



面向嵌入式设备的高实时微小目标跟踪检测方法

冒睿瑞,江波

(中国电子科技集团公司第三十二研究所,上海 201808)

摘要:传统主流目标检测算法在嵌入式平台无法兼具高实时性与高准确性,难以应用于边缘智能等领域。为解决微小目标跟踪检测在嵌入式平台实时应用的瓶颈,提出一种高实时微小目标跟踪检测方法。利用轻量化神经网络的骨干网络和路径聚合网络,对整体网络进行针对性的剪枝优化,同时深度融合相关滤波算法,提升针对微小目标跟踪检测的准确度和速度。在3D物体场景渲染器自建的军事微小目标数据集上的实验结果表明,在100像素的极小目标跟踪识别中,与DarkNet53-CSP方法相比,该方法检测精度大幅提高,在400~10 000像素的微小目标识别跟踪中,检测精度与检测速度优于DarkNet53和ResNeXt50+CSP等算法。

关键词:目标检测;嵌入式平台;多路影像;卷积神经网络;滤波算法

开放科学(资源服务)标志码(OSID):



中文引用格式:冒睿瑞,江波.面向嵌入式设备的高实时微小目标跟踪检测方法[J].计算机工程,2021,47(12):291-298.

英文引用格式:MAO R R, JIANG B. Highly real-time method of small target tracking and detection for embedded device[J]. Computer Engineering, 2021, 47(12): 291-298.

Highly Real-Time Method of Small Target Tracking and Detection for Embedded Device

MAO Ruirui, JIANG Bo

(The 32nd Research Institute of China Electronics Technology Group Corporation, Shanghai 201808, China)

[Abstract] Traditional mainstream target detection algorithms cannot provide both high real-time performance and high accuracy on embedded platforms, and thus have limited application in the edge intelligence field. In order to solve this real-time application bottleneck on embedded platforms, a highly real-time small target tracking and detection method is proposed. This method employs the backbone network of a lightweight neural network and a path aggregation network to prune the overall network for optimization. At the same time, it deeply integrates correlation filtering algorithms to improve the accuracy and speed of small target tracking and detection. The method is tested on a self-made small military target dataset built by using the 3D object and scene renderer. The experimental results show that in the case of 100 pixel small target tracking and recognition, the proposed method greatly improves the detection accuracy compared with Darknet53-CSP. In the case of 400~10 000 pixel small target tracking and recognition, the proposed method exhibits higher detection accuracy and speed than DarkNet53 and ResNeXt50+CSP.

[Key words] target detection; embedded platforms; multi-channel image; convolutional neural network; filtering algorithm

DOI: 10.19678/j.issn.1000-3428.0060220

0 概述

近年来,随着嵌入式设备智能化的发展,基于战场末端智能设备需求量不断增加,对其性能要求也愈加严格。其中,战场智能感知作为军事智能化的

关键一环,在各种军事行动如侦察、打击、决策以及毁伤评测中承担重要使命。在现代战争的复杂场景中,微小目标跟踪检测识别技术更贴近现实场景应用需求,相比普通目标识别,一直是目标感知领域中的技术难点^[1]。微小目标在整张图像中所占比重极

基金项目:上海市科学技术委员会科研计划项目“华东计算技术研究所与白俄罗斯国立信息技术及无线电电子大学联合研发与教育中心”(19510750200)。

作者简介:冒睿瑞(1989—),男,工程师、硕士,主研方向为人工智能、图形图像处理;江波(通信作者),研究员、博士。

收稿日期:2020-12-08 **修回日期:**2021-01-12 **E-mail:**jiangbo@ecict.com

小,像素点信息少^[2],当目标背景为复杂的环境时,对于提取微小目标的特征信息存在巨大挑战,其颜色、边缘轮廓等都较为模糊。

面对微小目标检测跟踪的挑战,国内外科研人员基于不同理论开展研究,并提出多种方法。CHEN等^[3]利用上下文模型和小区域提议生成器提出了最新的区域卷积神经网络算法,以提高微小目标检测性能;LI等^[4]提出一种新的感知生成对抗网络模型,该模型通过减少微小目标与正常尺寸目标之间的表示差异来提升微小目标的检测性能;吴言枫等^[5]提出一种动态背景下“低小慢”目标自适应实时检测技术,根据图像的亮度对比度获得显著性图,提取显著性特征的方法为通过形态学梯度,依据图像中目标所在占比的变化,以及质心距离、宽高比等信息筛选真实目标,从而进行微小目标检测。

在面向战场军事需求时,相比于运行高性能人工智能服务器,微小目标跟踪检测技术无法较好地运行在嵌入式设备上,其实际效果不能满足战场实时需求。现有的目标检测管道通常通过学习多个尺度上所有对象的特征表示来跟踪检测微小目标。但是,这种临时架构的性能增益通常仅限于偿还计算成本。当前,基于高性能人工智能服务器的主流移动目标跟踪检测系统仅对视频进行采集、存储和回看,考虑到战场实时性的要求,基于高性能AI服务器的移动目标跟踪检测方法已经不再适用直接移植至嵌入式平台,嵌入式平台的算力无法与高性能服务器相比,但其具有高可靠性和高实时性的优势,这使得面向嵌入式设备的目标跟踪检测技术在军事领域具有广阔的应用发展空间。

针对上述问题,本文提出一种面向嵌入式设备的高实时微小目标跟踪检测方法。该方法利用轻量级卷积神经网络,并融合相关滤波算法,以提升微小目标跟踪检测的精度和速度。

1 相关工作

1.1 目标跟踪检测方法

目标跟踪检测是机器视觉领域的一个重要研究课题,在战场感知领域有重要应用。近年来,目标跟踪检测的研究已经取得较大进步,但在实际过程中目标可能发生形变、剧烈运动、被遮挡等情况,从而导致目标丢失、跟踪失败^[6],因此,目标跟踪的研究仍具有较大挑战性。文献[7]提出一种基于快速多尺度估计的重新检测目标跟踪算法,在相关滤波算法的基础上构建一个新的自适应检测指标,该检测指标可以更加准确地检测出当前帧是否出现跟踪失败的情况。与其他指标不同,新的检测指标减少对最大响应值的依赖。当检测到目标丢失时,通过重新检测的方法寻回目标,恢复跟踪。若当前帧需要进行重新检测,则当前帧的模型拒绝更新,仍采用前

一帧的模型。文献[8]通过最小输出误差平方和(Minimum Output Sum of Squared Error, MOOSE)滤波器进行目标跟踪,当使用单帧图像进行初始化时,产生稳定的相关滤波器,在高速运动跟踪中对光照强弱、尺寸比例、目标姿态和非刚性变形的变化具有鲁棒性。

1.2 神经网络轻量化与优化

深度学习模型的训练和推断耗费大量的计算资源,其中包括内存、人工智能处理器等。当需要执行实时模型推断或在计算资源有限设备端运行模型训练时,造成计算瓶颈。

解决以上问题的有效方法是提高运算效率。模型剪枝优化加速是深度学习领域的一种有效模型压缩方法,可以有效提高推断效率,模型剪枝优化加速可以提高内存使用效率、降低能耗、缩小规模等。文献[9]提出一种基于稀疏卷积神经网络的模型剪枝方法,利用L1正则化在模型训练中的稀疏化作用,对模型的卷积层和批规范化(BN)层参数进行稀疏正则化训练,获得权值稀疏的神经网络模型,再根据滤波器的稀疏性和BN层的特征缩放系数对两者的重要性进行判断,最后用结构化剪枝方法对稀疏滤波器及对应的连接进行剪枝。文献[10]提出一种基于层融合特征系数的动态网络结构化剪枝方法,保证网络的结构性,使剪枝后的网络能结合现有的深度学习库和硬件装置加速方法,并且考虑层与层之间对改变特征图的综合影响,将卷积层和BN层操作表达为全连接形式,提取由滤波器参数、BN层缩放系数及特征图方差多个动态参数组成的层融合特征系数,通过层融合特征系数判定卷积网络每层中滤波器的重要性,结合动态剪枝策略,搜索最优的剪枝网络。

神经网络轻量化与优化的另一种方式是对算法进行针对性的改进,提高计算效率^[11]。例如,文献[12]提出一种基于ShuffleNet和集成随机权向量函数连接网络分类器的人脸识别方法,该方法克服了深度神经网络复杂结构的缺陷,在提高人脸识别准确度的同时,降低深度神经网络训练的复杂度,形成轻量级深层网络人脸识别模型。

2 高实时微小目标跟踪检测方法

由于微小目标的成像尺寸较小,因此只有少量的目标纹理信息,容易受到快速运动、背景、噪声、拖尾等因素的干扰,对检测算法的精度与鲁棒性要求很高^[13]。传统主流方法适用于算力与内存资源充分的高性能AI服务器,然而由于军事武器装备平台的特殊性,战场端的计算平台受空间、供电、散热等限制,只能使用体积小、功耗低、发热小的计算性能有限的嵌入式平台,因此算法的计算速度至关重要。

本文将基于嵌入式设备改进的 CSPDarkNet53 检测算法与基于相关滤波理论的目标跟踪算法相融合,并对整体神经网络进行轻量化改造,提出面向嵌入式设备的高实时目标检测跟踪方法。

2.1 高实时目标检测跟踪算法

当前主流的基于实时视频的目标检测方法是实时视频进行分帧处理,形成一系列连续的图像^[14],针对这些图像单独进行目标检测,最终将检测结果进行合成形成目标检测识别视频。但此类方法对于计算资源的要求高,不适于嵌入式装备有限的计算能力。

视觉目标跟踪是计算机视觉中一个基本问题,根据是否使用帧间关系,有两种解决思路:一种是利用单帧图像信息进行判决,常见方法是使用深度学习卷积神经网络;另一种是通常使用滤波算法与帧间前后关系进行统计运算^[15]。使用深度学习算法能够提取物体关键信息,精度高但运算量也较高。滤波算法计算相较于深度学习开销较小,因而处理速度快,但易受图像质量波动的影响。

本文对于实时视频分帧后形成的一系列图像^[16],按连续若干张形成一组,第1张图像采用面向嵌入式设备改进的 CSPDarkNet53 算法进行目标检测,之后若干张图像采用相关滤波目标跟踪算法对目标进行跟踪识别,相关滤波目标跟踪算法相比 CSPDarkNet53 算法具有优异的性能及良好的实时性,对于计算资源的开销较小,适用于嵌入式设备的高实时目标检测。

相关滤波目标跟踪算法的相关性是指不同事物之间的相关性。它通常用于描述各种事物之间的相似程度,是一种度量。根据共振原理,将相关性引入目标跟踪过程中,通过相关性与目标跟踪算法的结合,使得目标的跟踪通过计算目标模型和候选区域的相关性来达到跟踪目标的目的,所得到的相关性最大位置就是当前帧下目标的位置^[17]。

相关滤波是直接且有效的滤波方法,其在目标跟踪中的最早应用是在 MOSSE 滤波器中,如图1所示,2个目标越相关,相关值越大,即视频帧序列与初始目标越相似,所对应的值也就越大^[8]。

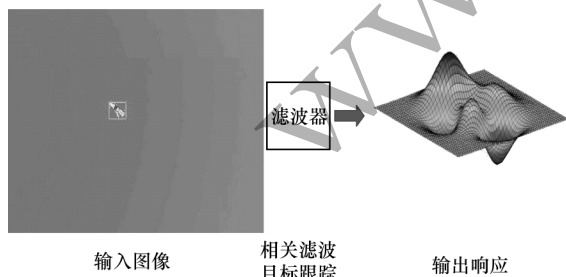


图1 相关滤波算法

Fig.1 Correlation filtering algorithm

由于在时间域进行卷积操作,计算量大且时间较长,因此将其转化成在频率域上的计算可以大幅

减少计算量^[18]:

$$G = F \times M^* \quad (1)$$

其中: G 表示输出图像; F 表示输入图像; M^* 表示滤波器。为在后续帧内找到最相似的目标图像,通过输出的平方误差最小得到最佳的滤波器,目标函数如下:

$$\min_H \sum_i |F_i \times M^* - G_i|^2 \quad (2)$$

通过式(2),可求得滤波器 M^* 的闭合解如下:

$$M = \frac{\sum_i F_i \times G_i^*}{\sum_i F_i \times F_i^*} \quad (3)$$

其中: M 为当前帧的模型; F_i 表示输入图像; G_i^* 表示期望响应的共轭。当前滤波器的模型仅适应于当前帧的模型,为使滤波器具有较强鲁棒性,滤波器的模型根

$$M^* = \frac{A_i}{B_i} \quad (4)$$

A_i 更新如下:

$$A_i = \mu G_i \times F_i^* + (1 - \mu) A_{i-1} \quad (5)$$

B_i 更新如下:

$$B_i = \mu F_i \times F_i^* + (1 - \mu) B_{i-1} \quad (6)$$

其中: μ 是学习率,即当前帧内的滤波器与上个滤波器模型的比值。根据应用需求进行模型的更新,以确保跟踪精度,从而匹配最优的滤波模板。

2.2 下采样算法优化设计

基于嵌入式设备的微小目标检测对模型的精度和速度都有较高的要求,本文基于 CSPDarkNet53 网络模型进行改进,在提高算法速度的同时保证目标识别准确率。

网络的第1次下采样对于微小目标的识别极为重要,如果采用简单的池化或单纯的卷积进行下采样,则无法保证很好地提取出微小目标的特征,导致后面的特征提取丢失微小目标^[19]。传统的下采样一般采用卷积、特征提取的组合,但是这在第1次下采样中带来大量的计算量。因此,本文采用亚像素卷积的逆过程来实现下采样,极大地减少了计算量和内存消耗,同时保持一定的特征提取性能。原始亚像素卷积如图2所示。

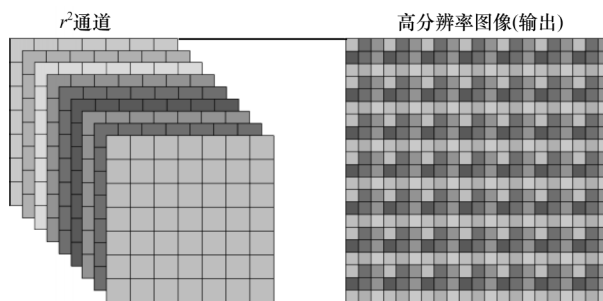


图2 亚像素卷积

Fig.2 Sub-pixel convolution

亚像素卷积可以将 $(C \times r \times r) \times H \times W$ 的输入转为 $C \times (r \times H) \times (r \times W)$,利用这一原理,本文采用其逆过程实现下采样,如图3所示。

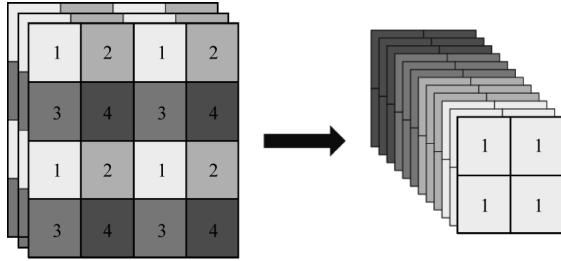


图3 下采样过程

Fig.3 Downsampling process

图3所示的亚像素卷积可以将 $C \times H \times W$ 的输入转化为 $(C \times 2 \times 2) \times (H/2) \times (W/2)$,从而达到下采样的目的。该下采样方法不仅节省大量的计算量,同时也给后面的特征提取保留丰富的细节信息,有利于小目标的特征提取。

2.3 轻量化网络模型

骨干(Backbone)网络一般用于下采样后的进一步特征提取,在丰富特征语义信息的同时防止网络

过深带来的梯度消失、梯度爆炸等问题。网络模型主要采取 CSPDarkNet53 骨干网络,该网络模型深达 161 层,参数以及计算量大,不适用于嵌入式设备上的目标识别。虽然 DarkNet19 网络大幅缩减模型层数和参数量,但是对于微小目标的识别效果不是特别理想。本文提出轻量化网络模型,对骨干网络 CSPDarkNet53 进行深度裁剪,轻量化 CSPResNet 模块,满足本文中嵌入式设备微小目标识别的需求,同时 CSP(Cross Stage Partial)可以兼顾网络轻量化与模型检测的准确性,增强卷积网络的学习能力,降低计算瓶颈以及内存成本。其次,本文采用改进的 PANet 网络,通过多次融合高低层特征提升微小目标检测的效果。

2.3.1 轻量化 CSPResNet 骨干网络

本文基于 CSPDarkNet53 提出一种轻量化的骨干网络,该骨干网络能够减少网络中的模块数量及每个模块中残差单元的数量,其次,减少卷积核参数,提高微小目标的识别时间,适用于嵌入式设备。具体的网络结构如图4所示。其中网络输入图片尺寸为 $1\,024 \times 1\,024 \times 3$ 。

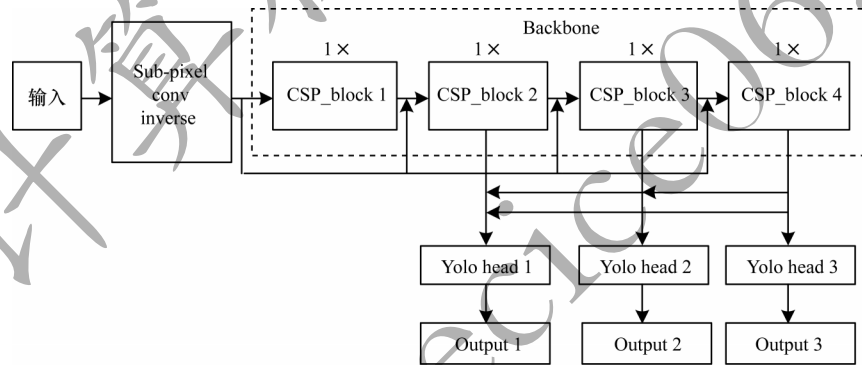


图4 骨干网络结构

Fig.4 Backbone network structure

如图4所示,虚线框中为轻量型骨干网络,首先相比于原来的 CSPDarkNet53 网络,本文极大地减少了 CSP 模块的数量,其次降低 CSP 模块中残差模块的数量,最后削减模块中卷积核的数量,降低网络主干网的参数。

主干网络中共包含 4 个 CSP 模块,这 4 个模块结构相同,如图 5(a)所示,其中卷积 2、3、4 部分为残差模块。输入特征进入 CSP_block 模块后,分成两路进行处理,组成 CSP 结构,在左边处理流程中,卷积 2、3、4 部分为残差模块。不同模块中的卷积核参数如图 5(b)所示。同时,本文将下采样后的底层特征图连接到接下来的每一层 CSP 模块后,可以增强底层特征信息的传输,避免丢失微小目标的少量特征。

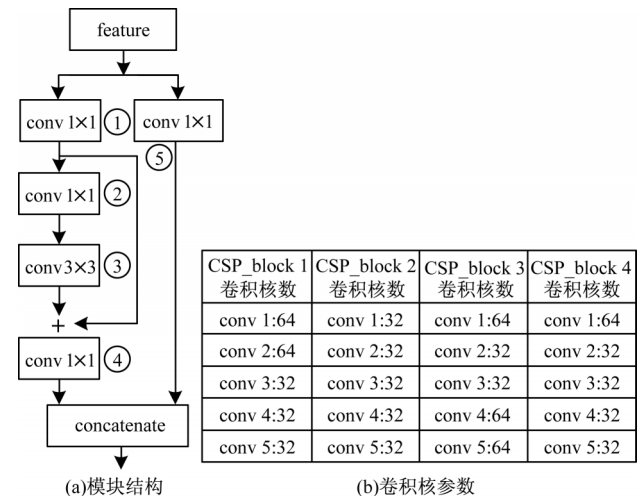


图5 CSP_block 模块

Fig.5 CSP_block module

2.3.2 轻量化路径聚合网络

路径聚合网络(PANet)通过增强自底向上的路径,充分融合浅层与高层之间的特征信息,从而将特征信息中的有用部分传递给子网络,缩短高层特征与底层特征的距离,以适用于小目标检测识别^[20]。但是由于创建多个路径通道,网络计算量较大,不适合部署在嵌入式设备上。本文提出了轻量化的PANet,通过减少原来 CSPDarkNet53 网络中的 Neck 部分,将主干网直接通过 PANet 与 CSPDarkNet53 head 相连接,极大地降低了网络层数和计算量,同时高低层特征通过多种路径相连接,强化了小目标检测效果,具体的轻量化 PANet 结构如图 6 所示。

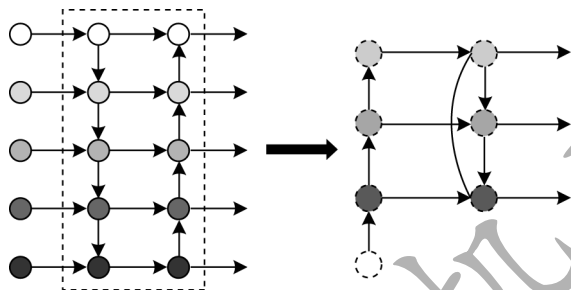


图6 轻量化 PANet 结构

Fig.6 Lightweight PANet structure

2.3.3 剪枝优化

由于神经网络模型计算量较大,在嵌入式设备上运行速率较低^[21],因此通过采用模型剪枝的方式,对模型中不必要的计算进行删除,从而简化模型来提高运行速度。对于裁剪网络的方式,其主要是判断每个卷积核对整个网络的贡献程度,本文主要利用深度神经网络中的批量标准化(BN)层中的参数 γ ,确定剪枝参数,以决定每一层中剪除的具体通道^[22]。BN层的相关参数加上 L1 正则化惩罚来训练模型,其新的损失函数如下:

$$L = \sum_{(x,y)} l(f(x,W),y) + \lambda \sum_{\gamma} g(\gamma) \quad (7)$$

利用小批量统计的方式规范化 BN 层内部激活函数,BN 层转换如下:

$$\hat{z} = \frac{z_{in} - \mu\beta}{\sqrt{\sigma_{\beta}^2 + \theta}}, z_{out} = \gamma\hat{z} + \beta \quad (8)$$

分析基准网络的每一层权重,采用基于判断权重值的大小判定其重要性来进行裁剪,对于过滤器,将其所有权重的绝对值求和,作为评价该过滤器的指标,将值低的过滤器裁掉,保留值较高的过滤器,以达到降低模型复杂度的目的;在裁剪的过程中,同时考虑每一层卷积核对于剪枝的敏感度,对于裁剪比较敏感的卷积层,对其裁剪程度适当减小;当删除部分的卷积核后,输出层的通道数发生变换,删除其对应的通道;在进行剪枝操作后,与原模型对比精准度,通过采用较小的学习率微调剪枝

后的模型以回升精度,具体模型剪枝步骤如图 7 所示。



图7 模型剪枝步骤

Fig.7 Model pruning steps

本文对整个网络进行排序,然后自行确定剪枝比例,去除参数值较小的通道,最后进行微调,重复操作,以到达最优的剪枝效果。

3 实验结果与分析

3.1 实验配置

本文实验硬件环境为 Jetson AGX Xavier 开发套件板,同时使用 JetPack SDK4.3 软件套件,主要软件环境为 L4T32.2、CUDA10.0、CUDNN7.6、OPENCV4.4。本文推理程序主要使用 C/C++ 编写,采用多种图像预读及处理多线程方法来提高程序执行效率,且使用平台最大电源功耗模式来考验嵌入式平台下极限推理性能。

本文实验数据采用三维物体场景渲染器自建军事场景进行图像采集,通过场景渲染器修改渲染大小来模拟现实场景目标,实验选用 5 种尺度目标大小考验算法的识别能力,如图 8 所示。数据集图像中分别包含 100 像素、400 像素、900 像素、2 500 像素和 10 000 像素的待测目标。



图8 微小目标数据集示意图

Fig.8 Schematic diagram of tiny target dataset

训练集内含 1 024 像素×1 024 像素分辨率的图片 10 000 张,验证集内含图片 2 000 张。本文算法主要考验对单类微小目标的识别和跟踪能力,因而采用目标检测中不同的交并比(IoU)阈值下的精确度

AP 作为准确度评价指标,AP25、AP50 和 AP75 表示在 IoU 阈值 25、50 和 75 下的检测成功率,采用推理速度 FPS,即推理时间倒数作为性能评价指标。

模型训练及前期预处理使用高性能服务器,搭载 4 块 NVIDIA Tesla P100 加速卡。所有的网络模型将其动量设为 0.9,学习率设为 0.002 6,权重衰减设为 0.000 5,并进行 10 000 次迭代,得到最终的训练结果。通过设定不同的稀疏策略进行训练,选出综合最优的精度和稀疏度,然后利用通道剪枝以及层剪枝分别压缩模型的宽度和深度,最后微调剪枝后的模型以回升精度,从而达到最优的模型剪枝效果。

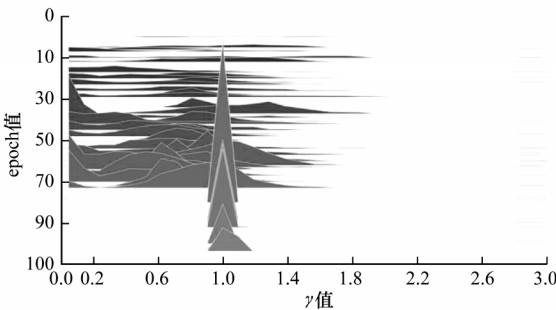
本次实验对比多个主流的目标检测网络,包含若干个单步法快速检测网络,如 DarkNet53、DarkNet19、CSP 改进的 DarkNet53 和 ResNeXt50+CSP。并且,本文将所提模型与剪枝后的模型分开进行对比,能够进一步验证剪枝优化的效果。

3.2 结果分析

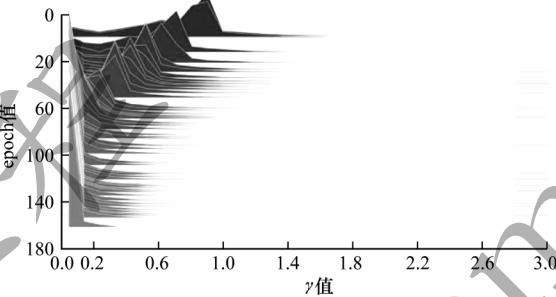
第 1 组实验使用图像像素尺寸为 1 024 像素×1 024 像素,对比 5 种不同像素尺度的微小目标检测准确率,基准为本文所提的轻量化检测网络,基准裁剪为裁剪优化过后的网络。

本文首先使用所提轻量网络训练一个基准模型,如图 9(a)所示,从图中可以看出基准模型的 γ 值大致服从期望为 1 的正态分布;然后使用稀疏训练出基准裁剪的模型,如图 9(b)所示,可以明显看出,经过稀疏过程,BN 层的 γ 值大部分逐渐被压到接近 0,而接近 0 的通道其输出值近似于常量,因而可以进行后续剪枝操作。

利用层剪枝以及通道剪枝策略,模型大小从 23.4M 压缩到 2.6M,在 900 像素目标下,AP 值减少 0.23,通过反复微调,AP 值回升到 80.05,基准裁剪的模型大幅压缩减少资源占用,提高了运行速度。



(a)压缩前 γ 系数分布



(b)压缩后 γ 系数分布

图 9 剪枝前后 BN 层 γ 值稀疏分布变化

Fig.9 The sparse distribution of γ values in BN layer before and after pruning

如表 1 所示,在极小目标情况下(100 像素),传统的快速单步法检测网络基本无法正常工作,改进后的 DarkNet53+CSP 依然表现很差,而本文所提卷积网络具有一定概率能够成功识别;在稍大一些的目标 400 像素~2 500 像素下,能够看出本文所提模型在识别能力上均领先对照模型,而改进的 DarkNet53 和 ResNeXt50+CSP 也比基础的算法有明显优势;但在很大目标识别上,由于 DarkNet53 和 DarkNet19 均缺乏多尺度特征提取的能力,预测精度变得很差,而本文的模型保证了很高的识别成功率。

表 1 多种算法的小目标检测精度对比

Table 1 Comparison of tiny target detection accuracy of multiple algorithms

目标尺寸/像素	指标	DarkNet53	DarkNet19	DarkNet53+CSP	ResNeXt50+CSP	基准	基准裁剪
100	AP ₂₅	1.33	0.00	2.31	1.82	19.45	18.23
	AP ₅₀	0.00	0.00	1.52	0.81	10.70	10.53
	AP ₇₅	0.00	0.00	0.71	0.00	7.83	7.61
400	AP ₂₅	27.56	18.98	32.12	29.76	56.46	49.81
	AP ₅₀	17.32	14.45	25.93	22.22	42.11	40.44
	AP ₇₅	7.70	5.55	18.45	12.24	28.20	25.50
900	AP ₂₅	56.40	67.64	79.41	71.42	83.36	80.05
	AP ₅₀	34.84	41.23	48.24	37.21	51.18	47.54
	AP ₇₅	21.18	29.46	34.79	26.65	40.04	38.29
2 500	AP ₂₅	56.24	81.25	85.49	80.52	86.56	87.42
	AP ₅₀	29.92	58.40	63.23	47.18	64.94	65.92
	AP ₇₅	18.46	39.82	40.21	22.82	42.80	45.95
10 000	AP ₂₅	8.32	78.92	90.92	78.84	92.90	94.82
	AP ₅₀	7.50	34.23	72.12	55.87	78.52	83.12
	AP ₇₅	2.40	17.90	59.34	32.52	69.52	71.56

如表2所示,本文测试了算法在多种图像输入尺寸下嵌入式平台的实时推理性能。在模型输入尺寸较大的情况下,推理延迟普遍较高,难以满足高实时高负载应用场景的性能需求,而本文所提轻量化模型经过裁剪有27.8 frame/s的速度,大幅领先对比算法。当卷积网络规模变小时,推理速度普遍提高,且本文模型均保持性能较优。

表2 多种图像输入尺寸下推理性能比较

Table 2 Comparison of inference performance under various image input sizes (frame·s⁻¹)

图像尺寸/ 像素	Dark- Net53	Dark- Net19	Dark- Net53+ CSP	ResNeXt50+ CSP	基准	基准 裁剪
1 024×1 024	5.4	9.2	4.6	4.5	10.1	16.8
608×608	9.0	13.2	7.8	7.2	18.0	27.8
416×416	16.8	22.0	15.2	14.7	24.0	45.1
256×256	24.8	26.0	23.2	22.8	27.0	52.9

针对特定现实场景应用,如在中高分辨率尺寸场景下,若在嵌入式平台进行多路图像处理,因其运算量成倍提高,则目前算法推理性能均难以满足高实时要求。本文采用的滤波跟踪补帧方法能够大幅提升等效推理性能。如表3所示的实验验证了基准裁剪算法在608像素×608像素输入下,帧间隔对推理速度和精度的影响:推理帧间隔为0时为原始算法性能,而当补帧间隔为2时,能够达到性能与精度的最佳平衡。

表3 推理帧间隔对性能和精度的影响

Table 3 The effect of inference frame interval on performance and accuracy

推理间隔	裁剪速度/(frame·s ⁻¹)	AP ₅₀ /%
0	27.8	83.12
1	57.0	79.82
2	86.0	74.64
3	124.0	57.19

4 结束语

针对嵌入式平台的算力瓶颈导致微小目标跟踪检测算法的指标性能不能满足现实战场需求的问题,本文提出一种面向嵌入式的高实时微小目标跟踪检测方法。将卷积神经网络进行轻量化改进,对模型进行剪枝优化,并深度融合相关滤波跟踪算法,以提高基于嵌入式平台的微小目标检测速度和精度。实验结果表明,该方法在Nvidia Jetson AGX Xavier嵌入式平台的微小目标跟踪检测中,检测精度和检测速度都优于目前主流目标检测算法。

参考文献

- [1] WOMG A, SHAFIEE M J, LI F, et al. Tiny SSD: a tiny single-shot detection deep convolutional neural network for real-time embedded object detection [C]//Proceedings of the 15th Conference on Computer and Robot Vision. Washington D. C. , USA: IEEE Press, 2018: 95-101.
- [2] 邵宇辉. 基于背景建模的微小运动目标检测算法研究[D]. 西安: 西安理工大学, 2019.
SHAO Y H. Research on small moving target detection based on background modeling [D]. Xi' an: Xi' an University of Technology, 2019. (in Chinese)
- [3] CHEN C, LIU M Y, TUZEL O, et al. R-CNN for small object detection [C]//Proceedings of ACCV'16. Berlin, Germany: Springer, 2016: 214-230.
- [4] LI J, LIANG X, WEI Y, et al. Perceptual generative adversarial networks for small object detection [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C. , USA: IEEE Press, 2017: 1222-1230.
- [5] 吴言枫, 王延杰, 孙海江, 等. 复杂动背景下的“低小慢”目标检测技术[J]. 中国光学, 2019, 12(4): 854-866.
WU Y F, WANG Y J, SUN H J, et al. LSS-target detection in complex sky backgrounds [J]. Chinese Optics, 2019, 12(4): 854-866. (in Chinese)
- [6] 葛宝义, 左宪章, 胡永江. 视觉目标跟踪方法研究综述[J]. 中国图象图形学报, 2018, 23(8): 1091-1107.
GE B Y, ZUO X Z, HU Y J. Review of visual object tracking technology [J]. Journal of Image and Graphics, 2018, 23(8): 1091-1107. (in Chinese)
- [7] 胡昭华, 黄嘉净. 基于快速多尺度估计的重新检测目标跟踪算法[J]. 微电子学与计算机, 2020, 37(9): 11-17.
HU Z H, HUANG J J. Re-detection target tracking algorithm based on fast multi-scale estimation [J]. Microelectronics & Computer, 2020, 37(9): 11-17. (in Chinese)
- [8] BOLME D S, BERERIDGE J R, DRAPER B A, et al. Visual object tracking using adaptive correlation filters [C]//Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington D. C. , USA: IEEE Press, 2010: 2544-2550.
- [9] 周勇. 改进的卡尔曼滤波在目标跟踪中的应用[D]. 哈尔滨: 哈尔滨工程大学, 2016.
ZHOU Y. Application of the improved Kalman filter in target tracking [D]. Harbin: Harbin Engineering University, 2016. (in Chinese)
- [10] 韦越, 陈世超, 朱凤华, 等. 基于稀疏卷积神经网络的模型剪枝方法[J]. 计算机工程, 2021, 47(10): 61-66.
WEI Y, CHEN S C, ZHU F H, et al. Model pruning method based on sparse convolutional neural network [J]. Computer Engineering, 2021, 47(10): 61-66. (in Chinese)

- [11] TRONG T D, NGUYEN V T, TRAN M T. Lightweight deep convolutional network for tiny object recognition[C]//Proceedings of ICPRAM'18. Washington D. C. , USA: IEEE Press, 2018:675-682.
- [12] 杨德胜, 马冬. 基于轻量级深层神经网络的人员安全管控平台研究[J]. 电力信息与通信技术, 2019, 17(6): 1-7.
- YANG D S, MA D. Research on personnel safety management and control platform based on lightweight deep neural network[J]. Electric Power Information and Communication Technology, 2019, 17(6): 1-7. (in Chinese)
- [13] 何洋. 时空域快速自适应视频前景检测与去噪算法研究[D]. 西安: 西安电子科技大学, 2017.
- HE Y. Research on fast adaptive video foreground detection and denoising algorithm in spatio-temporal domain[D]. Xi'an: Xidian University, 2017. (in Chinese)
- [14] 卢海伟, 袁晓彤. 基于层融合特征系数的动态网络结构化剪枝[J]. 模式识别与人工智能, 2019, 32(11): 1051-1059.
- LU H W, YUAN X D. Dynamic network structured pruning via feature coefficients of layer fusion [J]. Pattern Recognition and Artificial Intelligence, 2019, 32(11): 1051-1059. (in Chinese)
- [15] 尹宏鹏, 陈波, 柴毅, 等. 基于视觉的目标检测与跟踪综述[J]. 自动化学报, 2016, 42(10): 1466-1489.
- YIN H P, CHEN B, CHAI Y, et al. Vision-based object detection and tracking: a review [J]. Acta Automatica Sinica, 2016, 42(10): 1466-1489. (in Chinese)
- [16] 吴天舒, 张志佳, 刘云鹏, 等. 基于改进SSD的轻量化小目标检测算法[J]. 红外与激光工程, 2018, 47(7): 47-53.
- WU T S, ZHANG Z J, LIU Y P, et al. A lightweight small object detection algorithm based on improved SSD [J]. Infrared and Laser Engineering, 2018, 47(7): 47-53. (in Chinese)
- [17] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C. , USA: IEEE Press, 2018: 8759-8768.
- [18] 赵杰. 基于相关滤波的多策略融合运动目标跟踪算法研究[D]. 昆明: 昆明理工大学, 2019.
- ZHAO J. Research on multi-strategy fusion moving target tracking algorithm based on correlation filtering [D]. Kunming: Kunming University of Science and Technology, 2019. (in Chinese)
- [19] 王明丽, 王刚, 郭晓新, 等. 基于跨层复制连接卷积神经网络的遥感图像融合[J]. 吉林大学学报(理学版), 2020, 58(4): 913-922.
- WANG M L, WANG G, GUO X X, et al. Remote sensing image fusion based on cross layer copy connection convolutional neural network[J]. Journal of Jilin University Science Edition, 2020, 58(4): 913-922. (in Chinese)
- [20] 林成创, 赵淦森, 尹爱华, 等. AS-PANet: 改进路径增强网络的重叠染色体实例分割[J]. 中国图象图形学报, 2020, 25(10): 2271-2280.
- LIN C C, ZHAO J S, YIN A H, et al. AS-PANet: improved overlapping chromosome instance segmentation of path enhancement network[J]. Journal of Image and Graphics, 2020, 25(10): 2271-2280. (in Chinese)
- [21] 杨元挺. 基于粒子滤波的灰度目标跟踪算法[J]. 厦门大学学报(自然科学版), 2012, 51(1): 33-36.
- YANG Y T. A gray object tracking algorithm based on particle filter[J]. Journal of Xiamen University (Natural Science), 2012, 51(1): 33-36. (in Chinese)
- [22] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift [C]//Proceedings of the 32nd International Conference on Machine Learning. Washington D. C. , USA: IEEE Press, 2015: 448-456.