

DOI: 10.3969/j.issn.1009-9492.2023.06.056

王帅, 孙海江, 江山, 等. 基于嵌入式平台的实时多目标编批研究 [J]. 机电工程技术, 2023, 52 (06): 263-268.

基于嵌入式平台的实时多目标编批研究*

王 帅¹, 孙海江¹, 江 山^{1*}, 刘心睿², 徐婉莹¹, 刘 颖¹

(1. 中国科学院长春光学精密机械与物理研究所, 长春 130033; 2. 吉林大学神经外科, 长春 130031)

摘要: 针对嵌入式平台多目标检测与实时跟踪任务, 提出基于目标特征与空间位置信息联合的多目标实时编批方法。采用板卡间目标信息交互联动方式, 智能板卡负责目标检测、编批, 为提升网络特征提取能力, 增强检测效果, 检测方法以YOLOv5网络为基础加入SimAMC3注意力机制模块。通过对检测目标的编批和编号更新, 实时反馈编批目标与跟踪目标的空间位置信息, 该模式改变过去传统跟踪器单一点选跟踪目标方式。综合考虑重叠面积、空间距离、长宽比、交并比等多方面因素实现目标检测与目标队伍的数据关联。实验表明, 该方法在目标运动时能够快速建立目标编号并管理更新, 同时能够提高跟踪系统对目标遮挡情况下的跟踪能力, 对消失后目标重捕获能力得到显著提高, 算法简单、易于工程实现。

关键词: 编批; 目标检测; 目标跟踪; 数据关联

中图分类号: TP391.4

文献标志码: A

文章编号: 1009-9492 (2023)06-0263-06

Research on Real-time Multi-objective Batch Editing Based on Embedded Platform

Wang Shuai¹, Sun Haijiang¹, Jiang Shan¹, Liu Xinrui², Xu Wanying¹, Liu Ying¹

(1. Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China;

2. Department of Neurosurgery, Jilin University, Changchun 130031, China)

Abstract: Aiming at the task of multi-target detection and real-time tracking on the embedded platform, a multi-target real-time batching method based on the combination of target features and spatial position information is proposed. The target information interaction and linkage method between boards is adopted, and the smart board is responsible for target detection and batching. In order to improve the network feature extraction ability and enhance the detection effect, the detection method is based on the YOLOv5 network and the SimAMC3 attention mechanism module is added. Through batching and serial number update of detected targets, real-time feedback of spatial position information of batched targets and tracked targets, this mode changes the traditional single-click tracking target mode in the past. The data association between target detection and target teams is realized by comprehensively considering multiple factors such as overlapping area, spatial distance, aspect ratio, and intersection ratio. Experiments show that this method can quickly establish the target number and manage the update when the target is moving, and at the same time can improve the tracking ability of the tracking system when the target is occluded, and the ability to recapture the target after disappearing is significantly improved. The algorithm is simple and easy to engineer. accomplish.

Key words: batch editing; target detection; target tracking; data association

0 引言

目标检测跟踪技术在计算机视觉领域扮演着至关重要的角色, 其实质是根据所提供目标在像面上的初始位置状态, 通过图像处理算法实时预测目标在视频的后续图像帧的位置状态。图像跟踪在安防监控、火控观瞄、交通道路监测、武器制导等领域都有广泛应用, 良好的图像跟踪性能是光电系统性能的核心保证^[1]。

基于DSP+FPGA的传统嵌入式跟踪器具有尺寸小、功耗低、整合度高、操作灵活性强等特点, 被广泛应用^[2]。传统嵌入式跟踪器的跟踪算法主要是对跟踪目标进行建模并完成其相关性匹配^[3]。基于目标建模和匹配的跟踪算法一般步骤是首先对目标进行特征提取, 根据提取到的目标特征完成建模, 最后采用相应相似性度量

准则完成目标模板与候选目标的相似匹配。传统跟踪器虽然可以完成高帧频跟踪任务, 但是在场景复杂时跟踪效果不佳。目前卷积神经网络已经表现出对图像特征强大的学习和表达能力, 在图像处理领域已经广泛应用, 故考虑在传统跟踪器基础上加入智能处理板卡(NVIDIA Jetson AGX Xavier)实现多目标检测跟踪任务^[4]。目前用于目标检测的深度学习网络分为以R-CNN^[5]及其变体Fast R-CNN^[6]、Faster R-CNN^[7]为代表的两阶段目标检测算法和以SSD^[8]、YOLO^[9-11]系列算法为代表的单阶段目标检测算法。

本文在传统跟踪器的基础上加入智能板卡, 设计多目标检测跟踪实现方案。采用“识别跟踪并行”的总体方案, 即通过智能板卡完成识别算法, 对当前存在的目

基金项目: 吉林省自然科学基金资助项目(20200401091GX)

收稿日期: 2023-04-27

标数据信息完成检测及编批,同时与跟踪器实时进行多目标信息的交互关联,进而实现了多目标的精准跟踪。该方法不仅可以应用卷积神经网络提升对目标的检测效果,同时还可以对多目标完成自动编批并指导跟踪器进行目标选取及重捕获过程,实现了高帧频复杂背景下的多目标检测跟踪任务。

1 多目标编批流程

本文多目标编批算法主要由智能板卡和跟踪器两个板卡相互配合完成。智能板卡主要负责目标检测,跟踪器负责目标跟踪。在多目标批流程中,编号建立和编号销毁的关联匹配准则是编号更新的基础。本文建立一个固定数量的队伍来管理检测到的目标,首帧完成检测目标的编号初始化,接下来以此为基础进行后续队伍的编号管理。由于队伍编号数量固定,涉及到队伍编号的销毁、更新、增补等过程。当队伍初始化后每一帧的检测结果与当前队伍目标进行关联匹配,来保持队伍的实时更新。队伍确定后智能板卡与跟踪器实时进行信息交互,来提升跟踪器的跟踪效果。由于对目标已经完成编号,在选择跟踪目标时可以通过编号进行跟踪目标的选择,该选择方式提升了目标选择的稳定性。目标编批具体流程如图1所示。

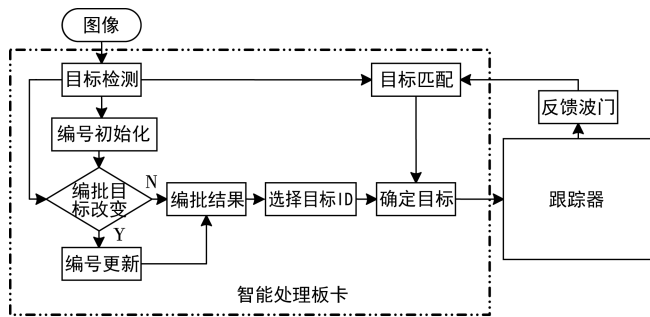


图1 多目标检测编批流程

2 检测算法

2.1 YOLOv5网络介绍

本文检测算法应用目前兼具时效性的YOLOv5。近年来YOLOv5作为单阶段目标检测算法被提出,与经典的两阶段网络相比,其具有更佳检测速度,同时其检测精度在众多单阶段检测网络中也脱颖而出并在工业界被广泛应用。YOLOv5共有4个不同网络模型YOLOv5S、YOLOv5M、YOLOv5L、YOLOv5X^[12-15],其网络架构相同,区别是具有不同的网络深度和宽度,其中YOLOv5S网络模型最小,网络模型依次增大,YOLOv5X网络模型最大,可以根据不同应用场景任务选择所需要的网络模型。YOLOv5S的深度和宽度是四种网络模型中最小的,参数量少,适用于嵌入式平台部署,且其网络模型检测精度也满足任务需求。其网络模型如图2所示。

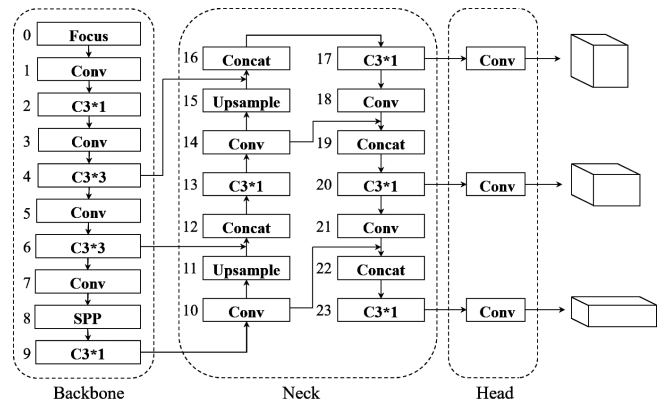


图2 YOLOv5S网络模型结构

YOLOv5S模型的构成主要包含特征提取层Backbone、特征融合层Neck以及检测层Head三部分,Backbone包含Focus层、卷积层、C3层和SPP层,用来进行目标特征的提取。Focus结构进行输入图片的切片操作,在减少参数计算量的同时可以降低图片在下采样过程中信息的丢失;卷积层主要负责提取图像的特征;C3结构是学习残差特征的重要模块,包含了3个标准卷积层和多个BottleNeck模块,其结构分为两支,其中一支只有一个标准卷积层,另一只经过一层卷积和多个BottleNeck层,最后两个支路完成融合。C3结构不仅能够提升网络对目标特征的学习和提取能力,还能够降低网络参数量,防止网络因层数增加而产生梯度消失或梯度爆炸的现象。空间金字塔池化SPP层,能够融合不同尺度的特征,其拥有4个最大池化层,池化核大小分别为 1×1 、 5×5 、 9×9 和 13×13 。Neck采用了FPN+PAN的特征金字塔结构,FPN是一个自上而下的网络结构,将高层特征逐渐传递下去,与之相对PAN是一个自下而上的网络,将低层的特征信息向上传播,两者结合将不同尺度的特征进行融合,从而提高模型对不同尺寸目标的检测能力。检测头部分用于进行分类回归预测,包含3个检测头,分别用来检测大、中、小目标。每一层特征图上,预设3个不同长宽比的预测框,每一个预测框上包含目标的位置信息和置信度,最后经过非极大值抑制(Non-maximum Suppression, NMS)算法将交并比(Intersection Over Union, IOU)^[16]超过设定阈值的重叠预测框丢弃,得到最终的检测结果。

除此之外,YOLOv5网络还会对图像进行一系列预处理操作,包括马赛克数据增强、自适应锚框计算以及自适应图片缩放,在进行特征提取之前完成这些操作可以提高检测的平均精度均值。

2.2 设计注意力机制

注意力机制(Attention Mechanism)类似于人类视觉系统中的视觉注意力机制,可以在目标检测任务中帮助

神经网络集中关注目标区域, 获取与目标相关的信息并抑制背景和干扰信号。在使用注意力机制时, 首先通过扫描图像生成权重, 判断每个区域的重要程度, 然后将注意力资源应用到目标区域, 以便更准确地提取目标的数据信息并对干扰背景进行抑制。注意力机制能够帮助网络更快速地筛选出目标的特征信息, 从而增强局部上下文信息的关联性。常见的注意力机制主要分成3种类型, 包括通道注意力机制、空间注意力机制和通道与空间混合的注意力机制。其中, 通道注意力机制沿着通道维度对特征进行处理, 生成一维权重, 可以对不同通道进行细化操作, 但对所有空间位置都视为同等重要; 空间注意力机制沿着空间维度对特征进行处理, 生成二维权重, 可以对不同位置进行细化操作, 但对所有通道同等看待。这些注意力机制能够提高网络的特征学习效果, 但在细化特征时只能考虑通道或空间维度, 限制了注意力权重学习的灵活性。

SimAM^[17]是一种简单有效的三维注意力模块, 其具有独特的能量函数设计, 可以在不增加网络参数的情况下发现每个神经元的重要性, 并推断出特征图的三维注意力权重。相比于现有的空间注意力模块和通道注意力模块, SimAM注意力机制具有更加灵活的特点, 并且可以在不增加参数数量的情况下提高网络性能。常见的注意力机制如图3所示。

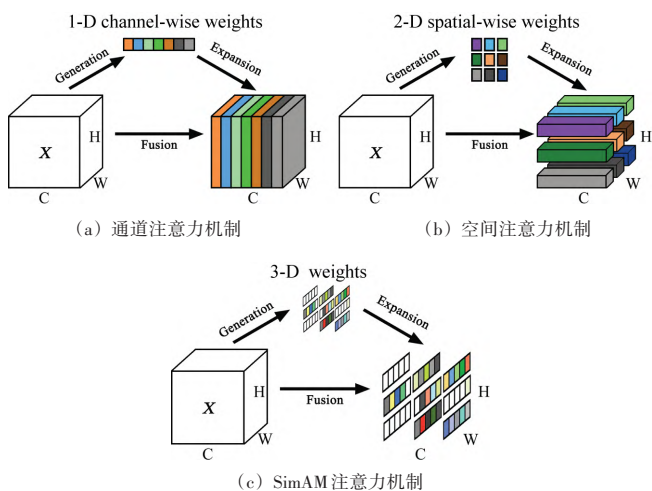


图3 注意力机制原理

SimAM注意力机制的实现原理基于视觉神经科学中的空间抑制现象^[18]。当一个神经元的放电模式与周围神经元不同时, 该神经元通常会具有更大的信息量, 并可能抑制其周围神经元的活动。在视觉处理中, 表现出空间抑制现象的神经元的重要程度更高, 应该具有更高的优先级。为了找到这些具有优先级的神经元, SimAM注意力机制给每个神经元定义了能量函数, 用于测量该神经元与其他神经元之间的线性可分性, 如式(1)所示:

$$e_i(w_i, b_i, y, x_i) = (y_t - \hat{t})^2 + \frac{1}{M-1} \sum_{i=1}^{M-1} (y_0 - \hat{x}_i)^2 \quad (1)$$

式中: $\hat{t} = w_t t + b$, $\hat{x}_i = w_i x_i + b_i$, 分别为 t 和 x_i 的线性变换; $M = H \times W$ 为通道上的神经元数量, 每个通道都有 M 个能量函数; w_i 和 b_i 为权重。

加入 SimAMC3 注意力机制模块后的网络整体结构图和 SimAMC3 注意力机制模块结构图分别如图4和图5所示。

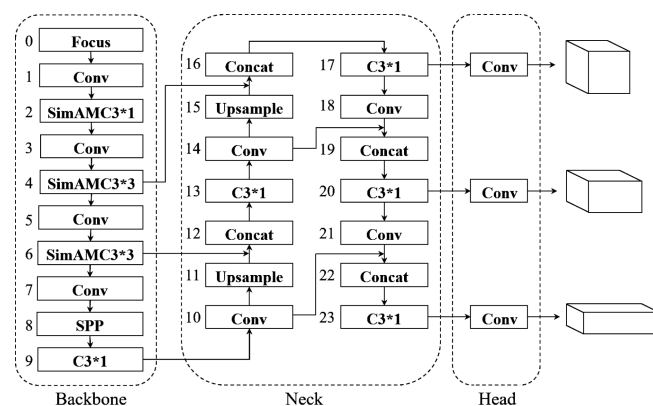


图4 加入 SimAMC3 模块后的网络结构

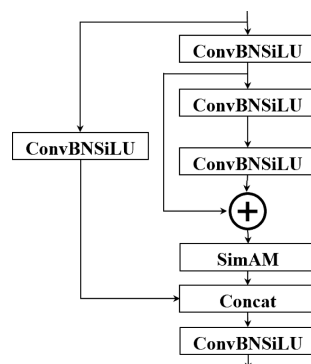


图5 SimAMC3 模块结构

SimAMC3 注意力机制模块应用在主干特征提取部分, 提升网络的特征提取层能力, 通过能量函数找到表现出具有空间抑制现象的神经元, 使网络能够更有效的提取目标的空间位置信息, 抑制复杂背景及噪声等冗余信息的干扰。由于 SimAM 注意力本身参数数量为零, 在不给原网络增加任何参数数量的情况下, 使得网络在检测目标时更加准确和鲁棒。

3 多目标编批算法详解

3.1 空间位置关联

数据关联是多目标编批问题的核心环节, 依次将当前帧的检测数据与已存在的目标队伍进行匹配, 依据匹配结果对目标队伍的信息不断更新。数据关联算法的优劣性对多目标编批的效果有着至关重要的影响^[19]。本文设计的数据关联算法流程如图6所示。

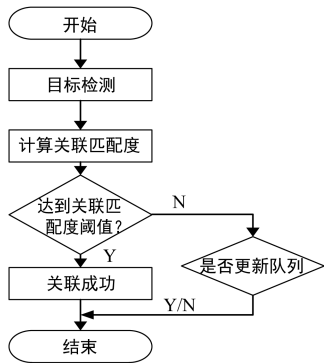


图6 数据关联流程

目前跟踪算法最常应用IOU来确定前后目标两帧的位置，即求出上一帧和当前帧之间的交集和并集之比。如果两帧重合则交并比等于1。IOU满足非负性、同一性、对称性、三角不等性，所以能够较好地反映相邻两帧间的位置关系。区域框交并比示意图如图7所示，IOU公式如下：

$$IOU = \frac{A \cap B}{A \cup B} \quad (2)$$

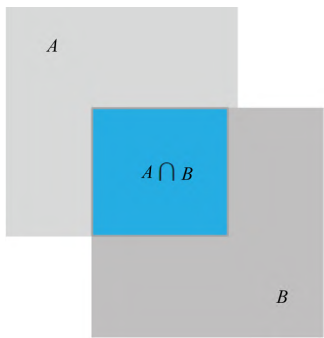


图7 区域框交并比示意

然而由于IOU并没有考虑框之间的距离，所以也有相应的缺点：在A框与B框不重合的时候IOU为0，不能正确反映两者的距离大小；而且IOU无法精确地反映两者的重合度大小。如图8所示，3种情况IOU都相等，但看得出来重合度是不一样的，左边的图的效果最好，右边的最差，这将对目标队伍编号更新产生影响。

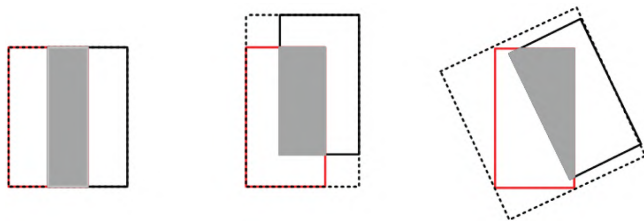


图8 相同IOU时不同重合度

因此本文采用重叠面积、中心点距离、长宽比、交并比等多方面因素，综合计算目标检测值与目标队伍的匹配相关度。如图9所示，空间位置关联依据如下式所示：

$$Corr = IOU - \frac{d^2}{c^2} - \alpha v \quad (3)$$

式中： d 为A框与B框中心点坐标的欧氏距离； c 为包住它们的最小外接矩形的对角线距离；空间位置关联公式中的 αv 是指前后帧两个框的宽高比，即当前后两帧框的宽高比，矩形对角线倾斜角度一致，两者形状一样。

$$\alpha = \frac{v}{(1 - IOU) + v} \quad (4)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (5)$$

长宽比示意图如图10所示。

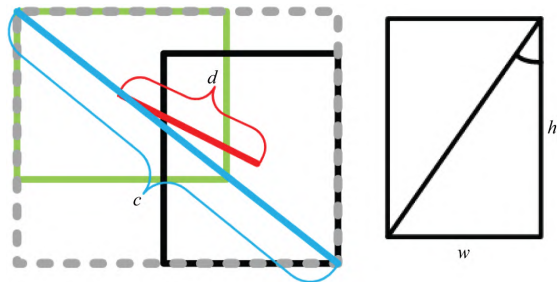


图9 空间位置关联依据示意图 图10 长宽比示意图

3.2 身份信息更新

每帧结束后都会对已有的目标队伍进行更新。数据匹配关联后，依次将当前帧的检测数据与目前存在的目标队伍进行匹配，根据匹配结果对目标队伍的信息不断更新。与现有编批算法身份信息编号序列不同的是，本文编批序号是固定不变^[20]。本文编批根据实际需求可以设置目标编号最多到 K 个，即图像内超过 K 个目标时只对其中 K 个赋予身份信息并完成目标身份信息的更新。同时对每个赋予目标的编号将保留 n 帧，即当该编号生成后目标消失超过 n 帧后，对编号进行销毁，并使用该

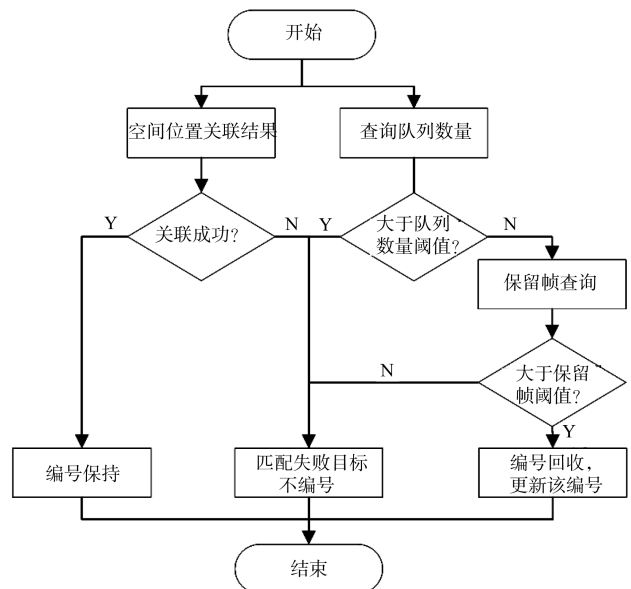


图11 编批更新流程



图12 复杂场景跟踪效果

编号为新的目标赋予身份信息。具体更新流程如图 11 所示。

在跟踪过程中保持目标身份信息不变依据空间位置关联判据, 具体更新流程可以理解为: 在原有目标未发生变化时且出现新的目标时, 将会判断目前队伍目标数量与预先设定编号数量, 当目前队伍目标数量大于预先设定编号数量时, 不赋予编号。当目前队伍目标数量小于预先设定编号数量且新出现目标连续出现 3 帧时, 将对闲置编号进行保留帧判断, 只有大于保留帧阈值时, 才会重新启用该编号。当编号启用后将每个赋予编号目标将保留 n 帧, 即当编号目标超过 n 帧依旧未出现即对编号进行销毁, 并使用该编号为新的目标赋予身份信息。

3.3 板卡联动

目前嵌入式跟踪器选择目标主要是点选目标模式, 当操作环境复杂或者跟踪目标很小时很难选中目标。智能板卡的加入可以解决该问题。智能板卡完成当前帧目标队伍编批后, 将信息发送给主控, 通过主控界面选择感兴趣目标进行跟踪, 当选择完成后智能板卡将编批目标信息发送给跟踪器, 完成跟踪目标选择。

智能板卡还将指导跟踪器完成目标重捕获过程。单一跟踪器在复杂场景或者背景与目标很相似时, 跟踪目标时常发生目标丢失现象。智能板卡与跟踪器通过信息交互, 在目标发生丢失时指导跟踪器完成重捕获, 大大提高跟踪精度。

3.4 多目标跟踪编批结果

为验证本文所提出的基于智能板卡联动的多目标编批方法效果, 在复杂背景且遮挡较多的自然环境下进行算法验证, 该场景包含了目标静止、目标机动、目标重捕获、目标交叉、多目标集中、目标遮挡、多目标分离等多种目标运动情况, 进而完成编批效果的充分验证。选取跟踪过程关键帧, 实验效果如图 12 所示。

在测试时, 初始根据智能板卡给定编号进行跟踪目标选取, 选择 1 号目标进行跟踪, 可以看到多个目标存在时目标编号保持稳定, 在跟踪过程中跟踪目标的编号稳定保持。当 2 号目标黑车消失后, 出现新的目标车辆, 编号能够重新回收利用, 能够及时更新目标信息。在跟踪目标行进过程中多目标相交分离后, 目标编号维持不变。由于在跟踪目标行进过程中智能板卡与跟踪器之间实时进行目标空间信息交互, 可以看到在穿过树、电线杆等遮挡时跟踪器发生了目标丢失, 跟踪器丢失目标后智能板卡提供丢失目标空间信息帮助跟踪器完成目标重捕获, 证明了编批算法的有效性。

4 结束语

为了实现高帧频多目标检测和跟踪任务, 本文提出了基于嵌入式平台的实时多目标编批方法。该方法通过智能板卡与跟踪器交互配合解决了多目标间的相互干扰问题, 在检测算法上加入注意力机制提高网络对目标的检测效率, 并综合考虑重叠面积、中心点距离、长宽比、交并比等多方面因素来实现目标队伍间的数据关联匹配。实验结果表明该方法可以快速实现目标编号建立、编号稳定保持、编号目标的及时更新, 证明了算法的有效性。该方法可以在原有工程基础上进行拓展实现性能提升, 在工程上可以快速实现应用。

参考文献:

- [1] 许剑清. 基于 FPGA+DSP 架构的图像目标跟踪器研制[D]. 哈尔滨: 哈尔滨工业大学, 2018.
- [2] 陈祥涛. 基于 DSP 技术的转台控制与图像跟踪[D]. 西安: 西安工业大学, 2022.
- [3] 罗仁. 基于小型嵌入式平台的实时目标跟踪算法优化和实现方法[D]. 天津: 天津大学, 2020.
- [4] ZUO C, QIAN J M, FENG S J, et al. Deep learning in optical metrology: a review. *Light Sci Appl*, 2022, 39(11).

- [5] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]// Proceedings of the 27th IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Columbus:2014:580-587.
- [6] GIRSHICK R. Fast R-CNN [C]// Proceedings of the 2015 IEEE International Conference on Computer Vision(ICCV). Santiago: IEEE, 2015:1440-1448.
- [7] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence. IEEE, 2017,39(6): 1137-1149.
- [8] LIU W, ANGUELOV D, ERHAN D, et al. SSD:Single shot multi-box detector [C]// Proceedings of the 14th European Conference on Computer Vision. Amsterdam:Springer,2016:21-37.
- [9] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C]// Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas:IEEE,2016:779-788.
- [10] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger [C]// Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE, 2017: 6517-6525.
- [11] REDMON J, FARHADI A. YOLOv3: An incremental improvement [J]. arXiv,2018:1804.02767.
- [12] WU T H, WANG T W, Liu Y Q. Real-time vehicle and distance detection based on improved YOLOv5 network[C]//2021 3rd World Symposium on Artificial Intelligence(WSAI). Guangzhou: IEEE,2021:24-28.
- [13] SONG X Y, GU W. Multi-objective real-time vehicle detection method based on YOLOv5[C]//2021 International Symposium on Artificial Intelligence and its Application on Media (ISA-IAM). Xi'an:IEEE,2021:142-145.
- [14] 王嘉琳. 基于YOLOv5和DeepSORT的多目标跟踪算法研究与应用[D]. 济南:山东大学,2021.
- [15] 张忠民, 吴泽. 基于改进YOLOv5的密集行人检测方法[J]. 应用科技,2023,50(1):33-39.
- [16] Yu J, Jiang Y, Wang Z, et al. UnitBox: An Advanced Object Detection Network[J]. ACM, 2016.
- [17] YANG L X, ZHANG R Y, LI LD, et al. SimAM: A simple, parameter-free attention module for convolutional neural networks [J] ICML 2021 Computer Science,2021.
- [18] WEBB B S, DHRUV N T, SOLOMON S G, et al. Early and late mechanisms of surround suppression in striate cortex of macaque [J]// Journal of Neuroscience,2020(50):11666-11675.
- [19] 姜鑫, 陈武雄, 朱明, 等. 基于实时递推最小二乘的多目标编批研究[J]. 国外电子测量技术,2020,39(2):59-64.
- [20] Durve M, Orsini S, Tiribocchi A, et al. Benchmarking YOLOv5 and YOLOv7 models with DeepSORT for droplet tracking applications[J]. 2023.

第一作者简介: 王 帅 (1993-), 男, 硕士, 研究实习员, 研究领域为目标检测与识别。

※通讯作者简介: 江 山 (1986-), 男, 硕士, 副研究员, 研究领域为图像处理、人工智能。