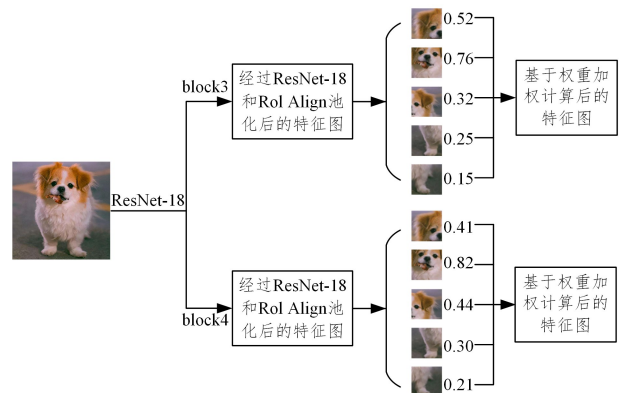


图2 模板帧注意力机制模块

Fig. 2 Template frame attention mechanism module

RoI Align的反向传播算法如下:

$$\frac{\partial L}{\partial x_i} = \sum_r \sum_j [i = i^*(r, j)] \frac{\partial L}{\partial y_{rj}}$$

其中, x_i 表示池化前特征图上的点, y_{rj} 表示第 r 个小区域中的第 j 个点, $i^*(r, j)$ 表示像素点的来源, 由公式可以看出, 当池化过程中利用到某个点 x_i 的像素值时, 才会在这个点处传回梯度。

RoI Align对回传梯度的方式做了改进:

$$\frac{\partial L}{\partial x_i} = \sum_r \sum_j [d(i, i^*(r, j)) < 1] (1 - \Delta h) (1 - \Delta w) \frac{\partial L}{\partial y_{rj}}$$

池化后的 $i^*(r, j)$ 是一个浮点数坐标, 在池化前, 每一个与这个坐标插值小于1的像素点, 都接受池化后 y_{rj} 回传的梯度, $d(i, i^*(r, j))$ 表示像素点距离。

3.2 基于注意力的异常选择机制

在单目标跟踪任务中,不仅可以在目标估计网络中添加注意力机制来增强模型的泛化能力,还可以在目标分类^[7-8]模块中增加注意力机制,来提高模型的定位准确度。目标分类模块不同于目标估计模块,没有庞大的网络架构,不能在网络的局部做出改变。如果在网络的局部增加注意力机制的小网络,则可能无法保证模型的实时性,因为随着网络的扩大,参数会不断增加,模型在进行在线更新时会更新更多的参数,导致模型的跟踪过程变慢。虽然这样的方法可以提高目标定位的准确性,但给网络带来的参数是无法估量的,是一种得不偿失的做法。为了在保证速度的同时提高定位的准确率,注意力会放在对样本的选择性更新策略上,前文提到了多种单目标跟踪算法的在线更新机制,MDNet^[9]采用 Long-term update 和 short-term updates 两种方式进行更新,这两种方式都是人为指定模型在跟踪的过程中每隔固定几帧就进行更新,不同的是间隔的时间长短。之后的 RT-MDNet^[10]算法对网络进行了改进,增加了池化层,对于准确率有了一定的提高,但是跟踪速度并没有提升。以上两种算法都是通过加入一些人为因素来控制模型的跟踪过程,这种方法虽然可以对模型的泛化能力带来一定的提升,但是不能控制在跟踪过程中一些不可变的因素,例如当跟踪目标突然消失或者发生严重遮挡时,模型收集到的全部是背景信息,如果在此时更新模型,那么后续的定位会发生严重偏移。对于以上这种更新情况,需要对模型增加一个自注意力机制,当跟踪目标从视频帧中消失时,模型应该注意到这一信息从而放弃更新,这种方法不仅提高了模型的跟踪速度也提高了模型定位的准确性。图3和图4截取了跟踪过程中部分样本的变化情况。

从图3中可以看出,在第一列前5个视频帧中,目标外观

比较清晰,而且不同帧也包含了目标的不同状态,对于目标分类模块的在线更新环节可以学到目标不同状态下的特征信息,这样对后续帧的目标追踪有一定的提升。对于后续的5个视频帧,可以明显看到目标逐渐变得模糊,视频帧被背景信息覆盖,到第8帧已经完全没有目标信息,如果在这时继续追踪目标,效果是非常差的,并且跟踪结果也是不可靠的。对于图4,在跟踪过程中背景信息混乱,并且在跟踪中有大量的相似目标,如果选择在这个时间段更新模型,则会对模型造成严重的影响,模型可能会跟踪到相似目标。通过改进网络以及人为规定更新时间 and 次数并不能应对这种突发性的情况。本文在处理这个问题时设置了一种异常处理机制。该机制计算跟踪后目标区域特征图的误差来确定模型是否需要更新,算法的流程如图5所示。



图3 OTB-100数据集 Bird1 视频部分的样本
Fig.3 Samples of Bird1 video in OTB-100 dataset

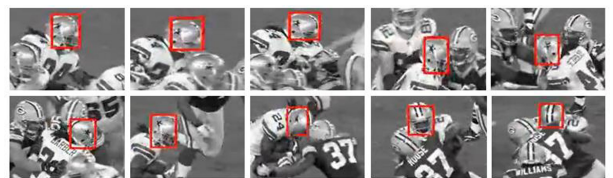


图4 OTB-100数据集 Football 视频部分的样本
Fig.4 Samples of Football video in OTB-100 dataset

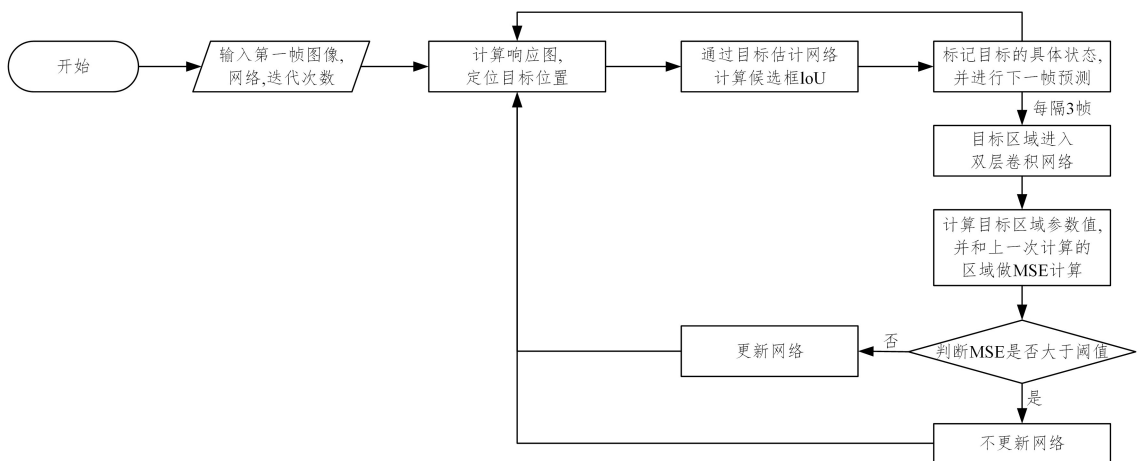


图5 目标分类网络在线更新异常检测机制

Fig.5 Anomaly detection mechanism for online update of target classification network

跟踪开始阶段,输入第一帧图像、目标分类网、迭代次数,通过网络之后生成以目标为中心的最大响应图,并以目标中心点为中心生成候选框,然后通过目标估计网络对候选框进行判断,选取IoU最大的候选框作为预测框。在模型经过3帧之后,要对目标区域进行判断,如果预测的目标区域和之前的目标区域参数的MSE相差很大,那么很可能是由于背景信息变的复杂和目标丢失引起的。这时模型不会进行更新,因为如果这时更新会给模型带来很多强干扰的背景信息,

反之,如果MSE相差不大,那么可以判断模型并没有受到跟踪环境的影响,可以更新模型。此方法在效果上可以增强跟踪模型的鲁棒性,模型的跟踪速度也没有受到较大影响。

4 实验结果与分析

4.1 评价指标

目标估计网络实际上是一个回归问题,在评价指标的选择上,采用了传统的均方误差(MSE),用于表现真实的IoU

值和预测的 IoU 值之间的误差:

$$MSE = \frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2$$

对于 MSE, 它本身的含义是求一批 (batch) 样本中 m 个样本的 m 个输出与期望输出的差的平方的平均值, 对于像目标估计这样的网络框架, 使用 MSE 相比其他的基于准确率的评价指标效果更好。对于最终的跟踪结果, 采用了准确率 (Accuracy)、鲁棒性 (Robustness) 和期望平均重叠率 (Expected Average Overlap, EAO) 这 3 个评价指标对跟踪结果进行比较。3 种评价指标的具体计算方法如下。

(1) Accuracy. 准确率用于来衡量跟踪器对目标跟踪情况的精确度, 对于每一帧的结果, 都采用了 IoU 的概念进行计算。

$$\phi_i = \frac{bb_i^g \cap bb_i^t}{bb_i^g \cup bb_i^t}$$

bb_i^g 是模型在 t 帧真实的目标标注框, bb_i^t 是模型在 t 帧预测的目标包围框, 分子取两个包围框的交集, 通常会在一段视频上运行 k 次来消除偶然因素, 取 k 次的平均结果作为当前帧的结果。在计算准确率时, 计算的是整个数据集的平均准确率, 即对所有有效帧的准确率求平均值, 有效帧即预测框和真实框的交集不为 0。最终 Accuracy 的表达式如下:

$$Accuracy = \frac{1}{N_{\text{test}}} \sum_{i=1}^{N_{\text{test}}} \phi_i$$

其中, N_{test} 表示测试集中所有有效帧数量。

(2) Robustness. 鲁棒性也是衡量单目标跟踪算法效果的有效评价指标之一。鲁棒性具体是指跟踪算法在跟踪过程中丢失目标的次数。

$$R = e^{-sM}$$

其中, s 是可选择参数, 通常取 30。

(3) Expected Average Overlap. 期望平均重叠率是将准确度和视频帧长度结合的一种方法, 表达式如下:

$$\bar{\phi}_{N_s} = \frac{1}{N_s} \sum_{i=1}^{N_s} \phi_i N_s$$

在计算之前, 首先对视频序列进行分类, 令算法在长度 N_s 的序列上进行测试, 得到每一帧的准确度 ϕ_i , 再取平均值得到该视频序列的准确度 $\bar{\phi}_{N_s}$ 。最后在所有长度为 N_s 的视频序列上进行测试, 并再次取得平均值, 从而得到算法在长度为 N_s 的视频序列上的 EAO 值。

4.2 对比分析

本实验中, 第一部分是探究基于注意力的权重分配机制对模型的影响, 第二部分是探究异常检测机制对模型的影响, 第三部分是对两种方法在不同跟踪情况下的跟踪结果进行分析。

第一部分: 探究基于注意力的权重分配机制对模型的影响

权重分配机制是目标估计网络中对于目标特征信息的再学习, 增加权重分配机制主要是对经过池化后的目标特征进行处理, 通过划分目标区域, 赋予目标区域数据权重, 对目标的各个区域做一个重要性程度的划分, 实验分析了在权重分配机制下目标估计网络的 MSE 曲线。图 6 给出了在增加权重分配机制后的目标估计网络的 MSE 曲线。从图中可以看出, 增加权重分配机制使得目标估计网络模型的精确度得到了巨大提升, 但是函数在接近 40 Epochs 时才收敛, 比原始的

方法慢了很多。这主要是由于在训练时要对区域的权重参数进行学习, 这部分权重参数相比网络的参数更难学习。因为对于不同的目标, 模型关注的区域会有差别, 重要程度也不同。

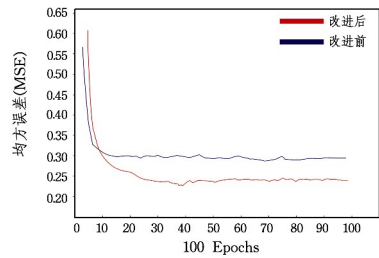


图 6 权重分配机制下目标估计网络的 MSE 曲线

Fig. 6 MSE curve of target estimation network under weight allocation mechanism

第二部分: 探究异常检测机制对模型的影响

在跟踪过程中, 模型要进行在线更新。传统的在线更新算法应用到网络中时通常都是人为规定模型的更新时间, 异常检测机制消除了这种人为因素, 模型通过参数之间的运算来判断此时是否需要更新。这种异常检测方法使得模型的在线更新机制更加具有主观性, 并且可以在合适的时机了解图片的变化, 做出合适的应对策略, 提高了对目标变化的敏感性。这部分实验加入了这种异常检测机制, 在对 3 个评价指标进行比较的基础上增加了跟踪速率 (FPS) 指标, 这个指标是比较模型在 1s 内跟踪的视频帧数。表 1 列出了跟踪的结果, 最后又通过绘制 A-R 图 (7) 来展示算法的有效性, 在 A-R 图中, 越靠左上角的算法综合能力越强。从表 1 中可以看出, 异常检测机制极大地提升了模型的鲁棒性, 但增加了少量参数, 因此准确率和跟踪速率略有下降, 并没有出现大幅度下降的情况, 这也证明了异常检测机制在本文中的有效性, 在算法的比较中也排在前列。

表 1 异常检测机制下的跟踪结果

Table 1 Tracking results under anomaly detection mechanism

	DRT	RCO	ATOM	Ours
EAO	0.356	0.376	0.401	0.432
Robustness	0.201	0.155	0.204	0.132
Accuracy	0.519	0.507	0.590	0.570
FPS	15	9	42	38

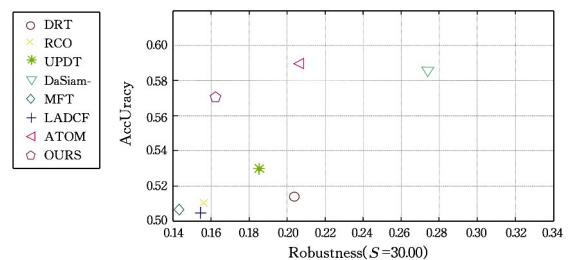


图 7 跟踪准确率和鲁棒性综合评价 (A-R 图)

Fig. 7 Comprehensive evaluation of tracking accuracy and robustness (A-R diagram)

第三部分: 不同跟踪环境下的跟踪结果分析

对于注意力机制的有效性, 通过展示在不同跟踪环境下的跟踪结果来证实算法的有效性。两种跟踪环境如图 8、如图 9 所示。分析简单环境下的跟踪结果可知, 在简单跟踪

环境中,3种算法都能够准确定位到目标的位置,但是对于目标的状态估计,本文算法和 ATOM 算法的效果更好,具体来说,这两种算法能够包围住目标的所有部分,并且目标框中不会出现赘余空白部分,这说明模型对于目标特征的把握是非常准确的,同时也能证明模型并没有受到背景信息的影响。在复杂环境下,3种算法的准确性和鲁棒性得到了充分的体现,图9中第一行对鸟的追踪,C-RPN 算法对目标的定位效果没有其余两种算法效果好,在0188帧中,鸟刚刚穿过云层,算法直接将目标定位到旁边的相似目标,0058帧也出现了类似的情况,说明此算法不能很好地应对相似目标遮挡的跟踪情况。在后面两个跟踪视频中,ATOM 算法和本文算法都能够准确定位到目标的位置,但对于目标状态的估计,本文算法可以更加准确地预测目标的状态,即使背景信息混乱,周围有大量相似目标,也不会影响算法的跟踪准确度和鲁棒性。

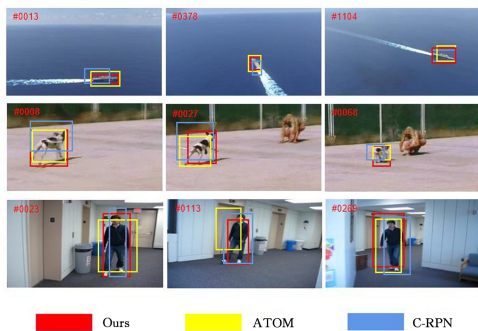


图8 简单跟踪环境下3种算法的跟踪结果

Fig. 8 Tracking results of three algorithms in simple tracking environment

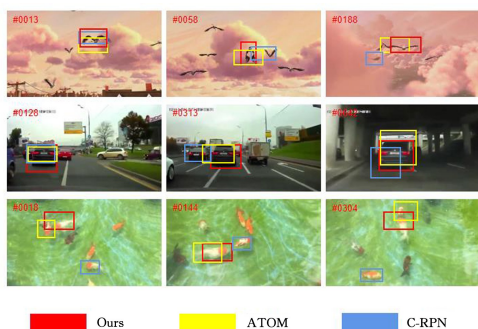


图9 复杂跟踪环境下3种算法的跟踪结果

Fig. 9 Tracking results of three algorithms in complex tracking environment

结束语 针对现有单目标跟踪算法存在的问题,本文对于目标信息的处理,不再只使用单纯的卷积和池化,而是增加权重分配机制,对于目标区域的各个部分都赋予重要性程度特征,并且通过训练优化这些区域权重,在实验阶段取得了很好的效果。在目标分类模块也使用了基于注意力的异常检测机制来确定模型更新的时机,两种方法结合使得跟踪算法在准确率、鲁棒性、速度等方面都有了明显的提升。

本文对在线更新的时机选择进行了一定的改进,并没有对在线更新算法做出改变,在目标分类网络中,在线更新算法要满足速度快、高效率的要求,这对算法的设计是很大的挑战,未来将在这个在线更新算法领域做出一定的设计,设计出更快、高效的更新算法。

参考文献

- [1] ZHANG W C,SUN C M. Corner Detection Using Multi-directional Structure Tensor with Multiple Scales[J]. International Journal of Computer Vision,2020,128(2):438-459.
- [2] ZHAO F,WANG J Q,WU Y,et al. Adversarial Deep Tracking [J]. IEEE Transactions on Circuits & Systems for Video Technology,2018,29(7):1998-2011.
- [3] YU L Y,FAN C X,MING Y. Improved Target Tracking Algorithm based on Kernelized Correlation Filter [J]. Journal of Computer Applications,2015,35(12):3550-3554.
- [4] WANG N,SONG Y B. Unsupervised Deep Tracking[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). Long Beach:IEEE,2020:1308-1317.
- [5] WU Y,LIM J. Object Tracking Benchmark[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9):1834-1848.
- [6] WU Y,LIM J. A Numerical Algorithm for the Coupled PDEs Control Problem[J]. Computation Aleconomics, 2019, 53(2): 697-707.
- [7] DANELLJAN M,BHAT G,KHAN F S,et al. ATOM: Accurate tracking by overlap maximization[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach:IEEE,2019:4655-4664.
- [8] KUMAR D,MONDAL S,KARUPPUSWAMI S,et al. Harmonic RFID Communication Using Conventional UHF System [J]. IEEE Journal of Radio Frequency Identification, 2019, 3(4):227-235.
- [9] ZHANG X M,ZHANG X H,DU X D,et al. Learning multi-domain convolutional network for rgb-t visual tracking[C]//2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI). Beijing: IEEE,2018:1-6.
- [10] JUNG I,SON J,BAEK M,et al. Real-time mdnet [C]//Proceedings of the European Conference on Computer Vision(ECCV). Cham:Springer,2018:83-98.
- [11] FERRAZ P A P,DE OLIVEIRA B A G,FERREIRA F M F, et al. Three-stage RGBD architecture for vehicle and pedestrian detection using convolutional neural networks and stereo vision [J]. IET Intelligent Transport Systems, 2020, 14(10): 1319-1327.



SUN Kaiwei, born in 1987, Ph.D, associate professor. His main research interests include machine learning, data mining and big data analysis.



WANG Zhihao, born in 1997, master's degree. His main research interests include computer vision and single target tracking.