

面向巡检机器人的电力设备状态检测算法研究

杨学杰¹, 李思毛¹, 李建业¹, 李宋林¹, 张旭²

(1. 国网淄博供电公司, 山东 淄博 255000;

2. 国网智能科技股份有限公司, 山东 济南 250101)

摘要:为更高效利用变电站巡检机器人开展电力巡检工作,满足电力行业发展对智能化巡检的需求,研究了面向电力巡检机器人的电力设备状态检测算法。首先,根据深度网络部署硬件芯片应用情况与性能对比,选择海思 Hi3559A 芯片作为算法移植的嵌入式平台。然后综合考虑各种检测算法的精度与速度,选用 YOLOv3 算法作为设备状态检测的基本判别模型。为了提升检测算法速度并减少模型体积,开展模型压缩算法及轻量型 YOLOv3 模型设计研究,分别提出了改进的小型化 YOLOv3 模型和基于通道剪枝与层剪枝结合的模型压缩方法,提高模型上下层的语义信息及剪枝后模型的精度保持。根据测试结果选择最优的模型在机器人前端部署,提出的轻量化 YOLOv3 模型很好地保持了设备目标与异物检测的精度,检测速度提升了 4 倍。

关键词:变电站巡检机器人;目标检测;YOLOv3;模型压缩;通道剪枝;层剪枝

中图分类号:TP391.41

文献标识码:A

文章编号:1673-629X(2021)03-0201-05

doi:10.3969/j.issn.1673-629X.2021.03.035

Research on Algorithm for Equipment Condition Monitoring Based on Inspection Robot

YANG Xue-jie¹, LI Si-mao¹, LI Jian-ye¹, LI Song-lin¹, ZHANG Xu²

(1. State Grid Zibo Power Supply Company, Zibo 255000, China;

2. State Grid Intelligent Technology Co., Ltd., Jinan 250101, China)

Abstract: In order to implement power equipment inspection work using substation inspection robot efficiently and meet the demand of the development of electric power industry for intelligent inspection, the status detection algorithm of electric equipment for electric power inspection robot is researched. Firstly, Hi3559A chip is chosen as an embedded development platform to implement arithmetic transplant according to the comparison of properties and application between different hardware chips. Taking the accuracy and speed of various detection algorithms, YOLOv3 algorithm is selected as the basic discriminant model of device state detection. In order to improve the speed of detection algorithm and reduce model size, the model compression algorithm and the lightweight YOLOv3 model are designed and studied. Both improved tiny YOLOv3 and model compression algorithm based on channel pruning and layer pruning are proposed to improve semantic information extraction and holding accuracy of original model. The optimal model is chosen to deploy on the hardware platform according to the test results. The experiment demonstrates that the proposed lightweight YOLOv3 model maintains the detection accuracy of equipment targets and foreign bodies well, and the detection speed is increased by 4 times.

Key words: substation inspection robot; object detection; YOLOv3; model compression; channel pruning; layer pruning

0 引言

目前,变电站巡检机器人已在国内变电站推广应用,利用携带的红外、可见光传感器实现设备温度自动测量、设备状态自动识别,自动代替人工完成变电站设备的巡检任务,可根据设备综合状态综合诊断设备故障^[1]。经过多年的机器人巡检技术和巡检模式研究,“机器人巡检+数据后台分析”模式已经成为变电站智

能运维的重要趋势。

利用巡检机器人采集图像可实现电力设备状态自动识别(表计读数、刀闸开合状态、断路器开关等)与设备缺陷自动诊断(鸟巢、绝缘子破损等),以上任务可以利用人工智能领域的目标检测技术完成^[2-3]。目前常用的目标检测算法主要包括 RCNN 系列网络^[4-6]、SSD 网络^[7]、RFCN 网络^[8]与 YOLO 系列网络

收稿日期:2020-04-03

修回日期:2020-08-06

基金项目:山东省电力科技项目(5206031900HW,2020A-108)

作者简介:杨学杰(1970-),男,工程师,研究方向为电气工程自动化。

模型^[9-11]等。YOLOv3 与其他目标检测算法相比,在检测准确率和速度上,均取得了较好的效果,并在电力领域得到了广泛应用^[12-13]。

尽管变电站巡检机器人及人工智能技术的应用提升了电力巡检的自动化和智能化水平,但基于图像的设备状态分析算法有效性、可靠性仍需进一步提高,其高存储、实时性差的问题仍然制约着机器人智能巡检的推广应用。随着深度神经网络模型日益深度化、复杂化,以上问题制约着人工智能技术在计算资源有限的机器人应用环境中的推广,如何在保证算法精度的前提下,实现深度模型的小型化和前置化是亟待解决的一个问题。

为了提升变电站巡检机器人智能分析能力和设备状态诊断效率,该文对电力设备状态检测前端化算法进行研究,以 YOLOv3 模型为基础网络,对模型结构进行剪枝压缩及整体结构轻量化改进,实现变电站设备定位及异物悬挂等典型缺陷检测的边缘计算,加强巡检机器人的图像分析、判别和智能决策能力,进一步提升电力巡检的效率及智能化水平。

1 设备状态自动检测算法前置化研究

实现电力设备状态检测算法在机器人前端部署,

表 1 Jetson TX2 与 Hi3559A 芯片对比

	Jetson TX2	Hi3559A
配置	GPU: NVIDIA Pascal, 256 CUDA cores CPU: 四核 ARM Cortex -A57 处理器 4G Bytes LPDDR4 16G Bytes eMMC	GPU: 双核 ARM Mali G71 @ 900 MHz CPU: 双核 ARM Cortex A73 @ 1.8 GHz + 双核 ARM Cortex A53 @ 1.2 GHz + 单核 ARM Cortex A53 @ 1.2 GHz 4G Bytes LPDDR4 8G Bytes eMMC
性能	可搭载 4G 通信、wifi 和 bluetooth 模块 支持 4k×2k 30 Hz 编码 (HEVC), 4k×2k 60 Hz 解码 (10 位) 支持 Linux 操作系统 单精度浮点运算速度 2TOPS	8K@30+1080P@30 H.265 编码下典型功耗 3W 支持 8K@30+1080P@30 或者 4K@120+1080P@30, H.265 编码 支持 2×4K@60 或 4×4K@30 或 8×1080P@30 视频录制, 支持机内硬化拼接 提供视觉计算处理能力, 算力可达 4TOPS
价格	¥4 000 左右	¥2 000 左右

综合考虑两种嵌入式 AI 平台的计算能力与生产成本,海思 Hi3559A 硬件平台在成本及算力上均优于 TX2,因此选用海思 Hi3559A 作为嵌入式平台进行深度学习检测算法在机器人前端部署。

1.2 设备状态自动检测算法

面向机器人的设备状态自动检测算法实现主要分成三个部分:(1)各类设备及缺陷的图像样本采集与标注,建立图像样本库;(2)利用目标检测模型对样本进行训练,得到设备状态自动检测判别模型;(3)将训练好的模型部署在机器人嵌入式硬件上,根据输入的

选择合适的嵌入式硬件平台进行检测算法的移植是首要前提。另外,为了满足模型部署需求,选择精度高、耗时少的目标检测算法,实现设备状态实时检测。

1.1 算法模型的嵌入式部署

目前深度学习网络的使用基本上在 GPU 平台下进行,其在嵌入式移动设备下的应用主要依靠 FPGA 硬件,同时专用于 DNN 的芯片研究也正在开展。在现有机器人设备的基础上,根据目前众多的深度网络部署硬件芯片应用情况与性能对比,主要考虑两种嵌入式硬件设备:英伟达 Jetson TX2 与海思 Hi3559A。

英伟达 Jetson TX2 是由英伟达公司研发的 Jetson 系列 GPU 嵌入式平台,主要用于深度学习模型的研究,其核心嵌入式组件体积与信用卡一般大小,可以嵌入到移动设备中使用,算法移植方便,但其属于进口芯片,价格较高,不适合机器人产品小型化、批量化生产。海思 Hi3559A 芯片属于国产硬件平台,包含了深度学习的 NNIE 内核,定制平台算力可达 4TOPS (Tera operations per second),是目前同等算力下性价比最好的平台之一。两款嵌入式平台的性能与硬件参数对比如表 1 所示。

巡检实时图像输出判别结果。

综合考虑当前硬件平台计算能力与巡检图像分析任务,巡检机器人采用前端检测与后台分析相结合的方式实现变电站设备状态的自动检测,将一些识别率高、需要现场确认的缺陷检测功能放在机器人前端实现,主要包含仪表读数、刀闸状态、绝缘子缺陷、鸟巢及异物四类目标的检测。

当前的目标检测算法根据设计差异总体上分为两种:基于候选区域的目标检测算法 (RCNN 系列算法) 和基于回归的目标检测算法 (SSD、YOLO 系列等)。

与以 Faster RCNN^[6] 为代表的基于区域的目标检测算法相比, YOLOv3 算法可以实现端到端的训练, 减少了模型训练的复杂度, 并且图像的检测速度大幅度提高。与 SSD 算法^[7] 相比, YOLOv3 拥有更高的检测精度以及更快的处理速度, 部署在嵌入式平台上更接近设备状态检测实时处理的需求, 因此选用 YOLOv3 算法作为电力设备状态检测的深度模型, 不同算法的检测结果如表 2 所示。

表 2 YOLOv3 与其他网络性能对比

方法	AP/%	单张耗时/ms
Faster RCNN	66.7	793
SSD	69.4	265
R-FCN	71.2	342
YOLOv1	71.7	257
YOLOv2	73.4	211
YOLOv3	75.6	145

YOLOv3 采用 Darknet53 的网络结构, 共包含 53 个卷积层, 训练时最大训练迭代次数设置为 100 000, 学习率为 0.001, 训练得到的参数量约为 60 M, 模型体积在 300 M 左右, 单张图像检测时间大于 100 ms, 目标检测速度和模型体积不适合往机器人的前端移植, 因此需要研究深度模型的轻量化算法, 在保证检测精度的基础上, 减少判别网络的计算量与模型体积。

2 深度模型的轻量化算法

由于机器人搭载的硬件计算能力与后台服务器的

计算能力无法比拟, 为了实现设备检测算法的前置化, 部署模型需要少量的算法参数, 较少的内存及运算时间, 因此这里将原始 YOLOv3 模型进行轻量化改进。一种方法是对网络模型进行轻量化设计, 精简模型结构, 减少模型参数; 另一种方法是进行模型压缩, 去除模型卷积层中不必要的通道。

2.1 小型 YOLOv3 模型

小型 YOLOv3 模型 (Tiny YOLOv3, T-YOLOv3)^[11] 是 YOLOv3 模型的简化, 保留了原始模型两个独立预测分支, 该网络模型主要是由卷积层和池化层构成, 结构更简单, 但精度有所下降。

T-YOLOv3 模型引入了多尺度特征提取模块, 输出 2 个尺度的特征图来保证对不同大小目标的检测性能。虽然这种模型结构可以减少卷积特征层的参数, 提高模型运算速度, 但在环境复杂的变电站场景下, 这种网络结构容易造成浅层卷积特征图信息丢失, 并且对小目标及局部特征的目标输出不敏感, 很容易对远处的小目标设备或存在遮挡的设备造成漏检和误检, 影响了目标的召回。

为了提高模型上下层的语义信息及小目标的检测能力, 对 T-YOLOv3 的网络模型进行改进, 增加两个卷积层和池化层, 提高网络提取特征对变电站目标的表征, 并在原有网络结构的基础上增加一层上采样层, 为了避免训练中的梯度消失并增强特征的复用, 在网络的分类回归之前引入残差模块, 改进后的网络记为 E-T-YOLOv3, 其网络结构如图 1 所示。

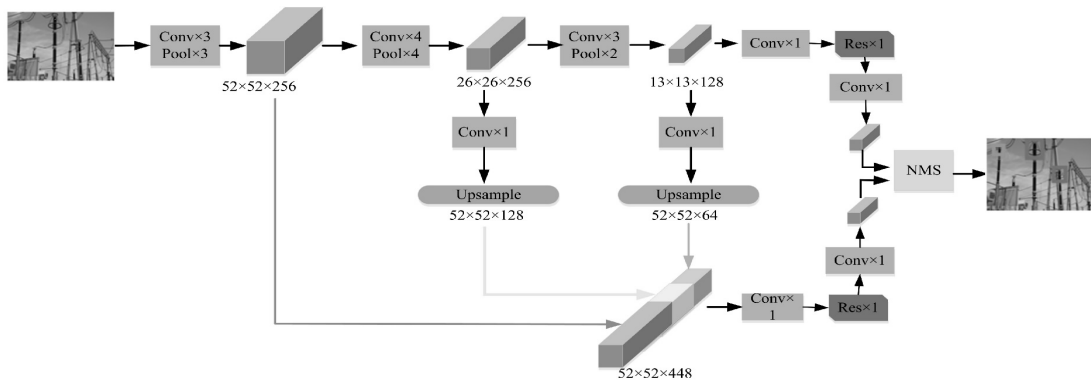


图 1 E-T-YOLOv3 模型结构

2.2 基于通道剪枝与层剪枝的模型压缩算法

为了实现 YOLOv3 模型的压缩, Zhang 等人提出了一种基于通道剪枝的 YOLOv3 模型压缩算法^[14], 并将压缩后的模型部署在无人机上, 取得了理想的效果。首先, 该算法在原始模型的基础上添加了三个空间金字塔池化 (spatial pyramid pooling, SPP) 模块, SPP 模块可以提取不同尺度的特征信息, 并将不同尺度的特征图以串联的方式进行特征融合, 提高原始网络的检测精度, 构成模型修剪的基本模型。接着, 为了删除模

型不重要的通道, 提出了 SlimYOLOv3-SPP3-50/90/95 三种型号的修剪模型, 此种模型修剪可以极大改善模型在机器人上的部署, 但具有一定的精度损失。具体剪枝步骤为:

(1) 对模型进行稀疏训练, YOLOv3 获得每个信道的比例因子, 评估模型中每个信道的重要性, 然后去除那些具有小比例因子的信道。

(2) 通过修剪得到的模型 SlimYOLOv3 在数据集上进一步微调获得检测结果, 对剪枝后的模型进行效

果评估,然后进入下一轮稀疏训练。

(3)迭代重复前两步通道剪枝过程,直到满足设定的模型剪枝比或检测速度要求。

此类通道自动剪枝策略对于 YOLOv3 残差块中的卷积网络具有很好的压缩效果,减少了整体网络的宽度,但是当冗余的参数主要分布在网络的深层次时,即使达到模型剪枝比停止训练,实际的加速往往达不到理想的效果,并带来一定的精度损失。为了进一步提高模型的压缩质量,在通道剪枝的基础上,引入了层剪枝策略,在保持模型精度的前提下,使得模型在宽度和深度上得到有效剪枝。层剪枝的策略主要针对 YOLOv3 模型的直连(shortcut)层进行剪枝,具体步骤为:

(1)对于每个直连层前一个基础模块进行重要性评价,基础模块包括一个 1×1 与 3×3 的卷积核,层重要性评价系数选用 Liu 等人提出的缩放因子 γ 系数^[15]。

(2)由于每裁剪一个 shortcut 层,会同时减掉与其相连的两个卷积层,为保证 YOLOv3 的结构完整,设置层剪枝 shortcut 剪枝数量阈值为 12,即剪枝层数不超过 36 层,精度保持阈值为 7%,即剪枝后模型的平均精确度(average precision, AP)下降不超过原模型 AP 值的 7%。

(3)根据计算的 γ 系数对各直连层进行排序,在满足阈值的前提下取最小的层进行剪枝。

这里将 SlimYOLOv3-SPP3-50 通道剪枝策略与层剪枝策略结合使用,整体过程如图 2 所示。待剪枝完成后,对得到的轻量模型再进行参数的微调。以剪枝后的模型进行训练参数的初始化,然后对模型进行重训练,训练中将最大训练迭代次数减小到 50 000,并在迭代 30 000 次后将学习率降低为原学习率 1/5。通过微调,最终得到压缩后的深度判别模型。

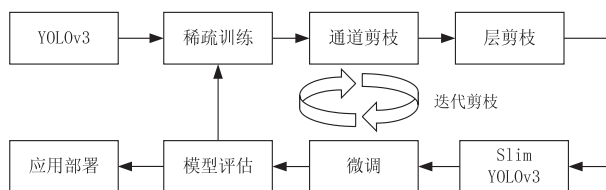


图2 通道剪枝与层剪枝模型压缩过程

3 实验结果与分析

为了验证轻量化模型的设备状态检测性能,首先对该文提出的算法训练模型与原始 YOLOv3 及其他轻量化算法进行对比实验。不同算法验证实验环境:CPU 为 E5-1607,内存为 16 GB, GPU 为 NVIDIA GeForce 1080TI,操作系统为 Ubuntu16.04。

该文利用变电站巡检图像作为数据集,图像分辨

率为 $1\,920 \times 1\,080$,数据集共有 4 类对象,包含仪表、刀闸、绝缘子、鸟巢及悬挂物。每类图像包含样本 1 000 张,其中 800 张图像作为训练验证集,200 张图像作为测试集。分别利用不同的模型算法对 4 类目标进行训练及测试,验证指标采用 AP 值及平均交并比(Intersection over union, Iou),AP 值越高表示模型在全部类别找那个目标的综合性能最高,Iou 值越高代表检测到的目标位置越精确。除此之外,对训练模型的大小及 4 类目标的平均检测速度进行比较,模型越小,检测速度越快,越适合嵌入式平台前端进行部署。

首先,为了确定最优层剪枝直连层的剪枝个数,设置不同的剪枝个数进行压缩模型的性能测试,如图 3 所示。横坐标“比率”为层剪枝后模型相对于原模型的体积与 AP 值变化,从图中可以看出,当剪枝直连层个数为 2 时,精度保持效果最好,但模型体积压缩较少。其中当剪掉层数为 1,2,4,6,8 时,满足设定剪枝层数与精度保持阈值的要求。为了兼顾精度与模型体积,选择剪枝直连层的个数为 8,模型共减掉 24 层,精度下降了 4.3%,模型体积仅有 31.3 M,压缩了一半以上。

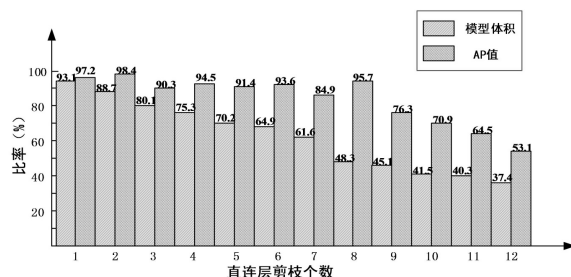


图3 不同层剪枝个数的模型压缩后性能

不同模型对 4 类目标的测试结果如表 3 所示,与原始 YOLOv3 模型相比,轻量化后的模型在精度上都有不同程度的下降,模型体积与单张图像检测耗时间都有明显提升,T-YOLOv3 模型检测速度最快,但精度下降比较严重,改进后的 E-T-YOLOv3 模型 AP 值提升了 4.1%。基于 SlimYOLOv3-SPP3 通道剪枝的策略当压缩比过高时精度下降比较严重。综合检测精度与速度因素,“SlimYOLOv3-SPP3-50+层剪枝”策略压缩后的模型检测速度降低至 37 ms/张,AP 值仅下降了 2.4%,Iou 下降了 0.08,完全适应机器人目标检测前端部署需求。另外,PR 曲线反映了精确率(precision)与召回率(recall)之间的关系,更能直观地看出算法间性能的好坏,精简模型与原始模型的 PR 曲线如图 4 所示。与原始 YOLOv3 模型相比,模型剪枝后的 YOLOv3 模型与精简 YOLOv3 模型在检测效果上有略微的下降,SlimYOLOv3-SPP3-50+层剪枝后的 YOLOv3 模型要略优于 E-T-YOLOv3 模型。

表 3 不同算法的检测结果比较

算法	AP/%	Iou	模型体积 /MB	时间 /ms
YOLOv3	86.9	0.82	238.5	145
T-YOLOv3	78.6	0.59	33.5	23
E-T-YOLOv3	82.7	0.69	52.4	48
SlimYOLOv3-SPP3-50	85.3	0.75	67.1	57
SlimYOLOv3-SPP3-90	78.9	0.7	49.7	52
SlimYOLOv3-SPP3-95	69.7	0.63	32.4	48
SlimYOLOv3-SPP3-50+ 层剪枝	84.5	0.74	31.3	37

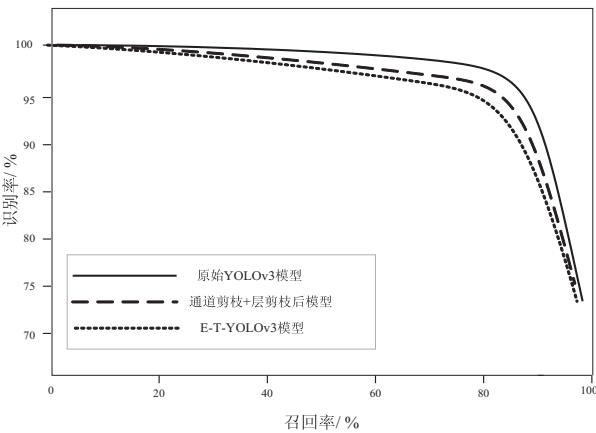


图 4 不同模型的 PR 曲线

为验证巡检机器人前端模型的识别效果,将经过“SlimYOLOv3-SPP3-50+层剪枝”策略压缩后的 YOLOv3 模型部署在变电站巡检机器人上,测试环境为济南 220 kV 许寺敞开式变电站,除了站内待检测设备的识别任务外,还人为制作了一些异物悬挂的场景进行测试。机器人巡检过程中前端输出的检测结果如图 5 所示,从输出结果可以看出部署的模型可以有效地检测出相应的电力设备及异物,从图 5 第三排图像可以看到当前算法对于遮挡的绝缘子仍存在漏报的情况,下一步将针对遮挡设备检测与进一步精简判别模型技术进行研究,并引入知识蒸馏思想^[16],进一步提升部署模型的性能。

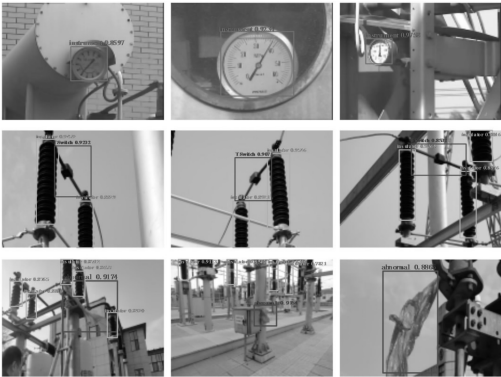


图 5 变电站巡检图像检测结果

4 结束语

针对变电站巡检机器人设备状态检测技术前端化部署需求,从成本及算力方面考虑,选择海思 Hi3559A 芯片作为算法开发的嵌入式平台,以 YOLOv3 算法训练模型作为基础模型进行模型轻量化设计。研究了基于模型剪枝及小型模型设计的轻量化技术,针对不同的模型轻量化算法,提出了改进的 T-YOLOv3 模型,增强模型上下层的语义信息特征提取能力。设计了基于通道卷积与层剪枝的模型压缩算法,在精度保持的前提下,进一步压缩模型的体积。最终将最优的轻量化模型在机器人平台上进行部署,实现设备状态检测技术在变电站巡检机器人中的前置化。实验结果及现场实际应用表明,前端化的网络模型大大减少了算法耗时,检测精度得到有效保持,满足了变电站设备状态巡检的需求。

参考文献:

[1] 蔡焕青,邵瑰玮,胡 霁,等. 变电站巡检机器人应用现状和主要性能指标分析[J]. 电测与仪表,2017,54(14):117-123.

[2] 孔英会,景美丽. 无人值守变电站监控视频异常模式识别方法[J]. 华北电力大学学报,2011,38(6):11-16.

[3] YANG J, AI X, JIA X F. Study of image recognition used for unattended substation[C]//Proceedings of international conference on power system technology. Chongqing: IEEE, 2006:1-6.

[4] GIRSHICK R, DONAHUE J, DARREL T, et al. Region-based convolutional networks for accurate object detection and segmentation[J]. IEEE Transaction on Pattern Analysis & Machine Intelligence, 2015, 38(1):142-158.

[5] GIRSHICK R. Fast R-CNN[C]//International conference on computer vision. Santiago: IEEE, 2015:1440-1448.

[6] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal network[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 39(6):1137-1149.

[7] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//European conference on computer vision. Amsterdam: Springer, 2016:21-37.

[8] DAI J F, LI Y, HE K M, et al. R-FCN: object detection via region-based fully convolutional networks[C]//Proceedings of the 30th international conference on neural information processing systems. Barcelona: Springer, 2016:379-387.

[9] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//IEEE conference on computer vision and pattern recognition. Las Vegas: IEEE, 2016:779-788.

[10] REDMON J, FARHADI A. YOLO9000: better, faster, stron-

(下转第 210 页)