

基于深度学习的烟包识别与分类

淡卫波¹, 朱勇建^{1,2}, 黄毅³

(1.浙江科技学院 机械与能源工程学院, 杭州 310023; 2.宁波敏捷信息科技有限公司, 宁波 315000;
3.长沙理工大学 汽车与机械工程学院, 长沙 410114)

摘要: **目的** 提取烟包图像数据训练深度学习目标检测模型, 提升烟包流水线拣包效率和准确性。**方法** 基于深度学习建立一种烟包识别分类模型, 对原始 YOLOv3 模型进行改进, 在原网络中加入设计的多空间金字塔池化结构 (M-SPP), 将 64×64 尺度的特征图下采样与 32×32 尺度的特征图进行拼接, 并去除 16×16 尺度的预测特征层, 提高模型的检测准确率和速度, 并采用 K-means++ 算法对先验框参数进行优化。**结果** 实验表明该目标检测模型平均准确率达到 99.68%, 检测速度达到 70.82 帧/s。**结论** 基于深度学习建立的图像识别分类模型准确率高且检测速度快, 有效满足烟包流水线自动化实时检测。

关键词: 深度学习; 烟包识别; YOLOv3; K-means++

中图分类号: TP391 **文献标识码:** A **文章编号:** 1001-3563(2023)01-0133-08

DOI: 10.19554/j.cnki.1001-3563.2023.01.015

Cigarette Pack Recognition and Classification Based on Deep Learning

DAN Wei-bo¹, ZHU Yong-jian^{1,2}, HUANG Yi³

(1. College of Mechanical and Energy Engineering, Zhejiang University of Science and Technology, Hangzhou 310023, China; 2. Ningbo Agile Information Technology Co., Ltd., Ningbo 315000, China; 3. College of Automotive and Mechanical Engineering, Changsha University of Science & Technology, Changsha 410114, China)

ABSTRACT: The work aims to extract the cigarette pack image data to train the deep learning target detection model, and improve the efficiency and accuracy of cigarette pack assembly lines. A cigarette pack recognition and classification model was established based on deep learning to improve the original YOLOv3 model. The designed multi-space pyramid pooling structure (M-SPP) was added to the original network. The downsampling of the 64×64 feature map was spliced with that of the 32×32 feature map. The prediction feature layer of 16×16 was removed to improve the detection accuracy and speed of the model, and the K-means++ algorithm was used to optimize the a priori frame parameters. The experiment showed that the average accuracy of the target detection model reached 99.68%, and the detection speed reached 70.82 frames per second. It is concluded that the image recognition and classification model established based on deep learning has high accuracy and fast detection speed, which can effectively meet the automatic real-time detection of cigarette pack assembly lines.

KEY WORDS: deep learning; cigarette pack recognition; YOLOv3; K-means++

收稿日期: 2022-01-19

基金项目: 国家自然科学基金 (51875048); 浙江省基础公益研究计划 (LGG21E050006)

作者简介: 淡卫波 (1995—), 男, 硕士生, 主攻计算机视觉与目标检测。

通信作者: 朱勇建 (1979—), 男, 博士, 副研究员, 主要研究方向为机器视觉与三维测量。

早在 2011 年,国家烟草专卖局提出“烟草物流是烟草行业核心业务,是中国烟草面向未来提升核心竞争力的重要支撑”,明确了卷烟物流在行业中的重要地位^[1]。烟包流水线当前拣包合规性主要依托于拣包系统,但是烟包经过冗长的流水线,伴随着订单流水号的分流、拆包、组包后,往往会随机出现烟包丢失、错乱的现象,仅依靠传统的人工方式复查难以实现精准、及时地发现错误,进而会造成后续一系列的拣包错乱。若无法及时监管到拣包错乱现象,会造成流水线停线,甚至是召回配送车,直接造成大量生产时间的浪费和资金的损失,因此,实现烟包流水线自动化实时监测,对烟草物流提质增效至关重要。对此,张毅等^[2]基于机器视觉对烟包图像进行特征点匹配,实现对烟包的识别,但未考虑不同品牌香烟包装信息的差异,具有一定的局限性。刘莹等^[3]通过卷积神经网络对 10 种品牌烟盒进行分类,该方法仅对单个烟盒进行分类,对实际烟包流水线复杂的场景并不适用。

二十世纪以来,人工智能技术逐渐应用于越来越多的工业领域,而深度学习^[4]作为人工智能领域重要的分支,在图像相关的实践工作中取得了显著的成果。尤其是随着计算机硬件的图形处理器(GPU)计算能力的快速发展,深度学习技术迎来了蓬勃发展的春天。在计算机视觉领域中,目标检测^[5]任务一直以来都具有举足轻重的地位,同时也充满挑战。传统目标检测技术通过人工提取目标图像的特征,2012 年以来随着 AlexNet^[6]图像分类网络的诞生,基于卷积神经网络的目标检测算法迎来了飞速发展。基于卷积神经网络的目标检测算法可以分为两阶段目标检测和单阶段目标检测^[7]。两阶段目标检测技术使用算法生成图像的候选框,由卷积神经网络对其进行分类;单阶段目标检测技术不生成候选框,而是把目标检测问题转化为回归问题。两阶段目标检测技术的优势在于准确率和定位精度更高,而单阶段目标检测技术的检测速度更加快。

烟包流水线自动化检测项目是在每条生产线的 2 个出入口各自安装订单对比系统,利用深度学习目标检测技术,快速地对原始订单流水号与出口图像进行智能比对,并通过信号灯触发报警纠错。首先,使用图像采集装置采集烟包流水线上的图像数据,制作烟包数据集。其次,根据烟包的图像特征对原始 YOLOv3^[8]模型进行改进,提出以下 3 点改进:为了提高对烟包图像数据的识别精度,采用 K-means++^[9]算法对烟包数据集进行聚类,获得 6 种宽高比的目标候选框。由于烟包在整幅图像中属于较小目标,因此将原始网络中预测大目标的 16×16 尺度的特征预测层去除,并将原网络中 64×64 尺度的特征图进行 2 倍下采样,与 32×32 尺度的特征图进行拼接,然后使用 32×32 和 64×64 尺度的特征预测层预测目标,提高网络对烟包目标的召回率及检测精确率。在主干网络

Darknet-53 后加入设计的多空间金字塔池化结构(M-SPP),通过对特征图不同尺度的池化操作,将特征图的局部和全局信息进行结合。将文中改进的网络与多种目标检测网络在流水线烟包数据集进行对比实验,实验结果表明改进后的网络准确率、召回率和检测速度都有提高。

1 实验

1.1 设备的搭建

文中需要对流水线烟包的图像进行采集,因此在流水线上搭建了烟包图像采集装置见图 1。此装置通过 LED 高强度光源照亮目标,形成合适的成像效果,烟包被传输到检测位置时触发 CCD 相机采集高清晰的图像。另外,文中算法运行的计算机硬件配置为 AIMB-706 主板,CPU 为 i7-9700,GPU 为 RTX 3090,Ubuntu16.04 操作系统,使用 Pytorch 深度学习框架。

1.2 数据采集与标注

基于深度学习的目标检测算法,数据集的制作至关重要,目标检测算法的性能表现很大程度上取决于其所用的数据集。数据集选取样本量大且具有代表性的 5 类烟包,通过图像采集装置采集 1 600 张原始图像,包含白沙(baisha)856 张、云烟(yunyan)745 张、利群(liqun)613 张、南京(nanjing)763 张和芙蓉王(furongwang)671 张,数据集内标签数量分别为白沙 2 567 个、云烟 2 143 个、利群 2 387 个、南京 2 073 个、芙蓉王 2 281 个,按照 9:1 的比例制作训练集和测试集用于训练和验证。

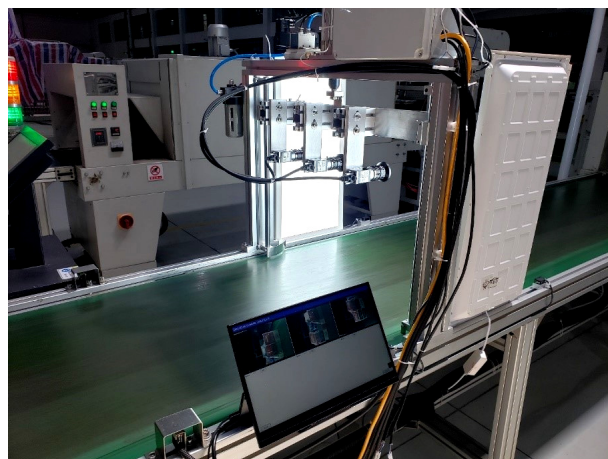


图 1 烟包图像采集装置结构图
Fig.1 Structure diagram of cigarette pack image acquisition device

通过 Labellmg 图像标注工具对采集到的烟包图像进行标注,使用标记框选择不同类型的烟包进行区

域标注, 并添加相应的类别标签, 生成对应的 xml 格式的文件。烟包图像标注界面见图 2。

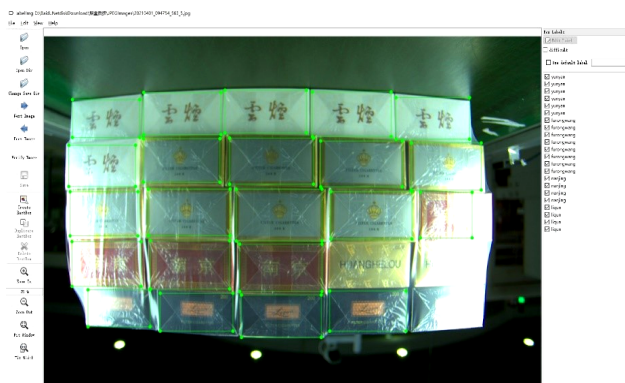


图 2 烟包图像标注界面
Fig.2 Cigarette pack image
labeling interface

2 识别与分类算法

YOLOv3^[8]网络是在 YOLOv1^[10]和 YOLOv2^[11]网络的基础之上改进而来。YOLOv3 通过 Darknet-53 主干网络来提取输入图像的特征, Darknet-53 的残差单元结构借鉴了 ResNet(Residual Neural Network)^[12]的网络结构, 弱化了卷积神经网络中层与层之间的紧密联系, 减弱了梯度消失现象, 同时加深了网络的深度。此外, YOLOv3 网络中没有池化层, 而是将卷积层的步长设置为 2 来达到下采样的效果, 同时将尺度不变特征传到下一层。批量归一化 (Batch Normalization)^[13]、卷积层和激活函数 (Leaky Relu) 共同构成 YOLOv3 网络的基本单元。

YOLOv3 网络的预测结构借鉴了特征金字塔 (FPN)^[14]策略, 采用 3 种尺度预测层对不同大小的目标进行检测, 提高了网络的检测能力。输入图像通过主干网络提取特征, 下采样 5 次, 输出 16×16 大小的特征图来检测大尺寸的目标; 然后将 16×16 特征图上采样与特征提取网络中 32×32 的特征图进行张量拼接 (Concat) 得到 32×32 大小的特征图, 融合 2 种尺度特征来检测中等尺寸的目标; 同样将 32×32 的特征图上采样与特征提取网络的 64×64 特征图进行张量拼接, 可以得到 64×64 大小的特征图, 用来检测小尺寸目标。另外, YOLOv3 网络沿用了 YOLOv2 网络中关于先验框的策略, 先验框的设置参考了 Faster RCNN^[15]和 SSD^[16]网络中的区域生成方法, 但没有根据工程经验人为设计先验框长宽比的大小, 而是通过 K-means 聚类的方法获得 9 个先验框的值。原始 YOLOv3 网络结构见图 3。

3 改进的 YOLOv3 网络

研究发现, 由于烟包数据集中的目标均属于较小

目标, 物体尺寸相似均为长矩形。直接使用原始 YOLOv3 网络模型对烟包进行识别, 准确率不能满足烟包流水线自动化实时识别的要求, 因此, 基于烟包图像的特征做出以下 3 点改进。

3.1 先验框的改进

YOLOv3 作为基于先验框机制的目标检测算法, 先验框的比例与尺寸设置十分重要。YOLOv3 网络总共输出 3 种尺度的预测特征图, 每种尺度的特征图预测 3 种尺寸的输出, 因此共有 9 种大小不同的先验框。先验框尺寸的设置是根据 K-means^[17]聚类算法在公共数据集上得到的, 而文中的数据为烟包图像, 与公共数据集的检测目标尺寸相差巨大, 因此, 需要对采集到的烟包数据进行聚类分析得到合适的先验框尺寸。

K-means 聚类算法的结果十分依赖初始点的选取, 若算法随机生成的初始点不合适, 将会导致最终的迭代结果陷入局部最优解。K-means++^[9]算法对聚类中心的选取采取以下流程。

1) 从数据集中随机选取 1 个点作为初始聚类的中心 c_1 。

2) 计算每个样本 x_i 与已有聚类中心点的距离, 用 $D(x)$ 表示; 然后计算每个样本被选为下一个聚类中心点的概率 $P(x) = \frac{D(x)}{\sum D(x)^2}$, 然后通过轮盘法选出一个聚类中心点。

3) 重复步骤 2 直到选择出 K 个聚类中心点。

综合, 由上述流程可以看出, K-means++ 算法选取了更加离散的初始点, 提高了先验框分布的合理性。由于数据集中的烟包均为长宽比例相似的矩形, 图像中的待测目标因为空间位置不同尺寸有所差异, 相对于整幅图像的尺寸属于较小的检测目标, 因此用检测较小尺寸的 2 个预测特征层上的 6 个先验框来预测目标。如图 4 所示, 烟包数据集使用 K-means++ 聚类算法生成了 6 种长宽比的先验框尺寸, 分别为 (22,16)、(35,25)、(48,36)、(64,48)、(85,45)、(95,49)。

3.2 改进的网络结构

深度卷积神经网络的特点是随着卷积层数的加深, 卷积层提取的特征越来越抽象。深度卷积神经网络可以作为有效的信息蒸馏管道, 向网络中输入原始数据 (文中的烟包图像), 通过反复对其进行不同尺度卷积, 将无关信息过滤 (比如物体的具体轮廓细节), 并放大和细化有用的信息。

由于文中待测目标即烟包的物体尺寸相同, 相对整幅图像属于尺寸较小的目标。为了减少误检, 提高网络对待测目标的检测准确率, 将原 YOLOv3 网络中用来预测尺寸较大目标的 16×16 预测特征层去除。

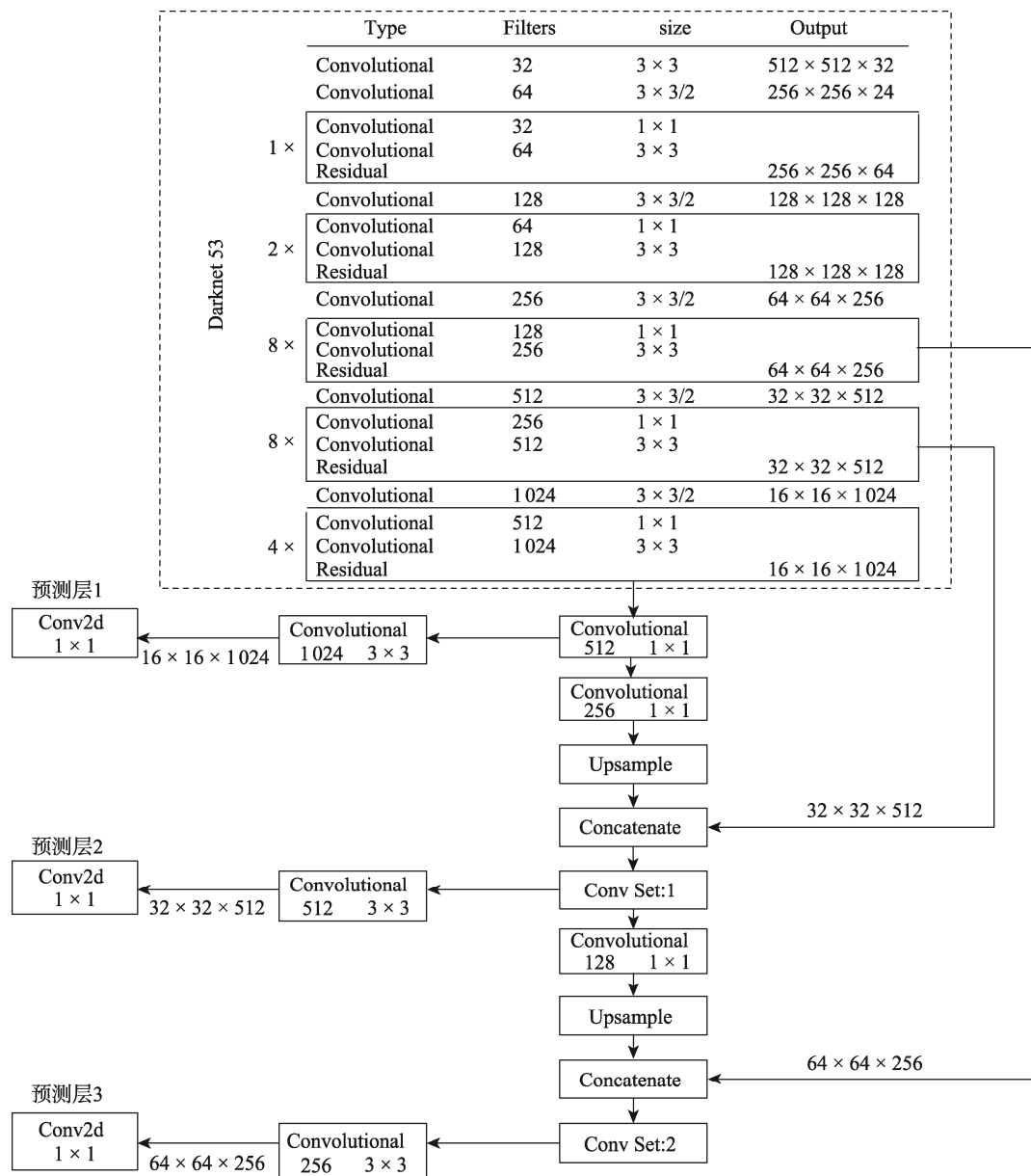


图3 原 YOLOv3 网络结构
Fig.3 The original YOLOv3 network structure

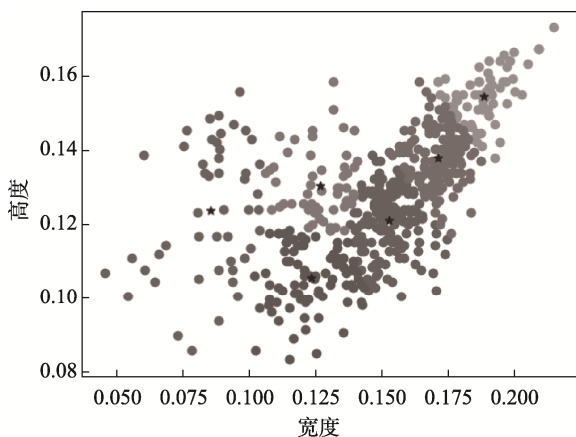


图4 烟包数据基于 K-means++ 的聚类结果
Fig.4 Clustering results of cigarette pack data based on K-means++

在深度卷积神经网络中较低层的特征图中包含待检测物体的边缘、颜色等基本信息，而烟包识别分类任务需要这些信息，因此，对原始 YOLOv3 网络进行不同尺度特征图的特征融合，提高检测网络对烟包图像的边缘、颜色等基本特征的识别效果。

YOLOv3 网络采用多尺度检测，使用 3 个不同尺度 16×16 、 32×32 、 64×64 的特征图来预测目标。在 64×64 的底层预测特征图中包含待测目标更多的边缘和颜色等基本信息，因此，在 64×64 的底层特征图做类别预测 2 之前，先提取其特征图，并通过 2 倍下采样与 32×32 的中层特征图进行张量拼接，不会损失特征信息同时参与类别预测，相当于在网络中添加自下向上的特征融合路径，使网络更好地运用待测目标的边缘、颜色等基本信息。改进后的网络结构见图 5。

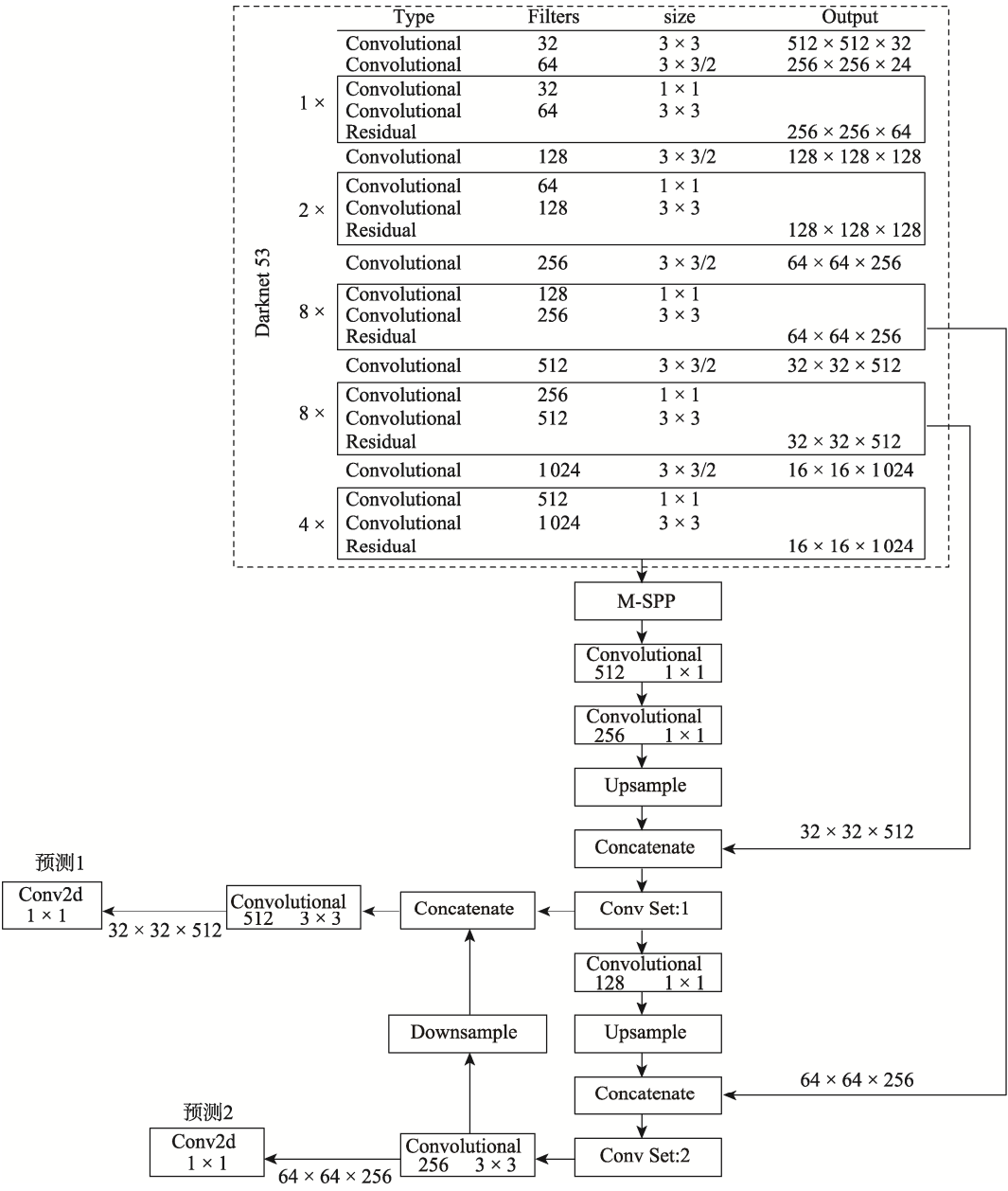


图 5 改进的 YOLOv3 网络结构
Fig.5 The improved YOLOv3 network structure

3.3 多空间金字塔池化结构 (M-SPP)

YOLOv3 中特征提取主干网络输出的特征图尺寸为 16×16, 不同尺度的特征图对目标检测的优势不同, 大尺度的特征图对目标定位更加准确, 小尺度的特征图包含的语义信息更加丰富^[18]。在网络中加入空间金字塔池化结构^[19] (SPP), 通过不同尺度的最大池化处理使特征图包含的语义信息更加丰富, 并融合不同层次的特征图信息。

图 6 所示为设计的多空间金字塔池化结构 (M-SPP), 此结构在原 SPP 的基础上增加 2 个不同尺度的最大池化核, 能够融合更多层次的特征图信息, 提高检测准确率。该结构添加在 YOLOv3 特征

提取网络 Darknet-53 之后, 由最大池化卷积核大小为 3×3、5×5、7×7、9×9、11×11 和一个跨层连接组成, 然后以 Concatenate 拼接的方式输出向量。

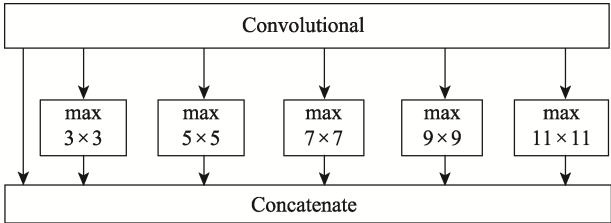


图 6 多空间金字塔池化结构 (M-SPP)
Fig.6 Multi-spatial pyramid pooling structure (M-SPP)

4 实验结果

4.1 模型的评价指标

模型的评价指标主要为均值平均精度 mAP (Mean Average Precision), 均值平均召回率 mAR (Mean Average Recall), 单张图像检测时间。采用的计算式见式(1) — (6)。

$$P = \frac{T_p}{T_p + F_p} \times 100\% \quad (1)$$

$$R = \frac{T_p}{T_p + F_N} \times 100\% \quad (2)$$

$$A_p = \int_0^1 P(R) dR \times 100\% \quad (3)$$

$$A_R = \int_0^1 R(P) dP \times 100\% \quad (4)$$

$$A_{mAP} = \frac{\sum_{i=1}^5 A_{p_i}}{5} \quad (5)$$

$$A_{mAR} = \frac{\sum_{i=1}^5 A_{R_i}}{5} \quad (6)$$

式中: T_p 为预测正确的目标个数; F_p 为预测错误的目标个数; F_N 为没有预测出的目标个数。由上述 P 、 R 可以绘制 P - R 曲线, A_p 则为 P - R 曲线围成的面积。 A_{mAP} 为 5 个类别的 A_p 平均值, 代表模型在 5 个类别上的检测准确率。

4.2 模型的训练

将制作的数据集送入模型中进行训练, 提前对网络参数初始化, 网络中的学习率设置为 0.001, 迭代次数 (epoch) 设置为 200, 批量大小设置为 64, 衰减系数设置为 0.000 5, 动量因子设置为 0.9。如图 7a 所示, 在训练过程中, 原 YOLOv3 网络在 60 个 epoch 附近开始收敛, 而使用 2 个预测层的改进网络模型在 50 个 epoch 附近开始收敛, 收敛速度更快。由图 7b 可以得出, 在 YOLOv3 网络中加入 M-SPP 结构比加入 SPP 结构在验证集上 mAP 值 (Iou 为 0.5) 提升约 0.5%, 提高了检测精度。

4.3 消融实验

为了对比模型不同改进部分的具体效果, 设计了如表 1 的消融实验, 表 1 中“✓”表示使用对应的改进方法。在相同的实验情况下, YOLOv3 网络中加入 M-SPP 结构, 用 K-means++ 算法优化先验框均可以提高检测精度, 两者结合使用可以提高 mAP 值约 1.5%, 而将网络中的预测层减少到 2 层可以提升检测速度, 检测时间减少约 2 ms。将 3 个改进部分同时应用于模型后, 对模型的检测精度和速度都有提升。

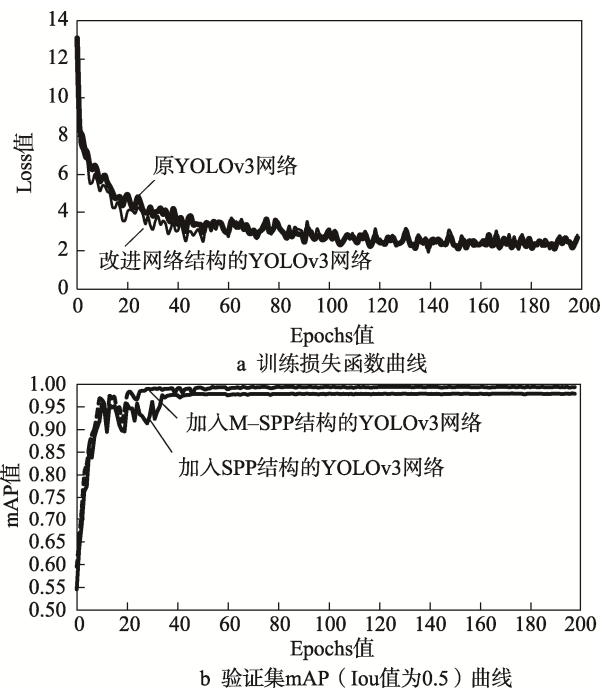


图 7 模型训练过程

Fig.7 Process of model training

表 1 消融实验数据

Tab.1 Data of ablation experiment

M-SPP	K-means++	2 个预测层	mAP 值/%	检测时间/ms
			98.16	16.81
✓			98.53	16.85
	✓		98.71	16.74
		✓	98.13	14.07
✓	✓		99.64	16.83
	✓	✓	98.72	14.11
✓	✓	✓	99.65	14.13

4.4 模型识别结果对比

图 8a、b 为原 YOLOv3 网络的识别结果, 图 8c、d 为改进的网络识别结果, 对比图 8a、c 和图 8b、d 可以发现原 YOLOv3 网络对图像中左下角的白沙 (baisha) 类别的烟包未检测出, 而改进的网络则正确地预测出图像中所有烟包的类别和数量, 因此, 改进的网络提高了对模糊目标和不完整目标的检测准确率, 提升了模型的检测能力。

4.5 不同模型的测试结果

为了对比改进后的网络检测效果, 选择当前主流的 Faster R-CNN、SSD、YOLOv3 和 YOLOv5^[20] 网络进行实验对比, 分别将这 5 个检测网络在制作的数据集上进行训练, 测试结果见表 2。

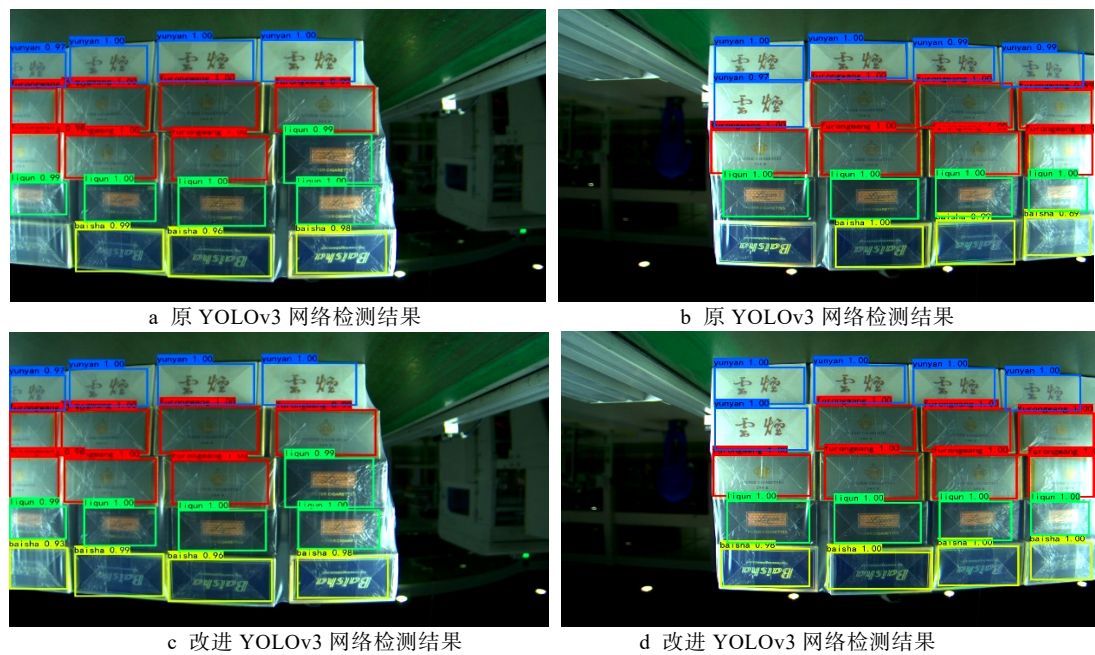


图 8 原 YOLOv3 网络与改进的 YOLOv3 网络对烟包的识别结果对比
Fig.8 Comparison of cigarette pack recognition results between the original YOLOv3 network and the improved YOLOv3 network

表 2 不同目标检测网络的检测结果对比
Tab.2 Comparison of detection results of different target detection networks

目标检测算法	Anchor Boxes	mAP 值/%	mAR 值/%	单张图片检测时间/ms
Faster R-CNN	9	99.22	82.10	426
SSD	9+	98.28	84.91	28.60
YOLO v3	9	98.16	83.26	16.78
YOLO v5	9	99.31	87.52	16.21
改进后的 YOLOv3	6	99.68	88.93	14.12

表 2 所示,虽然改进的网络在 mAP 指标上比 Faster R-CNN、SSD 网络领先不多,但在 mAR 指标上领先 SSD 网络约 4%,另外在检测速度上改进的网络远远领先 Faster R-CNN 和 SSD 网络,稍领先于 YOLOv5 网络,检测速度在烟包流水线自动化识别中至关重要。综合 mAP 和检测时间来看,改进的网络性能更加均衡,在保证检测精度的前提下提升了检测速度,可以满足流水线烟包实时、精准的识别要求。

5 结语

流水线烟包识别分类是一个长期制约烟草物流高质量发展的重要因素,而传统的检测方法及人工复查难以满足实时、精准的识别分类要求。文中提出了改进的 YOLOv3 网络,首先采用 K-means++算法针对烟包数据集进行聚类,获得最优的目标候选框:其次针对烟包图像目标的特点,去除原网络中预测大目标的 16×16 尺度的特征预测层,并将 64×64 尺度的特

征预测层 2 倍下采样与 32×32 尺度的特征预测层进行特征融合,进一步优化网络结构。经过实验验证,文中改进后的 YOLOv3 网络平均检测精度达到 99.68%,同时检测速度也进一步提升,可以实现流水线烟包实时准确地识别分类。

参考文献:

[1] 冰火, 禾木. 论烟草供应链物流建设[J]. 中国烟草学报, 2014, 20(2): 1-8.
BING Huo, HE Mu. On Construction of Supply Chain Logistics in Tobacco Industry[J]. Acta Tabacaria Sinica, 2014, 20(2): 1-8.
[2] 张毅, 王彦博, 付华森, 等. 基于机器视觉的不规则烟包校对码垛系统[J]. 烟草科技, 2019, 52(6): 105-111.
ZHANG Yi, WANG Yan-bo, FU Hua-sen, et al. Irregular Cigarette Parcel Stacking System Coupled with Machine

- Vision-Based Parcel Identification[J]. Tobacco Science & Technology, 2019, 52(6): 105-111.
- [3] 刘莹, 王晓宇, 徐卓飞, 等. 基于卷积神经网络的商品图像识别[J]. 数字印刷, 2020(6): 33-40.
LIU Ying, WANG Xiao-yu, XU Zhuo-fei, et al. Recognition of Commodities Images Based on the Convolutional Neural Network[J]. Digital Printing, 2020(6): 33-40.
- [4] 郑远攀, 李广阳, 李晔. 深度学习在图像识别中的应用研究综述[J]. 计算机工程与应用, 2019, 55(12): 20-36.
ZHENG Yuan-pan, LI Guang-yang, LI Ye. Survey of Application of Deep Learning in Image Recognition[J]. Computer Engineering and Applications, 2019, 55(12): 20-36.
- [5] 段仲静, 李少波, 胡建军, 等. 深度学习目标检测方法及其主流框架综述[J]. 激光与光电子学进展, 2020, 57(12): 59-74.
DUAN Zhong-jing, LI Shao-bo, HU Jian-jun, et al. Review of Deep Learning Based Object Detection Methods and Their Mainstream Frameworks[J]. Laser & Optoelectronics Progress, 2020, 57(12): 59-74.
- [6] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet Classification with Deep Convolutional Neural Networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [7] 刘洋, 战荫伟. 基于深度学习的小目标检测算法综述[J]. 计算机工程与应用, 2021, 57(2): 37-48.
LIU Yang, ZHAN Yin-wei. Survey of Small Object Detection Algorithms Based on Deep Learning[J]. Computer Engineering and Applications, 2021, 57(2): 37-48.
- [8] REDMON J, FARHADI A. YOLOv3: An Incremental Improvement[C]// IEEE conference on Computer Vision and Pattern Recognition, 2018.
- [9] ARTHUR D, VASSILVITSKII S. K-Means++: The Advantages of Careful Seeding[C]// Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, New York: ACM, 2007: 1027-1035.
- [10] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-Time Object Detection[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 2016: 779-788.
- [11] REDMON J, FARHADI A. YOLO9000: Better, Faster, Stronger[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA, 2017: 6517-6525.
- [12] HE Kai-ming, ZHANG Xiang-yu, REN Shao-qing, et al. Deep Residual Learning for Image Recognition[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA, 2016: 770-778.
- [13] 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, Lille, France, 2015: 448-456.
- [14] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature Pyramid Networks for Object Detection[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 2017: 936-944.
- [15] REN Shao-qing, HE Kai-ming, 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, 2017, 39(6): 1137-1149.
- [16] LIU Wei, ANGUELOV D, ERHAN D, et al. SSD: Single Shot MultiBox Detector[M]. Computer Vision - ECCV 2016. Cham: Springer International Publishing, 2016: 21-37.
- [17] 杨俊闯, 赵超. K-Means 聚类算法研究综述[J]. 计算机工程与应用, 2019, 55(23): 7-14.
YANG Jun-chuang, ZHAO Chao. Survey on K-Means Clustering Algorithm[J]. Computer Engineering and Applications, 2019, 55(23): 7-14.
- [18] 姜文涛, 张驰, 张晟翀, 等. 多尺度特征图融合的目标检测[J]. 中国图象图形学报, 2019, 24(11): 1918-1931.
JIANG Wen-tao, ZHANG Chi, ZHANG Sheng-chong, et al. Multiscale Feature Map Fusion Algorithm for Target Detection[J]. Journal of Image and Graphics, 2019, 24(11): 1918-1931.
- [19] HE Kai-ming, ZHANG Xiang-yu, REN Shao-qing, et al. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.
- [20] Ultralytics. YOLOv5[EB/OL]. [2021-03-14]. <https://github.com/ultralytics/yolov5>.