

DOI: 10.13652/j.issn.1003-5788.2020.03.026

改进 YOLOv3 算法的筷子毛刺缺陷检测方法

Defect detection method of chopsticks based on improved YOLOv3 algorithm

陈俊松 何自芬 张印辉

CHEN Jun-song HE Zi-fen ZHANG Yin-hui

(昆明理工大学机电工程学院, 云南 昆明 650500)

(Mechanical and Electrical Engineering, Kunming University of Science and Technology, Kunming, Yunnan 650500, China)

摘要:针对目前筷子质量检测机器并不能对带有毛刺的筷子进行有效分拣这一问题,提出了一种基于改进 YOLOv3 算法的筷子毛刺缺陷检测方法。通过删除 YOLOv3 网络多尺度检测中的 32 倍降采样的检测层,在 YOLOv3 网络中增加 4 倍降采样层,进一步得到深层特征并与第二次下采样中的浅层特征进行融合,让网络同时学习深层和浅层特征、重新聚类 anchor box 尺寸、改变 YOLOv3 网络的超参数,如减小抖动、减小权重衰减正则项、增大批尺寸、选择合适的动量值等方法,对原网络进行改进。当交并比(IOU)为 50 时改进后网络的平均检测精度由 89% 提升到了 94%,查准率提高了 4%,查全率提高了 9%,平均 IOU 提高了 3.5%,平均检测速度由 16.8 帧/s 增加到了 21.0 帧/s。试验结果表明,该方法相对于传统筷子质检机具有更高的检测效率,能满足筷子毛刺缺陷的检测需求。

关键词:YOLO 网络;筷子毛刺检测;卷积神经网络

Abstract: The current chopstick quality inspection machines on the market cannot effectively sort chopped chopsticks with burrs. Aiming at this problem, in this paper, a method for detecting burr defects of chopsticks based on improved YOLOv3 algorithm is proposed. By removing the 32x down-sampling detection layer in the YOLOv3 network multi-scale detection, adding a 4x down-sampling layer in the YOLOv3 network to further obtain deep features. Thereafter, it was fused with the shallow features in the second down-sampling, and let the network learn the deep and shallow features and re-cluster the anchor box size, with changing the hyper-parameters of the YOLOv3 network,

including reducing jitter and the weight-decay regular term, and increasing the batch size. Finally, a suitable momentum value was selected to improve the original network. When IOU = 50, the average detection accuracy of the improved network increased from 89% to 94%, and the accuracy rate increased by 4%, with the recall rate increasing by 9% and the average IOU increasing by 3.5%. The average detection speed increased from 16.8 to 21.0 frames per second. The experimental results showed that the method in this study had higher detection efficiency than the traditional chopstick quality inspection machine, which could meet the detection needs of chopstick burr defects.

Keywords: YOLO network; chopsticks burr detection; convolutional neural network

筷子在中国有着巨大的市场需求,据不完全统计^[1],中国目前每年大概需要消耗 450 亿双一次性筷子,因此筷子加工质量检测技术有着很大的市场需求。目前,中国的部分中小型筷子加工制造企业仍然采用人工检测的方法,随着筷子生产效率的进一步提高,传统的人工检测方法已经不能满足生产的需求。也有一部分中大型企业采用质检机检测,但现有的筷子质检机大多通过光纤检测出弯曲、细、半截料、发黑、发霉等不合格产品,再通过 PLC(处理器)控制气动电磁阀自动去除不合格的产品,但不能对带有毛刺的筷子进行准确分拣。

卷积神经网络是一种常用于分析和处理视觉数据的深度学习模型或多层感知机,先由卷积层对图像进行特征提取,再通过分类器对卷积层所提取到的特征进行分类。由于卷积神经网络具有简单、高效的特点,因此近年来卷积神经网络得到了快速发展,越来越多的传统工作中开始运用机器视觉与图像处理的方式来代替原来人眼的劳动^[2],例如针对自然场景下苹果图像的检测需求,程鸿芳等^[3]提出了一种基于改进 LeNet 卷积神经网络的苹

基金项目:国家自然科学基金(编号:61761024,61461022)

作者简介:陈俊松,男,昆明理工大学在读硕士研究生。

通信作者:何自芬(1976—),女,昆明理工大学副教授,硕士生导师,博士。E-mail:zyhhzf1998@163.com

收稿日期:2019-12-16

果目标识别模型,并利用该模型对不同场景的苹果图像进行识别训练与验证,试验结果表明该算法具有较强的抗干扰能力,图像识别速度快、识别率更高。胡小慧等^[4]提出基于图像分割的筷子瑕疵检测研究,利用图像处理技术对筷子图像进行增强、边缘检测及图像分割,获取筷子图像的 RGB 数据;对 RGB 三通道所占百分比及百分比方差进行统计分析最终能够对青皮、色斑、烧焦 3 种瑕疵进行处理分析。

笔者拟提出一种基于改进 YOLOv3(You Only Look Once, YOLO)算法的筷子毛刺缺陷检测方法,通过对 YOLO 算法的改进得出一个检测效果更好的网络模型,以解决带毛刺筷子的检测问题,进一步提高筷子加工业的自动化水平。

1 YOLO 的介绍

YOLOv1 算法是由 Redmon 等^[5]于 2015 年提出的,它主要是采用直接回归的思想,对输入的图像进行 $S \times S$ 的网格划分,通过物体中心所在网格对其进行预测,每个网格需要预测 B 个边界框以及 C 个类别概率。YOLOv1 算法通过设置阈值过滤掉概率低的候选框,再对保留的候选框进行非极大值抑制,得到最终的检测结果。该法存在诸多缺点,如对于密集的小目标种群以及近距离目标的检测效果不理想;检测结果的定位误差较大等。

Redmon 等^[6]针对 YOLOv1 缺点进行了改进,提出了 YOLOv2 网络模型。YOLOv2 网络主要是在分类的准确度不变的情况下,提高了网络的查全率以及定位精度。YOLOv2 模型通过提高训练图像的分辨率;在网络中添加 Batch Normalization 层,从而提高网络收敛性,同时还消除了其对其他形式的正则化的依赖。取消全连接层,使用 anchor boxes 来预测目标方框等措施进行改进。YOLOv2 网络包含了 19 个卷积层以及 5 个最大池化层,所以也被称为 Darknet19,最终 YOLOv2 模型在识别精度、定位精度、查全率等方面都有较大的提高。

针对两个近距离的同类或者异类目标的检测这一问题。很多新的算法不断地被提出,YOLOv3 便是其一,YOLOv3 网络在近距离目标与小目标的检测上有很好的鲁棒性,在一定程度上解决了这一问题。YOLOv3 网络是在 YOLOv2 模型的基础上,针对 YOLOv2 模型的一些问题又作了一些改进,连续使用 3×3 和 1×1 卷积层;参考残差网络,在新的网络的一些层之间添加快捷链路,形成一个深度达到 53 层的网络;利用多尺度特征进行对象检测,通过获得不同尺度的感受野,来加强对不同尺度对象的检测;用能够支持多标签对象的 Logistic 分类器替换 YOLOv2 模型中 softmax 分类器^[7]。经过改进最终提出了 YOLOv3 模型^[8],也被称 Darknet53。

2 改进的 YOLO 算法

文中通过删除 YOLOv3 网络多尺度检测中的 32 倍降采样的检测层、在 YOLOv3 网络中增加 4 倍降采样层,进一步得到深层特征并与第二次降采样中的浅层特征进行融合,让网络同时学习深层和浅层特征后进行检测,重新聚类与待检测目标最合适的 anchor box 尺寸、改变 YOLOv3 网络的超参数,减小抖动、减小权重衰减正则项、选择合适的动量值等方法对原网络进行改进。最终得到相对于原来的网络具有更高的平均精度、查全率、查准率、检测速度以及平均交并比(intersection over union, IOU)的新的网络模型。

YOLOv3 原网络中参考 SSD^[9]的思想分别在 32 倍降采样、16 倍降采样、8 倍降采样等 3 个不同的地方进行了 3 次检测。其中 32 倍降采样的感受野最大,主要是用于大型目标的检测,而 16 倍降采样适合检测中等大小的物体,8 倍的感受野最小,则是主要是用于小型目标的检测,由于筷子毛刺属于中小目标检测,因此文中删除了原 YOLOv3 网络中 32 倍降采样时的 YOLO 检测层,获得新的网络模型。

网络层越深,特征表达效果越好。在 YOLOv3 网络中使用 up-sample(上采样),能使 16 倍降采样和 8 倍降采样使用深层特征。YOLOv3 网络为了让网络同时学习深层和浅层特征,获得更好的表达效果,因此使用 route 层把 16 倍采样得到的 feature map 与第 4 次下采样得到的浅层特征进行融合,将 8 倍采样得到的 feature map 与第 3 次下采样得到的浅层特征进行融合。为了使网络得到更好的检测效果,在 8 倍采样后面增加了一个具有更小感受野的 4 倍降采样,并利用 route 层使其与网络中的第 2 次下采样的浅层特征进行融合然后用于 YOLO 层的检测,相对于原网络,改进后的网络平均检测精度提升了 1.5%、查准率提高了 5.0%、查全率提高了 1.0%。

Anchor box 的作用是解决传统检测中,一个窗口只能检测一个目标以及多尺度预测的问题。YOLOv3 先验框与 Faster R-CNN^[10]和 SSD 的设定方法不同,Faster R-CNN 和 SSD 采用人工设定的方法,在 YOLOv3 中采用 K-means 聚类的方式确定 Anchor box 尺寸。K-means 算法是基于距离的聚类算法,基本思想是随机选取初始聚类中心,再计算每个样本点到初始聚类中心的欧式距离,按照距离最近的准则将它们分配给相似度最大的聚类中心所代表的类。计算每个类别所有样本点的均值,更新聚类中心,直到目标准则函数收敛为止^[11],见图 1。通过 K-means 聚类的方法重新计算了多组,并与文中数据集相适应的且在 416×416 网络输入下的 anchor box 尺寸。当 anchor box 尺寸为 4,7,7,15,10,13,16,30,33,23,30,

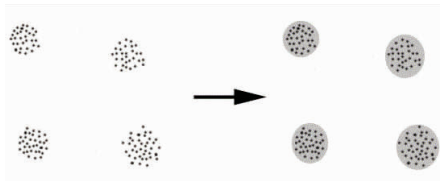


图 1 K-means 聚类法示意图

Figure 1 K-means clustering diagram

61,62,45,59,119,116,90 时的网络相对于原网络的平均检测精度提升了 2.6%、查准率降低了 2.0%、查全率提高了 6.0%。

利用数据抖动产生更多数据,防止过拟合。抖动是裁剪的参数,抖动 0.1,就在 0.0~0.1 中裁剪。试验中当抖动为 0.1 时,改进后的网络相对于原网络,平均检测精度提升了 4.31%、查准率提高了 2.00%、查全率提高了 8.00%。抖动为 0.1 时裁剪效果如图 2 所示。



(a) 原图 (b) 裁剪后

图 2 裁剪效果示意图

Figure 2 Crop effect diagram

在损失函数中加入一些正则项可以防止网络的过拟合。在随机梯度下降(SGD)^[12]中加入正则项对损失函数进行规范化:

$$\omega_i \leftarrow \omega_i - \eta \frac{\partial E}{\partial \omega_i} - \eta \lambda \omega_i, \quad (1)$$

$$Z = wx + b, \quad (2)$$

式中:

λ —— 权值衰减参数;

ω_i —— 权重;

b —— 偏置;

Z —— 输出值;

E —— 损失函数;

η —— 学习率。

由式(1)可知,权值衰减参数 λ 增大,权重 ω_i 则会衰减。由式(2)可知,权重 ω_i 的衰减使得 Z 值不断减小,实际上 Z 的取值范围很小,因此激活函数会相对线性,整个神经网络会计算非线性函数近的值,这个线性函数相对简单且不会发生过拟合。权值衰减参数 λ 越大对过拟合的抑制能力越强。在文中通过网络搜索的方式对 0.001 5、0.000 5(基准)、0.000 3、0.000 2、0.000 1 5 组参数进行训练,结果显示,当权重衰减正则取 0.000 1 时改进

后的网络相对于原网络,平均检测精度提升了 5.0%、查准率提高了 4.0%、查全率提高了 9.0%。

动量影响梯度下降到最优值的速度,由式(3)可知梯度下降法参数的更新只与当前梯度有关。但动量梯度下降法是先求每个 mini-batch 的梯度,然后再进行指数加权平均得到新的参数 V_{vw} 与 V_{vb} 用以更新之前的参数,由于指数加权平均后的梯度中含有之前梯度的更新方向的信息,所以利用指数加权平均之后的梯度对参数进行更新得出最终的更新方向,是一种较优的梯度下降方法,其过程如式(4)所示。动量梯度下降算法不仅能够增加稳定性而且还能避免局部最优。若当前梯度与历史梯度一致则增强此梯度,反之则梯度衰减抑制震荡,其作用效果如图 3 所示。通过试验,当动量取 0.9 时,平均精度达到最大值。

$$\begin{cases} W = W - \alpha \nabla W \\ b = b - \alpha \nabla b \end{cases}, \quad (3)$$

$$\begin{cases} V_{vw_0} = 0 \\ V_{vw_1} = \beta V_{vw_0} + (1 - \beta) \nabla W_1 \\ V_{vw_2} = \beta V_{vw_1} + (1 - \beta) \nabla W_2 \\ \vdots \end{cases}, \quad (4)$$

式中:

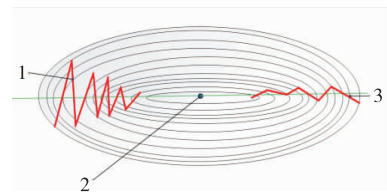
W —— 权重;

b —— 偏置;

α —— 步长;

β —— 加权参数;

V_{vw_i} —— 指数加权平均数。



1. 梯度下降法 2. 极值点 3. 动量梯度下降法

图 3 梯度下降示意图

Figure 3 Gradient descent diagram

3 超参数优化结果

表 1 中所列为 Anchor box 尺寸、权重衰减正则项、抖动、动量 4 个参数的不同取值组。用 darknet53.conv.74 预训练权重对表 1 中的不同参数依次由上往下,逐步使用网络搜索的方法进行训练,每次训练次数为 4 000 次,使得网络的平均检测精度、查全率、平均 IOU、F1-score 逐步得到提高,最终得到更优的网络模型。

通过网络搜索的方式对数组参数分别进行训练,取到其中的最优值。通过试验所确定的各组最优参数见表 2。

表 1 试验参数取值

Table 1 Experimental parameter value

参参数名	参数取参数取值
Anchor box 尺寸	16,14;24,15;20,18;27,19;33,16;34,20;40,17;50,15;51,20(基准) 17,14;24,15;23,19;36,16;31,19;47,15;41,19;52,17;55,20 17,14;23,15;26,19;32,16;39,17;34,20;48,15;46,20;57,18 4,7;7,15;10,13;16,30;33,23;30,61;62,45;59,119;116,90
抖动	0.1,0.2,0.3(基准),0.4,0.5
权重衰减正则项	0.000 1,0.000 2,0.000 3,0.000 5(基准),0.001 5
动量	0.90(基准),0.91,0.92,0.93,0.95

表 2 试验最优参数

Table 2 Experimental optimal parameters

Anchor box 尺寸	抖动	权重衰减正则项	动量
4,7;7,15;10,13;16,30;33,23;30,61;62,45;59,119;116,90	0.1	0.000 1	0.90

4 试验与结果讨论

4.1 数据集的准备

数据集主要通过相机拍摄的方式获取。文中的数据集全部是来源于 C 型一次性竹筷。为了增加模型的泛发性和训练精度,文中数据集在获取图像时,增加了许多不同形态特征、不同角度、不同位置的筷子毛刺的图片。筷子上既可能包含特征较为明显的毛刺也可能包含特征不明显的毛刺(见图 4)。毛刺可能出现在筷子上的任何位置,且由于拍摄角度和筷子本身姿态的影响,同一个筷子毛刺可能会在图像中表现出不同的特征(见图 5)。筷子上的毛刺数量不一,可能只有一个也可能有多个(见图 6)。数据集总共包含了 303 图片,其中 222 张图片作为训练集,81 张图片作为测试集。由于筷子毛刺检测环境并不复杂,因此不需要人工制造负样本数据用于网络训练,采用 YOLO 网络中的 Grid Cell 机制,来确定相应的正负样本。其作用是通过计算其与一个或者多个 Object 的 Ground Truth Box 的 IOU,再选择特定阈值对正、负样本进行分类。数据集标注时只标注毛刺这一类缺陷。

4.2 试验平台

试验在 PC 端完成。PC 主要配置为:i7 处理器、电脑运行内存为 16G、GPU(GTX-1060)、显存为 6G。开发环



(a) 不明显的毛刺特征 (b) 明显的毛刺特征

图 4 不同程度毛刺筷子的示意图

Figure 4 Schematic diagram of chopsticks with different burrs



(a) 姿态a (b) 姿态b

图 5 不同姿态毛刺筷子的示意图

Figure 5 Schematic diagram of chopsticks burrs in different poses



(a) 多处有毛刺



(b) 仅一处有毛刺

图 6 不同数量毛刺筷子的示意图

Figure 6 Illustration of burrs with different number of chopsticks

境为 Visual Studio2015,需要配置 CUDA9.0、CUDANN9.0、OpenCV3.0 和 YOLOv3 物体检测框架。

4.3 筷子毛刺缺陷检测工作流程

文中提出的改进 YOLOv3 算法的筷子毛刺缺陷检测方法具体的检测过程主要有以下几个步骤:① 通过相机获取待检测筷子的图像;② 将获取到的图像传输到训练好的网络模型中进行检测,判断是否具有毛刺;③ 获取模型的检测结果,若未检测到毛刺则为合格,若检测到毛刺则对此筷子进行剔除;④ 对合格的筷子进行收集。工程流程图见图 7。

4.4 结果与分析

用改进后的网络模型对 darknet53.conv.7 预训练模型进行训练,训练次数为 4 000 次,得到一系列检测模型,

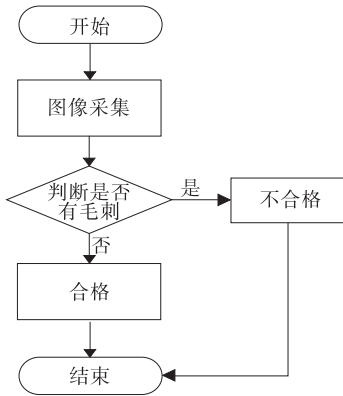


图7 筷子毛刺检测工作流程图

Figure 7 Flowchart of chopstick burr detection

其结果见图8。由图8可知,训练到4000次时检测模型的平均检测精度最高,此时的平均loss值降到了0.070。原网络的训练结果见图9。由表3可知,当IOU为50时改进后的网络的平均检测精度提高了5.0%,查准率提高了4.0%,查全率相比提高了9.0%,平均IOU提高到了3.5%,F1-score提高到了0.92,检测速度由16.8帧/s提升到了21帧/s,从而每小时检测的图像数为75600张,已经超过了目前市面上筷子质检机每小时70000只的检测效率。综上所述可知,相比原网络模型,改进后的网络模型在平均检测精度、查准率、查全率、检测速度等各方面都有较大的提升。

4.5 测试对比结果分析

IOU为50时,分别对原网络的检测模型与改进后的最优检测模型进行测试,其结果见图10、11。在图10(a)原网络中从左往右,两处缺陷的检测精度分别为92%,

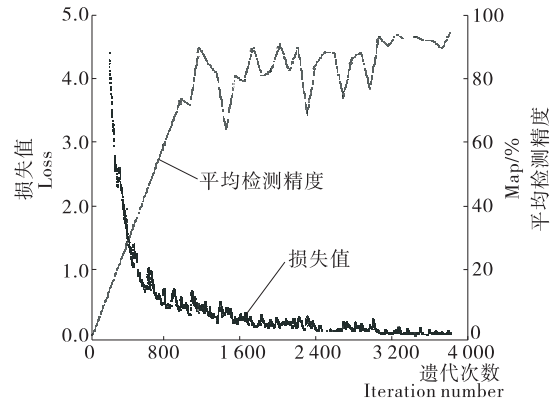


图8 改进后网络训练结果示意图

Figure 8 Schematic diagram of improved network training results

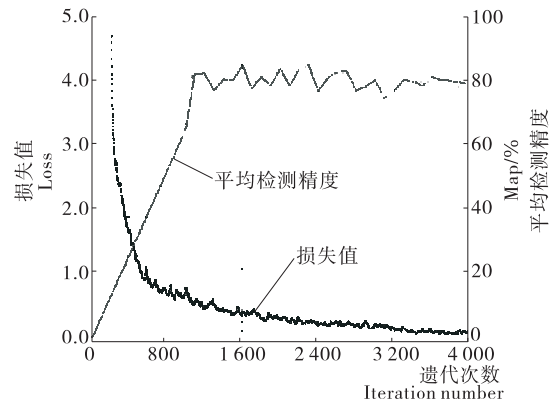


图9 原网络训练结果示意图

Figure 9 Schematic diagram of original network training results

表3 当IOU为50时最优模型检测结果

Table 3 Optimal model detection results when IOU=50

检测方法	平均检测精度/%	查准率/%	查全率/%	平均IOU	F1-score	每秒传输帧数
原网络	89	89	82	62.38	0.85	16.8
改进后的网络	94	93	91	65.87	0.92	21.0
提高率/%	5	4	9	3.49	0.07	4.2

65%。而在图11(a)改进后网络中从左往右,两个空闲车位的检测精度分别提升到了99%,91%。在图10(b)中原网络出现了误检的情况,图中从左往右,3个缺陷的检测精度分别为63%,76%,99%,但图中其实只有两处缺陷,分别为图中的筷子的中部和右部,筷子的左部并无毛刺缺陷,因此出现了误检的情况。而在图11(b)改进后网络中并未出现误检的情况,且图中从左往右,两个毛刺缺陷的检测精度分别提升到了99%、98%。由对比结果可以看出,由于改进后的模型有更高平均检测精度、查全率、查准率、平均IOU以及F1-score,所以检测效果明显

优于改进前的模型,不仅检测精度明显提高,且有更高的查全率以及更好的泛化能力,不易发生漏检的情况。

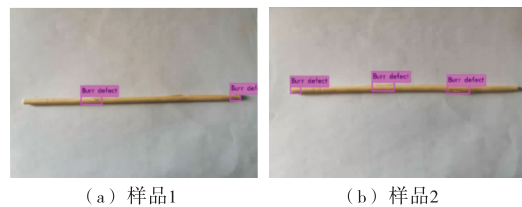


图10 原模型的测试结果

Figure 10 Test results of the original model



(a) 样品1 (b) 样品2

图 11 改进后模型的测试结果

Figure 11 Test results of the improved model

4.6 与其他目标检测网络的对比

采用具有较快检测速度与精度的 SSD 神经网络^[9]对文中的筷子毛刺缺陷数据集进行了训练检测,并与文中提出的方法进行对比,结果见表 4。通过对比可知,基于改进 YOLOv3 的检测方法与 SSD 神经网络算法相比,在准确率与检测速率的上都有明显的优势。

表 4 不同检测方法的训练结果

Table 4 Training results of different detection methods

I 检测方法	平均检测精度/%	每秒传输帧数
原网络	89	16.80
改进后的网络	94	21.00
SSD 神经网络	84	7.64

5 结语

研究提出了一种基于改进 YOLOv3 算法的筷子毛刺缺陷检测方法,通过删除 YOLOv3 网络多尺度检测中的 32 倍降采样的检测层、在 YOLOv3 网络中增加 4 倍降采样层并与第二次下采样进行融合、重新聚类 anchor box 尺寸、改变 YOLOv3 网络的超参数,减小抖动、减小权重衰减正则项、选择合适的动量值等方法对原网络进行改进,最终得到了新的网络模型。再利用新的网络模型进行训练,最后得到了一个相对原模型具有更好检测效果的检测模型,当交并比为 50 时其平均检测精度达到了 94%、查准率达到了 93%、查全率达到了 91%、平均检测速度 21 帧/s。该方法基本能满足筷子毛刺的检测需求。

在后续的研究中将通过进一步改进网络结构等方法继续提高模型的泛化能力和检测速度。

参考文献

- [1] 严晓明. 有关一次性筷子的研究报[J]. 数学大世界, 2013 (34): 32-33.
- [2] 郭静, 罗华, 张涛. 机器视觉与应用[J]. 电子科技, 2014 (7): 185-188.
- [3] 程鸿芳, 张春友. 自然场景下基于改进 LeNet 卷积神经网络的苹果图像识别技术[J]. 食品与机械, 2019, 35 (3): 155-158.
- [4] 胡小慧, 江虹, 郭秋梅. 基于图像分割的筷子瑕疵检测研究[J]. 微型机与应用, 2016, 35(2): 39-42.
- [5] 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 (CVPR), June 27-30, 2016. Las Vegas, NV, USA. [S.l.]: IEEE, 2016: 779-788.
- [6] REDMON J, FARHADIA. YOLO9000: better, faster, stronger[C]// Proceedings of the IEEE conference on computer vision and pattern recognition. [S.l.]: IEEE, 2017: 7 263-7 271.
- [7] 万磊, 佟鑫, 盛明伟, 等. Softmax 分类器深度学习图像分类方法应用综述[J]. 导航与控制, 2019, 18(6): 1-9.
- [8] 林健巍. YOLO 图像检测技术综述[J]. 福建电脑, 2019, 35 (9): 80-83.
- [9] LIU Wei, ANGUELOV D, ERHAN D, et al. SSD: Single shot multiBox detector[C]// European Conference on Computer Vision. Cham: Springer Press, 2016: 21-37.
- [10] REN Shao-qing, HE Kai-ming, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[C]// Advances in Neural Information Processing Systems. Cambridge: MIT Press, 2015: 91-99.
- [11] 丛思安, 王星星. K-means 算法研究综述[J]. 计算机技术应用, 2018(17): 155-156.
- [12] 王功鹏, 段萌, 牛常勇. 基于卷积神经网络的随机梯度下降算法[J]. 计算机工程与设计, 2018, 39(2): 441-445.

信息窗

韩国发布《食品添加剂法典》部分修改单

3月20日,韩国食品药品安全处(MFDS)发布了第2020-18号告示,修改《食品添加剂法典》的部分内容,其主要内容如下:

修改食用色素红色3号(赤藓红)、食用色素红色40号(诱惑红)、食用色素蓝色1号(亮蓝 FCF)、食用色

素黄色4号(柠檬黄)的使用标准,使其可以使用在咖啡中,使用限量均为0.1 g/kg以下,上述几种添加剂并用时,使用量之和应为0.1 g/kg以下。

(来源: <http://news.foodmate.net>)