

基于轻量化 YOLOv4 的黏稠食品灌装成品缺陷检测

doi:10.3969/j.issn.1674-7100.2021.02.005

张昌凡¹ 孟德志¹

王燕因²

1. 湖南工业大学

电气与信息工程学院

湖南 株洲 412007

2. 广州达意隆包装机械

股份有限公司

广东 广州 510530

摘要: 针对黏稠流质食品的自动化灌装中, 传统灌装成品缺陷检测方法难以同时对多目标高速检测的难题, 提出一种基于 YOLOv4 目标检测算法的轻量级灌装成品缺陷检测方法。MobileNetV3 主干特征提取网络能对输入样本进行轻量级特征提取; 增强特征提取网络采用深度可分离卷积策略, 以降低参数计算量, 然后通过设计的全面路径聚合网络 (FPANet) 和引入的通道注意力机制 (ECA) 提升增强特征提取网络对于目标特征的靶向表达。将设计的轻量化网络进行模型训练和精度测试, 并在同一数据集下与其它目标检测算法进行对比, 以分析本文方法的优劣。实验结果表明, 本文方法能够在保持精度的前提下提升检测速度, 实现了黏稠食品灌装成品缺陷的多目标高速检测。

关键词: 黏稠食品; 缺陷检测; 深度可分离卷积; 通道注意力机制

中图分类号: TP391

文献标志码: A

文章编号: 1674-7100(2021)02-0037-09

引文格式: 张昌凡, 孟德志, 王燕因. 基于轻量化 YOLOv4 的黏稠食品灌装成品缺陷检测 [J]. 包装学报, 2021, 13(2): 37-45.

1 研究背景

目前, 在黏稠流质食品灌装生产线中, 灌装成品的缺陷检测是必不可少的环节。有封盖缺陷的产品随着时间推移会出现密封不严的问题, 易导致产品提前变质。贴标缺陷会影响产品口碑和消费者购买意愿。因此, 对灌装成品的封盖和贴标进行检测极其必要。

在现有研究中, 已有许多学者提出了灌装成品缺陷检测方法。C. Arun Prakash 等^[1] 利用机器视觉技术对瓶盖进行检测, 该方法简单、省时, 检测正确率高; C. Toxqui-Quitl 等^[2] 提出了一种基于图像频率滤波的

聚对苯二甲酸乙二酯 (poly(ethyleneterephthalate), PET) 瓶身缺陷检测方法, 对采集的瓶身图像进行伽马自适应校正, 实现了对于瓶面、瓶壁和瓶底较高准确度的检测; 彭玉等^[3] 针对灌装成品防盗环难以检测的问题, 提出了一种基于轮廓曲率的防盗环断裂检测方法, 该算法检测正确率能满足生产要求。尽管上述机器视觉技术在灌装成品缺陷检测上得到了较好的应用, 但需要大量的样本构建检测模型且难以同时检测多类缺陷。

近年来, 深度学习技术在模型调优、特征提取、目标检测方面的优势显而易见^[4]。深度学习标志性算

收稿日期: 2021-01-13

基金项目: 国家重点研发计划基金资助项目 (2018YFD0400705)

作者简介: 张昌凡 (1960-), 男, 湖北武昌人, 湖南工业大学教授, 博士生导师, 主要从事复杂工业工程控制、包装工业自动化方面的研究, E-mail: zhangchangfan@263.net

通信作者: 孟德志 (1995-), 男, 河北唐山人, 湖南工业大学硕士生, 主要研究方向为工业智能检测,

E-mail: mengdezhi1995@outlook.com

法 AlexNet 的问世^[5], 使深度学习技术应用于产品缺陷检测成为可能。2018 年, 俞芳芳等^[6]利用卷积神经网络 (convolutional neural network, CNN) 对食用油灌装生产线全流程包装缺陷进行检测。传统 CNN 网络结构复杂, 参数多, 运算时间长, 因此在保持精度阈值的前提下对网络结构轻量化设计尤为重要。林景栋等^[7]从网络剪枝与稀疏化、张量分解、知识迁移和精细模块设计 4 个方面归纳了 CNN 结构优化技术。MobileNet^[8]系列算法将传统卷积变成深度可分离卷积 (depthwise separable convolutions, DSC), 这种新的结构能在不影响准确率的前提下大大减少运算时间和参数量。

深度学习目标检测算法发展至今, 已形成了以 R-CNN (regions with CNN)^[9]为代表的两步骤检测方法和以 YOLO (you only look once)^[10]算法为代表的单步骤检测算法, 其中前者更注重精度, 后者更注重精度与速度的平衡^[11]。但是目前目标检测应用于灌装产品缺陷检测鲜有研究。如何利用单一标准实现灌装产品缺陷的多目标检测成为缺陷检测领域亟待解决的问题。

综上所述, 本课题组提出了一种基于 YOLOv4^[12]的黏稠食品轻量级灌装成品缺陷检测方法, 在单步骤目标检测算法的基础上对模型进行轻量化设计, 仅利用单一规则实现对多缺陷目标快速实时的检测。

2 YOLOv4 算法

2.1 网络结构

YOLOv4 算法是一种端到端的单步骤目标检测方法, 它省去了两步骤检测法中生成候选区域的中间步骤, 从而加快对于目标的检测与识别。在框架上, 其很大程度上沿袭了前代算法 YOLOv3^[13]的主体结构, 其网络结构大致分为主干特征提取网络 (Backbone)、增强特征提取网络、多尺度结果输出 3 个部分。

主干特征提取网络采用的是 CSPDarknet53, 先经最小卷积单元 DBM (DarknetConv2D-BN-Mish), 再经 Resblock-body 复杂卷积环节实现卷积操作。Resblock-body 借鉴了 CSPNet (cross-stage-partial-connections network)^[14]结构, 由 DBM、残差结构和 Concat 构成。增强特征提取网络对 Backbone 得到的特征图进行复用, 以增强其对于目标特征的表达。YOLOv4 算法引入了 SPPNet^[15]和 PANet^[16] (path aggregation network) 两种算法, 前者可增加小尺度

特征的感受野, 分离出最显著的图像特征, 后者则是对 Backbone 所得到的另外两种尺度的特征图和 SPPNet 输出的最小尺度特征图进行反复地上采样和下采样操作, 以通过多尺度特征的融合来实现对于目标特征的优质表达。经过增强特征提取网络的特征融合, 得到 3 种不同尺度的特征图。

2.2 预测机制

首先, YOLOv4 将输入的样本分成 $S \times S$ ($S \in \mathbf{N}$) 个非重叠区域, 然后使用 Anchor boxes 对每个区域进行边界盒 (bounding box) 回归, 如图 1 所示。通过 K-means 聚类算法对 Anchor boxes 进行聚类分析, 从而得到一组尺寸固定的先验框, 这些不同尺寸的先验框能够提升其对于真实目标位置预测的概率。

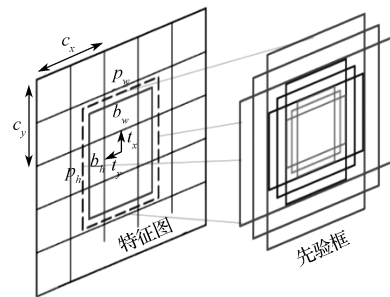


图 1 边界盒回归

Fig. 1 Bounding box regression

若某预测目标的中心位置落到某个区域, 则该目标由此区域的一组先验框进行预测, 这会产生多个预测边界盒。假设每个区域预测出 B 个边界盒和 C 个类别, 那么模型的输出结果为 $S \times S \times (5B+C)$ 的张量, 其中每个边界盒均包含 b_x 、 b_y 、 b_w 、 b_h 和置信度 (confidence) 5 个数值, 其计算公式如下^[17]:

$$b_x = \sigma(t_x) + c_x; \quad (1)$$

$$b_y = \sigma(t_y) + c_y; \quad (2)$$

$$b_w = p_w e^{t_w}; \quad (3)$$

$$b_h = p_h e^{t_h}; \quad (4)$$

$$\text{Confidence} = P(\text{object}) \times \text{IoU}_{\text{pred}}^{\text{truth}}, P(\text{object}) \in \{0, 1\}. \quad (5)$$

式中: c_x 、 c_y 分别为待检测区域与检测样本左上角的横纵距离;

t_x 、 t_y 分别为检测区域中心点至左边界和上边界的距离;

$\sigma(\cdot)$ 为 Sigmoid 激活函数;

p_w 、 p_h 分别为先验框的宽和高;

t_w 、 t_h 分别为检测区域的宽和高;

b_x 、 b_y 分别为边界盒的中心点 x 、 y 坐标;

b_w 、 b_h 分别为边界盒的宽和高。

预测边界盒会出现冗余, 因此引入非极大值抑制 (non-maximum suppression, NMS), 以保留概率最大的边界盒, 移除分数较低的边界盒。

3 灌装成品缺陷检测模型

本课题组设计了基于轻量化 YOLOv4 的黏稠食品灌装成品缺陷检测模型, 其结构如图 2 所示。其主要分为 6 个部分, 分别为成品缺陷数据集、MobileNetV3 主干特征提取网络、空间金字塔池化网络 (spatial pyramid pooling, SPPNet)、施加通道注

意力 (efficient channel attention, ECA) 机制的深度可分离全面路径聚合网络 (DSC-ECA-FPAnet)、多尺度结果输出 (Yolo Head) 和检测结果可视化。

本检测模型的具体工作流程如下: 首先, 将灌装成品缺陷数据输入 MobileNetV3 主干特征提取网络, 提取多尺度特征, 进而获得 3 个不同尺度的特征图; 然后, 将得到的特征图分别输入至 SPPNet 和 DSC-ECA-FPAnet 中以实现对于多尺度特征的融合; 最后, 将融合后的特征图进行多尺度结果输出, 从而得到既包含位置信息又包含类别信息的检测结果。

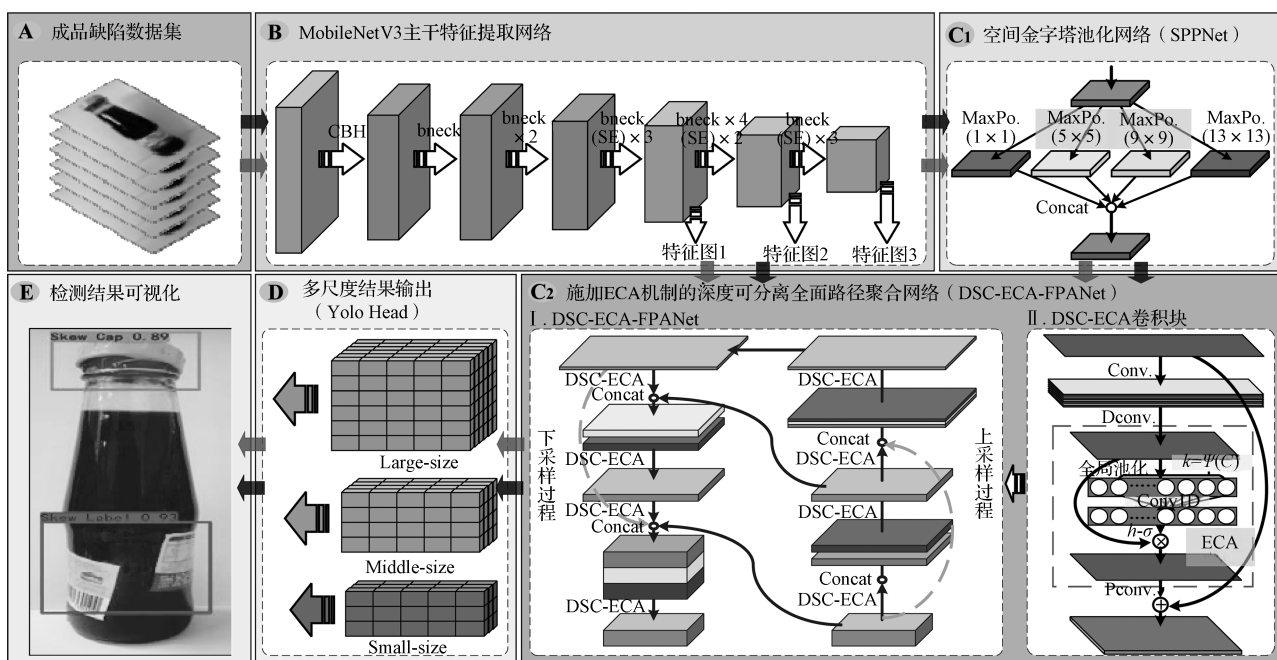


图 2 基于 YOLOv4 算法的轻量级灌装成品缺陷检测结构框图

Fig. 2 Structure diagram of filling product defect detection of lightweight based on YOLOv4 algorithm

3.1 建立灌装成品缺陷数据集

实验中所需的图像样本均来自灌装生产流水线, 采集设备主要包括传送带、LED 光源、CCD (charge-coupled device) 高速相机、光电传感器以及用于图像存储和缺陷检测的计算机, 如图 3 所示。数据采集过程为: 光电传感器检测传送带上的灌装成品, 并将检测信息转换成电信号传给 CCD 相机, 触发相机采集产品图像信息, 再将图像信息传至计算机中保存, 以构建用于后续的成品缺陷数据集。根据实际需求, 本课题组采集了 2 种缺陷、7 种类别共计 1200 张样本, 其中封盖缺陷包含无盖、歪盖、正常盖 3 种类别, 贴标缺陷包含无标、歪标、损伤标、正常标 4 种类别。

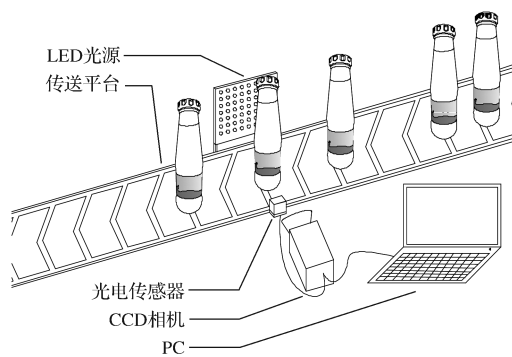


图 3 图像采集平台

Fig. 3 Image acquisition platform

本数据集的样本分布情况如图 4 所示。用 LabelImg 软件对样本进行了真实标签标定, 得到

1200 个具有缺陷类别和位置信息的 .xml 文件。文件和样本均用于模型的训练和测试。

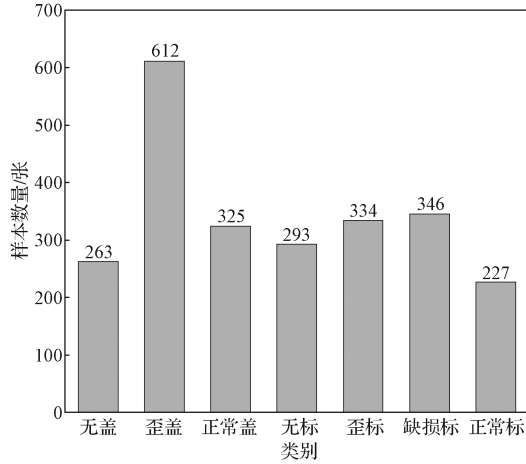


图 4 采集数据类别数量情况

Fig. 4 Number of collected data categories

3.2 改进主干特征提取网络

在 YOLOv4 中, Backbone 有复杂的残差结构, 标准 3×3 卷积操作占据了主体, 导致运算时间长, 算法可移植性较差。针对上述问题, 本课题组引入了轻量级神经网络 MobileNetV3^[18], 并根据目标检测要求重置了参数大小, 设计了网络特征图的多尺度输出, 以便用于后续增强网络的特征复用。

MobileNetV3 的核心是 DSC^[19], 这使参数量大幅降低, 检测速度提升。公式 (6) ~ (8) 揭示了 DSC 参数量减少的原理^[19]。

标准卷积操作是在多个输入通道上使用多个卷积核进行卷积操作。其参数计算量 C 为

$$C = D_k \times M \times N \times D_F, \quad (6)$$

式中: D_k 为卷积核的大小; M 为输入通道数; N 为卷积核数量; D_F 为输入图像大小。

DSC 分两步完成卷积操作, 先在 M 个输入通道上使用一个卷积核逐通道卷积 (Depthwise Conv.), 然后通过逐点卷积 (Pointwise Conv.) 即利用 N 个 1×1 卷积核调整输出通道数。参数计算量 C' 为

$$C' = D_k \times M \times D_F + M \times N \times D_F. \quad (7)$$

将式 (6) 和 (7) 求比值, 可得

$$\frac{C'}{C} = \frac{1}{N} + \frac{1}{D_k} < 1. \quad (8)$$

通常, 卷积核大小取 3×3 。可见, 采用 DSC 能大幅降低网络参数。

MobileNetV3 主要由 Conv2D-BN-h-swish (CBH) 卷积单元和 bneck 卷积块构成, 如图 5 所示。CBH

和普通卷积结构基本相同, 区别在于引入了 h-swish 激活函数^[18], 即

$$\text{h-swish}(x) = x \frac{\min[\max(0, x+3), 6]}{6}. \quad (9)$$

h-swish 激活函数在深层特征中更能对特征有效映射, 因此在网络中, bneck 卷积块使用了 ReLU 和 h-swish 两种激活函数。图 5b 中, 第 2~4 层卷积采用 ReLU 激活函数, 最后 3 层卷积采用 h-swish 激活函数。而在 bneck 中, 先用 1×1 的 CBH 提取特征, 再用 3×3 的 DSC 进行卷积, 施加 SENet (squeeze-and-excitation networks)^[20] 之后, 进行 1×1 卷积操作以调整通道, 最后输出 bneck 结果。SENet 是一种轻量级通道注意力机制, 通过学习的方式对不同通道重要性进行预测, 以提升有用的特征, 抑制用处不大的特征。

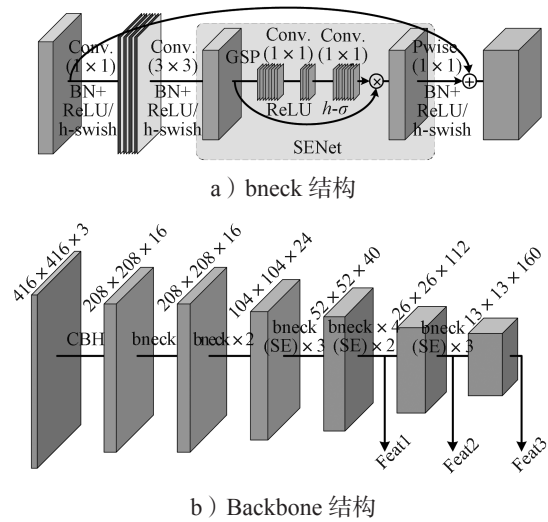


图 5 MobileNetV3 网络结构

Fig. 5 Network structure diagram of MobileNetV3

3.3 重构增强特征提取网络

YOLOv4 的增强特征提取网络包含 SPPNet 和 PANet。经增强特征提取网络的特征复用得到 3 个不同尺度特征图。SPPNet 对最小尺度特征图进行处理后, 同另外两个特征图输入至 PANet 进行特征融合。本课题组对增强特征提取网络进行重构, 使其轻量化的同时, 保证检测精度。

3.3.1 PANet 轻量化

和 Backbone 类似, PANet 计算繁重, 采用多个 3×3 的 DBL (DarknetConv2D_BN_LReLU) 卷积, 在上采样与下采样中, 这些卷积操作参数量极大, 导致缺陷检测时间过长。针对上述问题, 本课题组引入

了 DSC 卷积策略, 通过卷积模块替换实现对 PANet 网络的轻量化设计。图 6 为 PANet 轻量化过程中内部卷积模块的替换示意图。

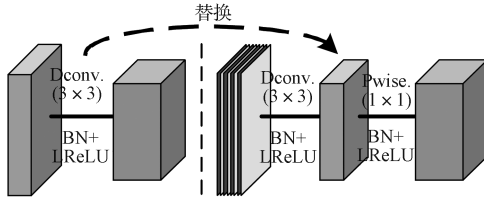


图 6 PANet 内部卷积结构替换示意图

Fig. 6 Internal convolutional structure replacement of PANet

将 3×3 的 DBL 卷积替换 3×3 的 DSC 卷积后, DSC-PANet 网络的参数计算量骤减, 检测速度大幅提升, 但是检测精度下降, 这表明过分使用 DSC 卷积会使特征不能被充分利用。因此, 轻量化后的 PANet 网络还需提升精度。

3.3.2 全面路径聚合网络

PANet 通过增加不同尺度特征层之间的信息传输路径, 对不同尺度特征进行融合, 以实现对于目标特征的优质表达。但在 PANet 中, 相邻特征层之间进行了特征路径聚合, 而顶层特征与底层特征之间尚不存在融合路径。因此, 针对 PANet 轻量化后导致特征融合能力不足的问题, 本课题组将顶层特征与底层特征进行跨越融合 (见图 7 虚线标识), 提出了全面路径聚合网络 (full path aggregation network, FPNet)。FPANet 是通过增强目标特征全面表达来提升网络精度, 其结构如图 7 所示。

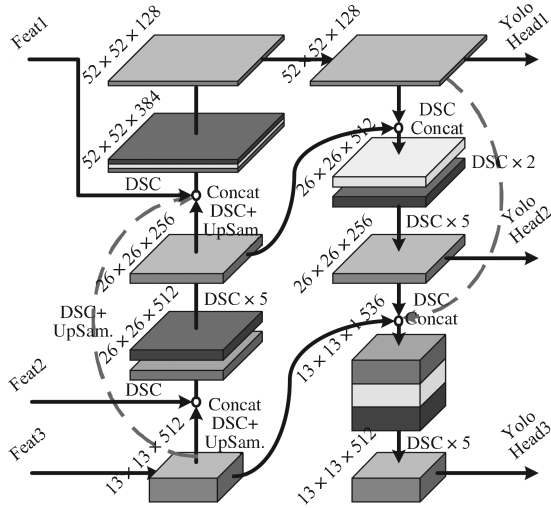


图 7 FPNet 结构图

Fig. 7 FPNet structure diagram

FPANet 接受来自 Backbone 和 SPPNet 的多尺度特征图的输入, 其中 Feat3 为 SPPNet 处理后的特征图。顶层特征和底层特征的跨越融合仍然通过上采样和下采样实现。上采样是通过 1×1 的 DSC 卷积和线性插值法分别负责通道调整 and 尺度扩张。下采样是通过 DSC 卷积来实现通道和特征图尺度的调整。轻量级 FPNet 是通过层间特征充分融合提升网络精度, 且检测速度无显著降低。

3.3.3 DSC-ECA 卷积结构

通道注意力 (ECA) 机制是深度神经网络中有效提升特征提取能力的重要方法。因此, 本课题组在 DSC 卷积中引入 ECA 机制^[21], 提出了 DSC-ECA 模块, 其结构如图 8 所示。

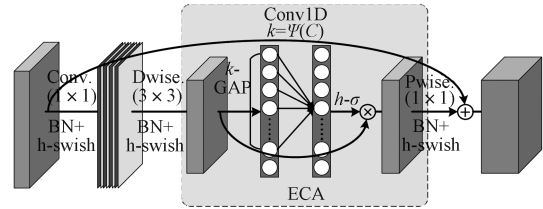


图 8 DSC-ECA 结构图

Fig. 8 DSC-ECA structure diagram

ECA 的结构和 SENet 的相似, 但 ECA 是通过执行大小为 k 的一维卷积来生成权重通道, 避免了降维所带来预测精度的下降。 k 是由输入特征图的通道数自适应选得, 具体公式^[21]为

$$k = \psi(C) = \left\lfloor \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma_{\text{odd}}} \right\rfloor. \quad (10)$$

式中: C 为输入样本通道数; b 和 γ 为经验参数, 设定 $b=2, \gamma=1$; “odd” 为对结果取奇数。同时, 本文保留了最外围的残差边, 以利于权重更新、趋于稳定。

本课题组对 YOLOv4 的增强特征提取网络进行改进后, 形成了全新的施加 ECA 机制的深度可分离全面路径聚合网络 (DSC-ECA-FPNet)。

4 实验结果与分析

4.1 模型训练与评价指标

本实验环境为 Ubuntu18.04 操作系统, 配置 Anaconda 的 Python 3.7、NVIDIA RTX 2070 GPU、CUDA 10.1、cuDNN 7.6.5.32, 使用 Keras 库编写程序。训练数据时, 本课题组从数据集中选取 1000 张样本, 其中 900 张用于模型训练, 100 张用于交叉验证, 而其余 200 张样本用于测试模型。鉴于数据量较少不利

于模型泛化能力的提升,本课题组选择迁移学习方式训练模型。先采用 VOC2007 数据集对网络进行预训练,接着在正式训练过程导入预训练模型分段调优。参数设定如下:输入数据的大小为 416×416 ;为了加快训练速度,前 50 个 Epoch 冻结主干特征提取网络,只对增强特征提取网络进行权重更新,Batch size 设置为 8,初始学习率为 0.001,而后 50 个 Epoch 解冻 Backbone 权重层,进行全局网络优化,Batch size 设置为 4,初始学习率为 0.0001。

在目标检测模型评估中,常采用精确度 P (precision) 和召回率 R (recall) 两个指标评价模型的优劣,其计算方式^[22]如下:

$$P = \frac{TP}{TP + FP}, \quad (11)$$

$$R = \frac{TP}{TP + FN}. \quad (12)$$

式中: TP 为样本正确检测的数量; FP 为样本错误检测的数量; FN 为样本未被检测出的数量。

通常,分类问题存在一个决策边界,改变置信度阈值会得到不同的准确率和召回率,即得到 PR (precision-recall) 曲线。PR 曲线与坐标轴所围成的面积为平均精确度 (average precision, AP), 可用于衡量模型检测效果的好坏。为了衡量模型对于所有类别 AP 的优劣,本课题组使用所有检测类别平均精度均值 (mean average precision, mAP) 评估模型。

4.2 模型评估与性能对比分析

4.2.1 模型训练评估

MobileNetV3_DSC-ECA-FPANet 模型 (改进网络训练后得到的最优模型) 的训练和验证结果如图 9 所示。其中,实线为训练集 loss 值,虚线为交叉验证集 loss 值。

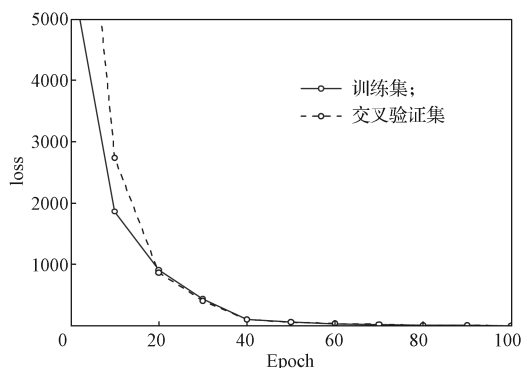


图 9 模型 loss 函数曲线图

Fig. 9 Loss function curve of the model

由图 9 可以看到,模型的 loss 值随着 Epoch 的增加而逐渐收敛,最后基本趋于稳定,这表明模型的预测结果和真实结果不断接近,预测的误差逐渐下降。此外,通过观察训练集和验证集 loss 曲线的吻合度可以看出,模型泛化能力达到了最佳状态。

4.2.2 性能测试

本课题组将训练好的模型导入检测程序后,对 200 张样本进行测试,瓶盖缺陷和贴标缺陷的部分检测结果如图 10 所示。从检测结果可以看出,本文所提方法可以实现对于瓶盖、贴标 2 种检测目标的 7 种缺陷类别进行有效检测。



图 10 样本检测结果示例

Fig. 10 Test results of sample

本课题组设置不同的置信度阈值,从而得到 7 种类别的精确度和召回率,每种缺陷类别的 PR 曲线如图 11 所示。其中,实线为贴标检测结果,虚线为瓶盖检测结果。

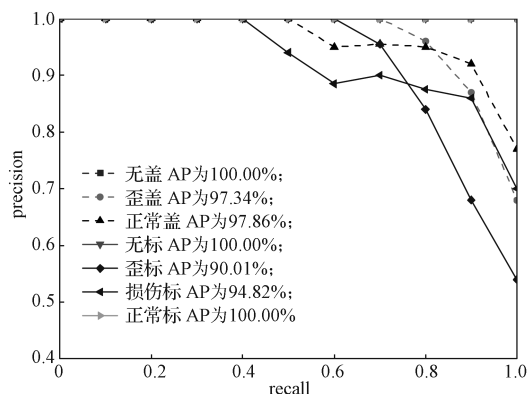


图 11 PR 曲线图

Fig. 11 Curve of PR

由图 11 可以看出，本文所提方法的检测 mAP 值可达 97.15%，能满足实际检测需要。此外，本模型对于歪标、损伤标的缺陷检测 AP 值略低，这可能是由于这两种缺陷的训练样本较少而导致模型对于这些缺陷的泛化能力不足。

表 1 改进中各模型性能测试比较

Table 1 Comparison of performance tests of each model in this study

检测方法	参数量 / 10^7	AP/%							mAP/%	检测时间 /ms
		无盖	歪盖	正常盖	无标	歪标	损伤标	正常标		
YOLOv4	6.443	100.00	98.10	99.62	100.00	87.09	95.76	100.00	97.22	184
MobileNetV3_PANet	4.004	95.94	88.82	94.66	100.00	83.33	93.39	100.00	93.73	117
MobileNetV3_DSC-PANet	1.179	94.31	70.46	92.41	100.00	97.69	88.47	87.42	90.11	52
MobileNetV3_DSC-FPANet	1.244	95.69	89.65	94.55	100.00	83.33	93.39	100.00	93.80	53
MobileNetV3_DSC-SE-FPANet	1.548	99.83	82.99	91.89	100.00	86.75	94.58	100.00	93.72	60
本文方法	1.451	100.00	97.34	97.86	100.00	90.01	94.82	100.00	97.15	61

由表 1 可以看出，只有当主干网络和增强网络均采用 DSC 策略时，模型才能真正做到轻量化，但随之带来的是 mAP 值急剧下降；本文方法的 mAP 值仅比 YOLOv4 低 0.07 个百分点，但单个样本的检测时间缩短，约为 YOLOv4 的 1/3。由此可见，改进后算法在保证精度的前提下极大提升了检测速度，这比较符合黏稠食品高速灌装生产成品缺陷检测的需要。

此外，为了综合评价本文模型的性能优劣，本课题组将已有成熟目标检测算法 YOLOv3^[13]、YOLOv4^[12]、SSD^[23]、Faster-RCNN^[17] 与本文方法进行对比。各种检测算法的测试结果如表 2 所示。

表 2 各目标检测算法性能测试比较

Table 2 Comparison of performance tests of each object detection algorithm

检测方法	参数量 / 10^7	mAP/%	测试时间 /ms
YOLOv3	6.200	92.60	171
YOLOv4	6.443	97.22	184
SSD	2.628	97.55	104
Faster-RCNN		98.93	835
本文方法	1.451	97.15	61

在表 2 中，由于 Faster-RCNN 参数计算量统计困难，因此本文未能给出。由表 2 可以看出，YOLOv3 的精度和检测速度均无法满足要求；虽然 Faster-RCNN 的检测精度最高，但是检测速度最慢；本文方法既保证了精度与其它算法相差无几，又实现了检测速度的大幅提升，因此在实际生产中本文方法更具有优势。

4.2.3 性能对比分析

本课题组在加入新的策略、设计新的网络时，做了大量的成品缺陷检测实验，进而不断优化模型，以得到最优模型。改进中各模型性能测试结果如表 1 所示。

5 结语

本研究提出了一种基于 YOLOv4 目标检测算法的黏稠食品轻量级灌装成品缺陷检测方法，并在灌装成品缺陷数据集上验证了该方法的有效性。研究结果表明：1) 本文方法能够实现对于黏稠物料灌装成品缺陷的多目标、多类别同时检测，有效降低了生产线检测设备的成本；2) 深度可分离卷积策略的应用使得模型参数计算量得到大幅下降，检测时间大大缩短，约为 YOLOv4 的 1/3，满足了灌装成品缺陷高速检测要求；3) 增强网络中，顶层特征与底层特征的融合设计和 ECA 的引入，加强了模型对目标特征的优质表达，进而实现了在模型参数下降的基础上保持较好的检测精度。

参考文献：

- [1] ARUN PRAKASH C, SAILESHWAR C S, RAJKUMAR R, et al. Design of a Vision Based Bottle Cap Inspection System[J]. Applied Mechanics & Materials, 2015, 787: 917-921.
 - [2] TOXQUI-QUITL C, CARDENAS-FRANCO J, PADILLA-VIVANCO A, et al. Bottle Inspector Based on Machine Vision[C]//Image Processing: Machine Vision Applications VI. Burlingame: SPIE, 2013, 8661: 1-10.
 - [3] 彭玉, 王耀南, 周显恩, 等. 基于轮廓曲率的 PET 瓶防盗环断裂检测 [J]. 电子测量与仪器学报, 2017, 31(3): 422-429.
- PENG Yu, WANG Yaonan, ZHOU Xian'en, et al.

- PET Bottle Tamper-Evident Bandbroken Detection Based on Contour Curvature[J]. *Journal of Electronic Measurement and Instrumentation*, 2017, 31(3): 422–429.
- [4] 张昌凡, 孙琳, 孙晓. 基于LSTM-Attention的高黏稠食品灌装流量检测[J]. *包装学报*, 2020, 12(1): 36–42.
ZHANG Changfan, SUN Lin, SUN Xiao. Detection for Filling Flow of High Viscosity Food Based on LSTM-Attention[J]. *Packaging Journal*, 2020, 12(1): 36–42.
- [5] KRIZHEVSKY A, SUTSKEVER I, HINTON G. ImageNet Classification with Deep Convolutional Neural Networks[J]. *Communications of the ACM*, 2017, 60(6): 84–90.
- [6] 俞芳芳, 董蓉, 李勃, 等. 基于深度学习的食用油灌装质量检测系统[J]. *中国粮油学报*, 2018, 33(3): 112–118.
YU Fangfang, DONG Rong, LI Bo, et al. An Edible Oil Filling Quality Detection System Based on Deep Learning Networks[J]. *Journal of the Chinese Cereals and Oils Association*, 2018, 33(3): 112–118.
- [7] 林景栋, 吴欣怡, 柴毅, 等. 卷积神经网络结构优化综述[J]. *自动化学报*, 2020, 46(1): 24–37.
LIN Jingdong, WU Xinyi, CHAI Yi, et al. Structure Optimization of Convolutional Neural Networks: A Survey[J]. *Acta Automatica Sinica*, 2020, 46(1): 24–37.
- [8] HOWARD A G, ZHU M L, CHEN B, et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications[J]. 2017, arXiv: 1704.04861.
- [9] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE, 2014: 580–587.
- [10] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-Time Object Detection[C]//2016 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). Las Vegas: IEEE, 2016: 779–788.
- [11] ZOU Z X, SHI Z W, GUO Y H, et al. Object Detection in 20 Years: A Survey[J]. 2019, arXiv: 1905.05055.
- [12] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal Speed and Accuracy of Object Detection[J]. 2004, arXiv: 2004.10934.
- [13] REDMON J, FARHADI A. YOLOv3: An Incremental Improvement[J]. *Computer Vision and Pattern Recognition*, 2018, arXiv: 1804.02767.
- [14] WANG C Y, LIAO H Y M, WU Y H, et al. CSPNet: A New Backbone that can Enhance Learning Capability of CNN[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops(CVPRW). Seattle: IEEE, 2020: 1571–1580.
- [15] HE K M, ZHANG X Y, REN S Q, 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.
- [16] LIU S, QI L, QIN H F, et al. Path Aggregation Network for Instance Segmentation[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). Salt Lake City: IEEE, 2018: 8759–8768.
- [17] REN S Q, HE K M, 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.
- [18] HOWARD A, SANDLER M, CHEN B, et al. Searching for MobileNetV3[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul: IEEE, 2019: 1314–1324.
- [19] 何静, 程涛, 黄良辉, 等. 深度可分离卷积神经网络在自动分拣中的应用[J]. *包装学报*, 2018, 10(6): 33–40.
HE Jing, CHENG Tao, HUANG Lianghui, et al. Depth Separable Convolutional Neural Network and Its Application in Automatic Sorting[J]. *Packaging Journal*, 2018, 10(6): 33–40.
- [20] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-Excitation Networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(8): 2011–2023.
- [21] WANG Q L, WU B G, ZHU P F, et al. ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE, 2020: 11531–11539.
- [22] LI KY, CAO L. A Review of Object Detection

Techniques[C]//2020 5th International Conference on Electromechanical Control Technology and Transportation (ICECTT). [S. l.]: IEEE, 2020: 385–390.

Conference on Computer Vision(ECCV). [S. l.]: Springer, 2016: 21–37.

[23] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single Shot MultiBox Detector[C]//14th European

(责任编辑: 邓 彬)

Defect Detection in Filling Products of Viscous Food Based on Lightweight YOLOv4

ZHANG Changfan¹, MENG Dezhi¹, WANG Yannan²

(1. College of Electrical and Information Engineering, Hunan University of Technology, Zhuzhou Hunan 412007, China;

2. Guangzhou Tech-Long Packaging Machinery Co., Ltd., Guangzhou 510530, China)

Abstract: The lightweight filling product defect detection method based on YOLOv4 object detection algorithm was proposed aimed at the issue of difficulty in detecting multiple targets at the same time in the automatic filling of viscous liquid food with the traditional filling product defect detection method. The lightweight feature extraction was performed on the input samples through the MobileNetV3 backbone feature extraction network, and the depthwise separable convolution strategy was used to reduce the computational cost of the enhance feature extraction network. Then, the full path aggregation network (FPANet) was designed and efficient channel attention (ECA) mechanism was introduced to improve the target feature expression of the enhance feature extraction network. The model training and precision testing were carried out on the designed lightweight network, and the performance of other object detection algorithms was compared in the same dataset to reveal the superiority and disadvantages of the proposed method. The experimental results showed that the proposed method could improve the detection speed while maintaining the accuracy, and the multi-objective high-speed detection of the filling product defects of viscous food was realized.

Keywords: viscous food; defect detection; depthwise separable convolution; attention mechanism