基于深度学习的工业自动化包装缺陷检测方法

李建明[1]，杨挺[1]，王惠栋[2]
（1.天津大学，天津 300072；2.北京工业大学，北京 100124）

摘要：目的 针对目前工业自动化生产中基于人工特征提取的包装缺陷检测方法复杂、专业知识要求高、通用性差、在多目标和复杂背景下难以应用等问题，研究基于深度学习的实时包装缺陷检测方法。**方法** 在样本数据较少的情况下，提出一种基于深度学习的 Inception-V3 图像分类算法和 YOLO-V3 目标检测算法相结合的缺陷检测方法，并设计完整的基于计算机视觉的在线包装缺陷检测系统。**结果** 实验结果显示，该方法的识别准确率为 99.49%，方差为 0.000 050 6，只使用 Inception-V3 算法的准确率为 97.70%，方差为 0.000 251。**结论** 相比一般基于人工特征提取的包装缺陷检测方法，避免了复杂的特征提取过程。相比只应用图像分类算法进行包装缺陷检测，该方法在包装缺陷区域占比较小的情况下能较明显地提高包装缺陷检测精度和稳定性，在复杂检测背景和多目标场景中体现优势。该缺陷检测系统和检测方法可以很容易地迁移到其他类似在线检测问题上。

关键词：缺陷检测；Inception-v3；YOLO-V3；TensorFlow Serving；MQTT；迁移学习
中图分类号：TB487 **文献标识码：**A **文章编号：**1001-3563(2020)07-0175-10
DOI：10.19554/j.cnki.1001-3563.2020.07.025

An Industrial Automation Packaging Defect Detection Method Based on Deep Learning

LI Jian-ming[1], *YANG Ting*[1], *WANG Hui-dong*[2]
(1.Tianjin University, Tianjin 300072, China; 2.Beijing University of Technology, Beijing 100124, China)

ABSTRACT: The work aims to study a real-time packaging defect detection method based on deep learning, in view of the problems such as complexity, considerable professional knowledge, poor generality, and difficulty in application under multi-objective and complex background of the current packaging defect detection methods based on artificial feature extraction in industrial automation production. In the case of small sample set, a defect detection method combining the Inception-V3 image classification algorithm and YOLO-V3 target detection algorithm based on deep learning was proposed, and a complete online packaging defect detection system based on computer vision was designed. Experimental results showed that the recognition accuracy rate and variance of the proposed method were 99.49% and 0.000 050 6 respectively. The accuracy rate of using only Inception-V3 algorithm was 97.70% and its variance was 0.000 251. Compared with the general packaging defect detection method based on artificial feature extraction, the proposed method avoids the complex feature extraction process. Compared with the packaging defect detection only with image classification algorithm, the proposed method can obviously improve the accuracy and stability of packaging defect detection especially when the defect occupies a relatively small proportion, and performs well in complex detection background and multi-objective situation. At the same time, the defect detection system and detection method designed herein can be easily migrated to other

收稿日期：2019-09-05
作者简介：李建明（1989—），男，天津大学硕士生，主攻计算机视觉。
通信作者：杨挺（1979—），男，博士，天津大学教授，主要研究方向为人工智能与大数据、泛在电力物联网。

similar online detection problems.

KEY WORDS: defect detection; Inception-v3; YOLO-V3; TensorFlow Serving; MQTT; transfer learning

受各种因素的影响，工业生产中不可避免地会出现产品包装缺陷。随着经济的发展、生产效率的提高，基于传统的人工包装缺陷检测方法，不仅检测速度慢、效率低，而且检测过程中容易出错，难以满足现在高效的工业自动化生产要求。对此部分学者在包装缺陷检测方面提出了不同的方法：杨祖彬等[1]利用小波变换改进算法对图像边缘的检测，提出了基于图像配准的食品包装印刷缺陷检测方法；刘学福[2]采用改进的Itti/Koch视觉注意力算法和图像分块方差-加权特征值方法、主成分分析法实现了对缺陷药卷的识别和检测；方文星等[3]研究了基于特征提取SURF算法、视觉词汇BOW算法和一类支持向量机SVM算法的缺陷检测模型，实现了药品包装缺陷识别。这些检测方法虽然实现了包装缺陷自动检测，但是在多目标和复杂背景检测场景下受到限制，由于人工特征提取不能有效覆盖缺陷全部特征，模型的准确率有待提高，当被检测物体发生变化，所有的规则和算法都需要重新设计，很难满足工业生产实际需求。

近年来，人工智能-深度学习技术得到了快速发展，其已广泛应用到交通、医疗、气象等多个领域[4—12]。然而，深度学习在包装缺陷检测方面应用还较少。文中研究基于深度学习的工业自动化包装缺陷检测方法，设计基于gRPC、Redis和MQTT的模块化包装缺陷检测系统，并将其应用到某工厂自动化生产中的电线产品包装缺陷在线检测问题上，取得了较好的效果。

1 系统结构

文中提出的包装缺陷检测系统整体结构见图1，主要包括图像数据采集服务、图像识别服务、系统报警服务、数据转发服务、数据存储服务等5部分组成，它们共同完成了缺陷检测工作。图像数据采集服务：由部署在树莓派上的采集程序调用摄像头进行图像采集，经YOLO-V3模型处理得到可疑缺陷区域，将一个或者多个可疑区域截取后发送到图像识别服务器上。图像识别服务：由Inception-V3模型提供图像识别服务，对接收的可疑缺陷区域进行检测，依据检测结果决定是否触发报警系统和下发控制状态指令。系统报警服务：根据接收的控制状态指令，触发相关设备的控制单元PLC执行相应程序，控制声光报警装置和相应设备动作，同时将报警信息推送到微信设备管理群中。数据转发服务：数据转发主要采取了2种方式，图片数据转发采用Redis的分布式的发布订阅机制；PLC装置、报警装置的状态控制信息，采用了MQTT协议消息发布订阅机制。数据存储服务：记录缺陷检测服务中产生的图片数据和报控制数据，以满足后续研究和模型更新训练需要。

2 包装缺陷检测理论

文中提出的缺陷检测方法是基于卷积神经网络的深度神经网络。在缺陷检测中可以降低对图像数据预处理要求，避免了复杂烦琐的特征提取工作。卷积神经网络权值共享的结构，大大减少了神经网络的参数量，避免了过拟合的同时降低了神经网络的复杂度，解决了多层感知机网络中存在的梯度爆炸和梯度弥散问题，提升了识别的准确率[13—15]。

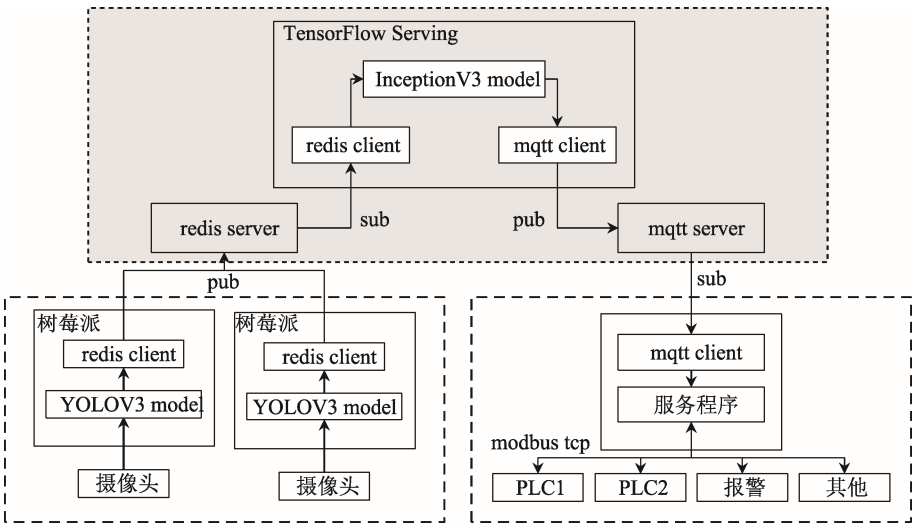


图 1 系统结构

Fig.1 System structure

2.1　YOLO-V3 缺陷区域检测

2.1.1　缺陷检测网络结构

为实现待检测图片中的缺陷区域识别，采用YOLO-V3目标检测算法，网络中没有RPN（Region Proposal Network）过程，将物体类别和位置检测统一为一个回归问题。缺陷特征提取网络采用Darknet-53，共有53层卷积层，其包含的ResNet的残差结构，提高了网络特征表达能力。YOLO-V3结构见图2，网络中共有75个卷积层，用卷积操作替换全连接层，减少了模型参数，可使用不同输入大小的图像；网络中没有最大池化层，用步长为2的卷积操作达到降维的效果；图2中res1，res2，res4，res8代表残差结构，通过一条捷径获取之前网络的基础特征，将学习目标从学习完整信息变成学习残差，缓解了深层网络中梯度弥散和爆炸问题。网络在多个特征图上进行多尺度预测detection 1，detection 2，detection 3，这种FPN（Feature Pyramid Network）结构，使当前较低层的特征图可以获取高层特征图的信息，高低阶特征有效融合，提升了检测精度[16—20]。

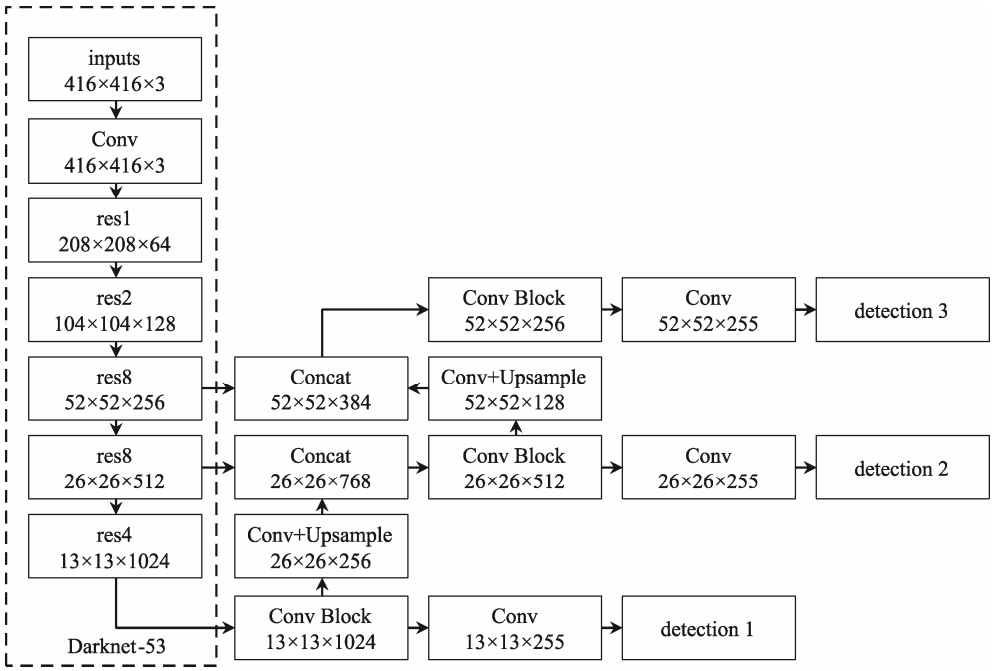


图 2　YOLO-V3 网络结构

Fig.2 Net structure of YOLO-V3

2.1.2　缺陷区域检测原理

可疑缺陷区域检测原理：YOLO-V3 是在 3 个尺度上，将输入图分成 $N\times N$（13×13，26×26，52×52）个格子，在每个格子上直接用逻辑回归预测 3 个框的参数，包括：中心坐标（x,y），框宽和高 w，h，是否包含有物体，以及每个类的置信度，共 $N\times N\times 3\times(4+1+C)$个参数，其中 $N\times N$ 为特征图的尺寸，C 为目标类别数。然后通过非极大值抑制（NMS）获取最优的框，得到目标类别置信度。边框预测公式见式（1—5），物理含义见图 3[16—18]。

$$b_x=\sigma(t_x)+c_x \tag{1}$$

$$b_y=\sigma(t_y)+c_y \tag{2}$$

$$b_{\mathrm{w}}=p_{\mathrm{w}}e^{t_{\mathrm{w}}} \tag{3}$$

$$b_{\mathrm{h}}=p_{\mathrm{h}}e^{t_{\mathrm{h}}} \tag{4}$$

$$P_{\mathrm{r}}(object)\times IOU(b,object)=\sigma(t_o) \tag{5}$$

式中：t_x，t_y，t_{w}，t_{h}，t_{o} 为网络要学习的参数，将

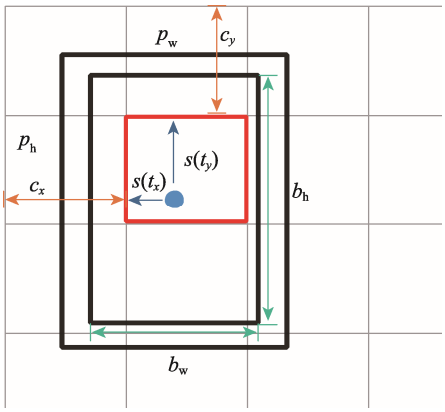


图 3　边框位置预测

Fig.3 Location prediction of bounding box

这些参数代入对应公式，可以求出预测边框的中心坐标b_x，b_y，宽度b_{w}和高度b_{h}，以及边框置信度；c_x，c_y为所在网格到图像左上角的距离；p_{w}，p_{h}为先验框的宽度和高度；$P_{\mathrm{r}}(object)$为是否含有目标；

IOU（*b*,*object*）为预测框和真实框的交并比，两者相乘作为预测边框的置信度；σ为Sigmoid激活函数见式（6）。

$$\sigma(x)=\frac{1}{1+e^{-x}} \tag{6}$$

σ函数将预测的中心点坐标b_x，b_y约束在红色标注的单元格中。这种约束使得模型更容易学习，预测更稳定。

式（7）是网络的损失函数由坐标误差e_{crd}，IOU误差e_{iou}，分类误差e_{cls}这3部分组成。各部分见式（8—10）。

$$loss=\sum_{i=0}^{s^2}\left(e_{\text{crd}}+e_{\text{iou}}+e_{\text{cls}}\right) \tag{7}$$

$$e_{\text{crd}}=\lambda_{\text{coord}}\sum_{i=0}^{S^2}\sum_{j=0}^{B}I_{ij}^{\text{obj}}\left\{\left[\left(x_i-\hat{x}_i\right)^2+\left(y_i-\hat{y}_i\right)^2\right]+\left[\left(\frac{w_i-\hat{w}_i}{\hat{w}_i}\right)^2+\left(\frac{h_i-\hat{h}_i}{\hat{h}_i}\right)^2\right]\right\} \tag{8}$$

$$e_{\text{iou}}=-\sum_{i=0}^{S^2}\sum_{j=0}^{B}I_{ij}^{\text{obj}}\left[\hat{C}_i\log(C_i)+(1-\hat{C}_i)\log(1-C_i)\right]-\lambda_{\text{noobj}}\sum_{i=0}^{S^2}\sum_{j=0}^{B}I_{ij}^{\text{noobj}}\left[\hat{C}_i\log(C_i)+(1-\hat{C}_i)\log(1-C_i)\right] \tag{9}$$

$$e_{\text{cls}}=-\sum_{i=0}^{S^2}I_i^{\text{obj}}\sum_{c\in\text{classes}}\left[\hat{p}_i(c)\log(p_i(c))+(1-\hat{p}_i(c))\log(1-p_i(c))\right] \tag{10}$$

式中：S^2为特征图划分网格数；B为每个网格预测的边框数目；i为网格序号；j为候选边框序号；I_{ij}^{noobj}为第 i 个网格的第 j 个候选边框中不存在对象；I_{ij}^{obj}为第 i 个网格的第 j 个候选边框中存在对象；x_i，y_i为第 i 个网格中边框的中心坐标；w_i，h_i为第 i 个边界框的宽度和高度；C_i为第 i 个网格中存在检测物的预测置信度；$p_i(c)$为第 i 个网格中检测物属于某一类的预测概率；λ_{coord}为调节预测边框的位置误差权重；λ_{noobj}为调节不存在对象的预测边框置信度的权重。训练检测模型时，通过最小化损失函数 loss 调节网络参数，使得网络可以准确地检测出所有可疑区域。在多目标和复杂背景下的模型检测可疑缺陷区域的效果见图 4。

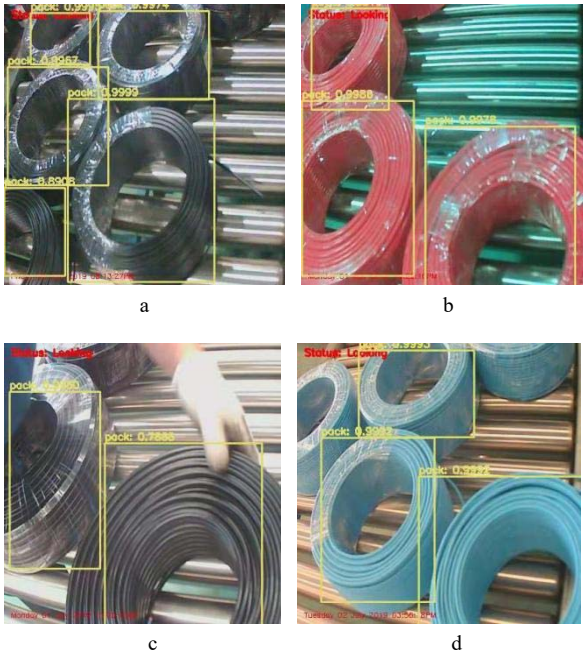


图 4 目标图片中可疑缺陷区域检测

Fig.4 Detection of the suspicious defect areas in the target image

表 1 Inception-V3 结构

Tab.1 Structure of Inception-V3

类型	卷积核/步(或注释)	输入尺寸
卷积	3×3/2	299×299×3
卷积	3×3/1	149×149×32
卷积	3×3/1	147×147×32
池化	3×3/2	147×147×64
卷积	3×3/1	73×73×64
卷积	3×3/2	71×71×80
卷积	3×3/1	35×35×192
Inception 模块	inception models×3	35×35×288
Inception 模块	inception models×5	17×17×768
Inception 模块	inception models×2	8×8×1280
池化	8×8	8×8×2048
线性	逻辑回归	1×1×2048
softmax	分类	1×1×2048

2.2 Inception-V3 缺陷识别

2.2.1 缺陷识别网络结构

经YOLO-V3缺陷检测模型检测后得到可疑缺陷区域，截取后分别对其做数据增强，得到一组或多组可疑区域数据组，使用迁移学习得到的Inception-V3缺陷识别模型，对这些可疑区域组进行缺陷识别后取平均值，得到识别结果。原始Inception-V3的网络结构见表1。首先3个3×3的卷积层串联，经池化层连接3个3×3卷积层，然后利用3类inception模块继续提取特征，最后经过和feature map相同大小的8×8卷积核池化操作转换成一维进行回归和分类[13,15,21]。文中迁移学习中，需要将最后的分类层根据样本集替换成新问题的分类层。

2.2.2 网络修改和迁移学习

将Inception-V3网络最后的dropout层、全连接层和softmax层替换为与缺陷样本集相符合的全连接层、softmax层。包装缺陷合格1类，不合格3类，新增全连接层需要4个输出神经元表示4个分类结果。根据实际需求，不需要对缺陷类别分类，故可以简化为

合格和不合格2种情况。

迁移学习有2种方式，当样本数据不足时，固定Bottleneck feature之前特征提取部分参数不变，仅训练修改后的全连接层网络参数。当样本数据充足时，首先固定Bottleneck feature之前特征提取部分参数不变，仅训练修改后的全连接层网络参数，训练几轮后（warm up），“解冻”部分或全部特征提取网络部分的参数，设置较小的学习率，微调整个网络[22]。文中考虑的不合格样本数量有限，选用了第1种迁移学习方式。

3　缺陷检测实例

3.1　数据集构建

3.1.1　数据采集和增强

由部署在树莓派3B中的程序采用帧差法调用USB摄像头进行数据采集，线盘的颜色有红、绿、蓝、黄、黑、白共6种，包装带有透明薄膜和白色塑料带2种。图5a是多个线盘的情况，图5d是只有边角的线盘情况，图5e是包装缺陷和包装合格共存的情况。图6中的图片是对图5a的水平翻转、垂直翻转、增加20%的亮度、降低60%的饱和度、灰度降低30%、中心按60%比例裁剪，将1张图片增强为7张图片，增加了样本容量。

3.1.2　数据集组成

将采集到的原始数据分为3类：全部带包裹（合格）、含不带包裹或者包裹有间隙（缺陷）、其他。其中‘其他’指类似于图5d没有目标线盘的图。根据实际不需要对缺陷进行分类，可简化模型，将样本分为2类：包装合格、包装缺陷。其中Inception-V3的数据集为：训练集[1280张，1266张]，测试集[100张，100张]。需要说明的是：包装合格的样本容易获取，训练集和测试集中的合格样本均为原样本图，包装缺陷样本相对不易获取，训练集和测试集中的缺陷样本是原200张缺陷图经数据增强后得到的1400张图片，剔除其中一些裁剪后无线盘或者线盘显示不足一半的图片后得到1366张增强后的图片，训练时将训练集的10%作为验证集；YOLO-V3的数据集是从Inception-V3的数据集中随机抽取120张图片，使用LabelImg工具对线盘位置标注生成VOC格式的xml文件，再通过脚本生产YOLO格式的txt文件，数据集结构见表2。

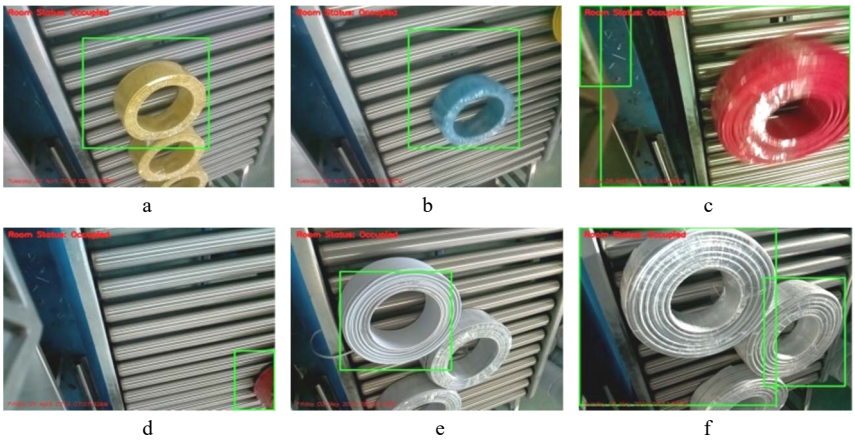


图 5　数据集示例

Fig.5 Samples of the data set

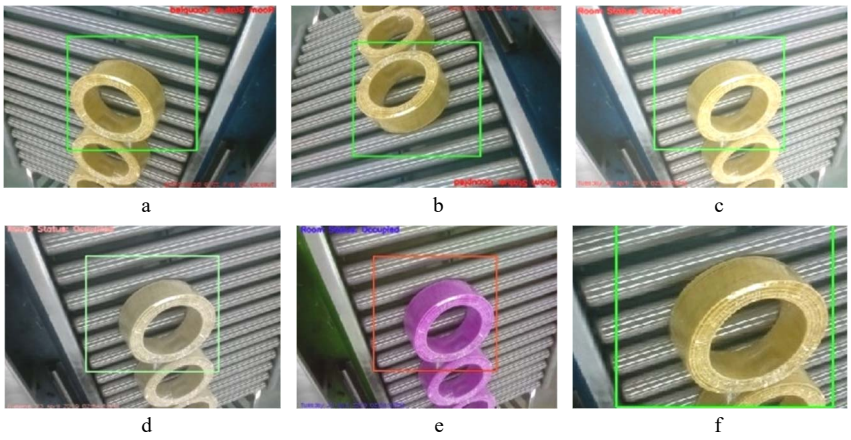


图 6　数据增强

Fig.6 Data augmentation

表 2 数据集组成

Tab.2 Composition of data set

训练类别	训练数量		测试数量	
	合格	缺陷	合格	缺陷
Inception-V3	1280	1266	100	100
YOLO-V3	50	50	10	10

3.2 模型训练

3.2.1 YOLO-V3 模型训练

为适应实验数据集，通过K-means聚类分析获取先验框尺寸，修改YOLO-V3的先验框尺寸，减小训练模型时微调先验框到实际位置的难度，使模型快速收敛。实验数据k=9时聚类点分别为：(109,146), (157,73), (180,160), (204,218), (242,117), (254,170), (285,240), (292,215), (440,295)。同13×13尺度上的默认先验框尺寸(116,90), (156,192), (373,326)相近，属于大目标，这和实验中使用原参数时只有13×13尺度上能检测到目标，其他尺度上未检测到目标的情况一致。实验发现比使用自己聚类得到的多个尺寸，在多尺度上训练模型效果要好，笔者认为，这是由于实验数据集较小，线盘目标均属于大目标，相比26×26和52×52尺度，13×13尺度上的特征图上包含的分辨率信息较少语义信息大，能更好地区分目标和背景，因此实验最终采用模型默认参数，能满足检测需求。

实验机为ubuntu 16.04系统 内存16 G，Intel i7-4980HQ 2.8 GHz处理器，NVIDIA GeForce GTX 980M 4 G显卡，设置批大小16，迭代次数5000，学习率策略：0～3000轮为0.01，2000～3000轮为0.001，3000～5000轮为0.0001。训练结果见图7—8，分别为模型训练平均损失变化曲线、13×13尺度上的IOU曲线。平均损失曲线在迭代开始损失较大，随着迭代次数增加损失快速降低，约在2000次基本处于稳定状态，在0.0045左右波动。IOU曲线同样在2000次后处于稳定状态，在0.88左右波动。

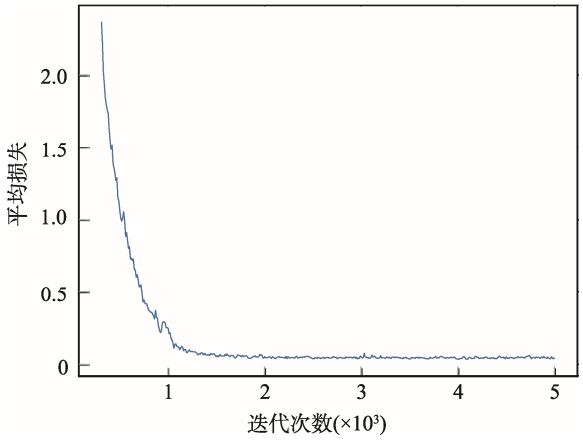


图 7 平均损失曲线

Fig.7 Average loss curve

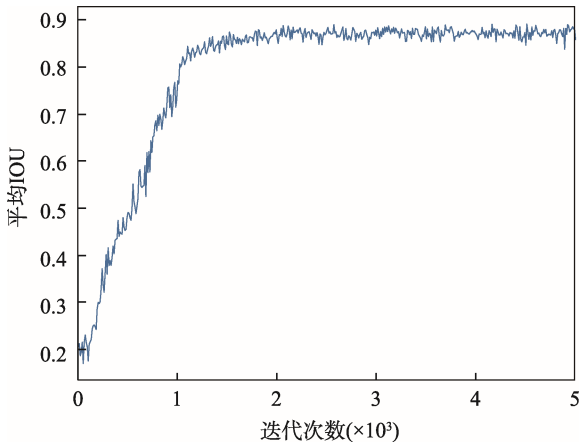


图 8 13×13 尺度上的平均 IOU 曲线

Fig.8 Average IOU curve on 13×13 dimension

3.2.2 Inception-V3 迁移学习

首先将Inception-V3的数据集，经训练好的YOLO-V3模型处理并整理得到新的数据集：[1350张，1206张]。如对图5中的样本缺陷区域检测，结果见图9，线盘均被精确地用黄色矩形框标出（图9d中线盘面积太小未被检测到）。对可疑区域进行截取见图10，

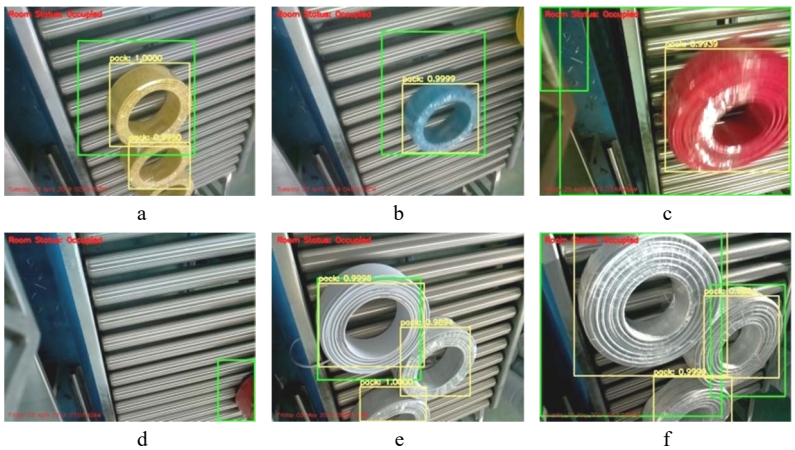


图 9 缺陷区域检测

Fig.9 Defect areas detection

图 10　可疑区域截取

Fig.10 Cropping of suspicious areas

由于图9e中最下面的线盘矩形框太小，未达到设定的阈值80×80，程序中被滤除。

采用迁移学习训练Inception-V3缺陷识别模型。训练参数设置为：批次大小100，迭代步数800，学习率为0.01。未使用和使用YOLO-V3处理得到的Inception-V3模型在训练集、验证集上的准确率和交叉熵损失曲线见图11—12。观察训练集和验证集上的曲线，两模型均没有出现明显过拟合迹象。对比图13所示的两模型曲线可看出，采用YOLO-V3处理得到的Inception-V3缺陷识别模型在400步后准确率稳定在99.49%，其方差为0.000 050 6，相比未使用YOLO-V3处理获得的模型准确率为97.70%，方差为0.000 251，曲线波动幅度均比较小，总体更加平稳，收敛速度更快。

为体现YOLO-Inception-V3检测方法在多目标和复杂景下的优势，在原测试集[100,100]基础上增加了50张同时包含缺陷和合格线盘的图片，50张没有线盘目标的图片，得到新测试集：[100,100,50,50]。如表3所示，实验结果表明YOLO-Inception-V3检测方法在只包含合格或者缺陷线盘图片上的识别准确率更高，对“均有”的图片能准确将图片识别为缺陷图片，对“其他”图片能完全过滤掉无目标线盘图片；而Inception-V3检测方法对“均有”类型的图片会将部分图片识别为合格，如实验中将50张“均有”缺陷图片中23张识别为合格，对无目标的“其他”图片无法进行过滤，实验中50张“其他”图片中29张识别为合格。因此，文中提出的YOLO-Inception-V3检测方法能应用于

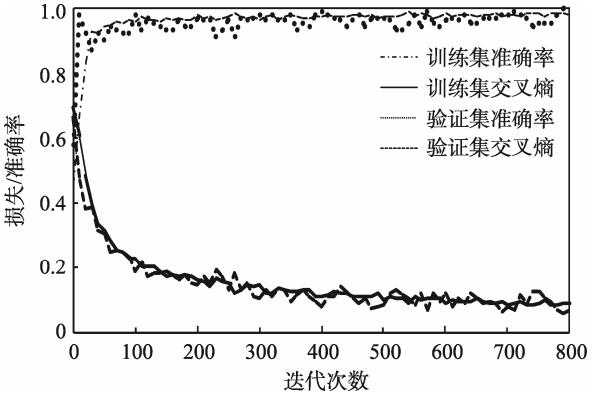


图 11　Inception 模型曲线

Fig.11 Curves of Inception model

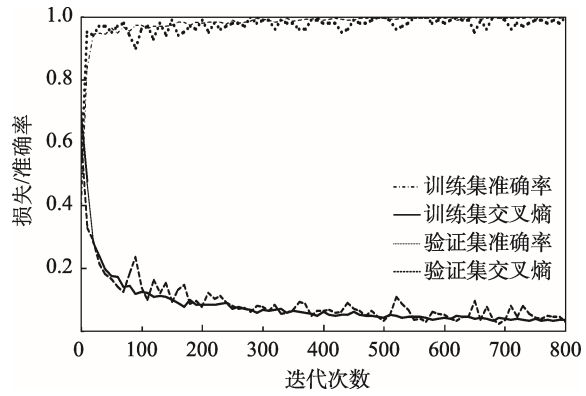


图 12　YOLO-Inception 模型曲线

Fig.12 Curves of YOLO-Inception model

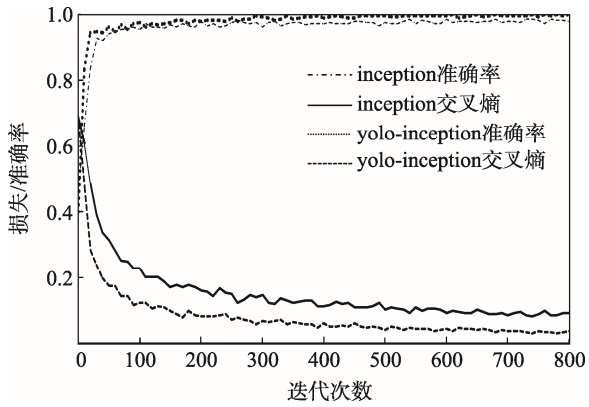


图 13　Inception 模型和 YOLO-Inception 模型曲线对比

Fig.13 Comparison of curves between Inception and YOLO-Inception models

表 3　模型在测试集上的运行结果

Tab.3 Model test results on test data set

模型类型	合格	缺陷		其他
	有且只有合格线盘(100 张)	有且只有缺陷线盘(100 张)	同时含有合格和缺陷线盘(50 张)	不含有线盘(50 张)
Inception-V3	93(合格)，7(缺陷)	5(合格)，95(缺陷)	23(合格)，27(缺陷)	29(合格)，21(缺陷)
YOLO-Inception-V3	98(合格)，2(缺陷)	0(合格)，100(缺陷)	0(合格)，50(缺陷)	滤除

复杂和多目标检测场景，可对不含检测目标的图片进行过滤，提高了检测效率和准确率。

4　模型部署

4.1　TensorFlow Serving 服务

在机房的服务器 ubuntu 16.04 上使用 docker 容器技术搭建了 TensorFlow Serving 服务，部署 Inception-V3 缺陷识别模型，远程多个客户端可通过 gRPC API 或 RESTful API 访问模型服务。TensorFlow Serving 支持热部署，方便模型迭代和管理，其结构见图 14，model 1—model 3 为模型迭代更新，model 3 为目前提供服务的模型。实验中通过 gRPC API 调用训练好的模型服务得到检测结果和置信度，与设定阈值比较后决定是否调用 MQTT 客户端向 MQTT 服务器推送报警和控制等消息。图 15 显示了 TensorFlow Serving 提供在线包装缺陷检测服务的日志，包括：请求机器编号、请求时间、识别结果及请求响应耗时等。

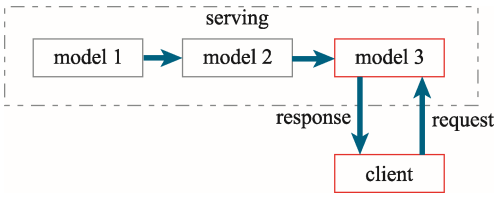


图 14　TensorFlow Serving 结构

Fig.14 Structure of TensorFlow Serving

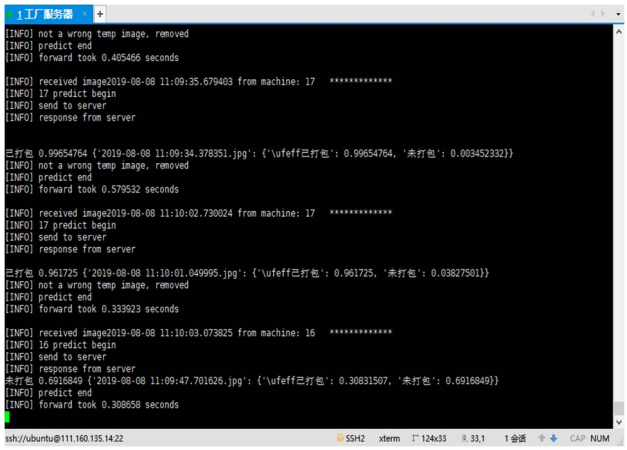


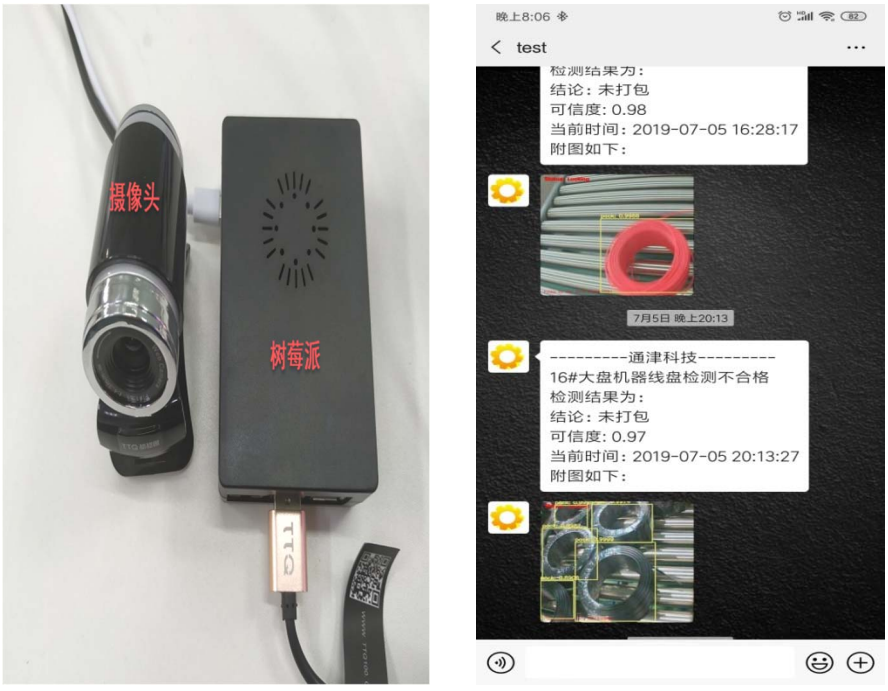
图 15　在线模型服务日志

Fig.15 Service log of the online model

4.2　采集端硬件和信息接收端

图 16a 为工业现场使用的数据采集硬件：USB 摄像头和树莓派 3B。系统检测结果为不合格时，服务程序调用微信 API 接口，将捕捉到的图片推送至微信设备群中的效果，可以看到 16#打盘机的检测结果显示未打包，置信度 0.97，见图 16b。

图 17 中，①和②是生产现场中某 2 台采集终端的实时日志和图像信息，每个窗口左侧显示的是图像采集、预处理、发送到远程服务器的程序执行日志，右侧显示了摄像头拍摄的实时画面；③是服务器端的图像识别服务的执行日志，包含了图片的接收、识别、



a　　b

图 16　图像采集端和信息接收端

Fig.16 Image acquisition end and message receiving end

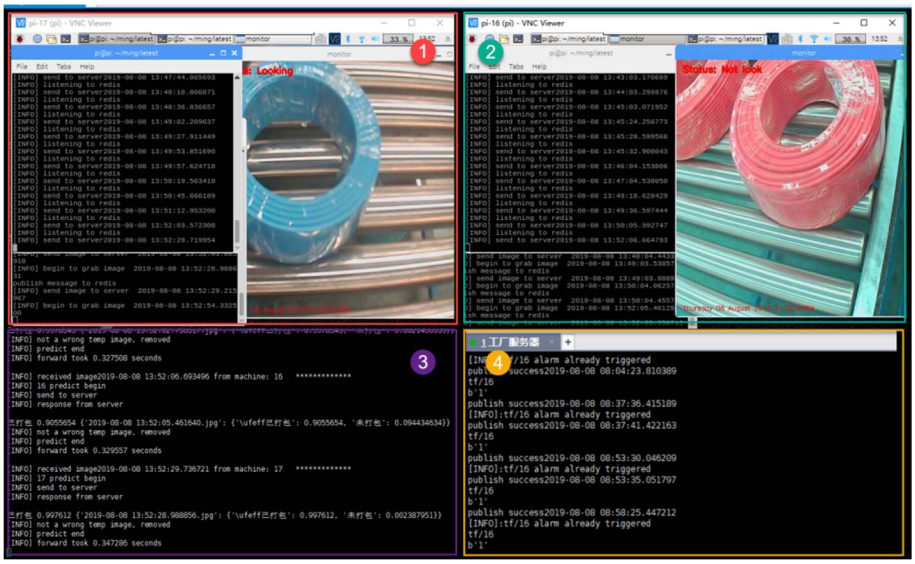


图 17　终端和服务端监测
Fig.17 Monitoring of terminal and server

得到识别结果过程；④为 MQTT 服务程序的实时日志信息，记录了缺陷检测不合格时通过 MQTT 消息触发对应报警装置、PLC 设备的日志信息。

实际生产中将文中提出的检测方法和采集设备应用在打盘机上的情况，见图 18。图 18 中，①是 USB 摄像头，通过延长线连接到了树莓派 3B 上；②是树莓派 3B，负责现场图片采集和预处理，通过车间 WIFI 连接到网络中。

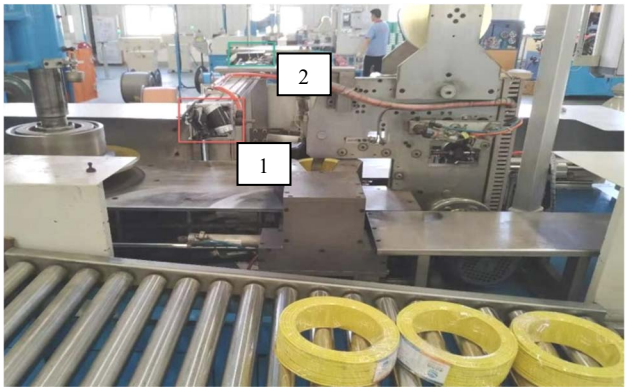


图 18　检测方法在生产环境中的应用
Fig.18 Detection method applied in the production environment

5　结语

文中基于计算机视觉技术，在样本集较小的情况下，提出了 YOLO-V3 和 Inception-V3 两模型相结合的包装缺陷检测方法，提高了多目标和复杂背景下缺陷检测精度，解决了传统传感器技术难以对产品包装质量实时检测的难题，避免了一般包装缺陷检测算法的复杂特征提取过程和模型泛化能力差等问题。该包装缺陷检测方法及其系统设计方案，适用于多客户端并发缺陷检测任务，并且可容易地迁移到类似的在线检测问题上，对企业提高生产效率，提升产品质量具有重要意义。

该研究不足之处是，图片数据网络传输用时较多，从捕获图片到得出检测结果用时 0.2～0.8 s。对实时性要求较高的场合，可考虑简化系统结构，将图片识别和缺陷检测服务部署在本地。在样本数据量足够的情况下，可考虑人工标注增大样本数据，使用 YOLO-V3 模型同时完成目标识别和缺陷检测任务，提升检测速度。迁移学习缺陷识别模型时，应根据缺陷大小、明显程度等情况选择合适的模型，避免“缺陷”在过深的卷积层中消失。

参考文献：

[1] 杨祖彬，代小红. 基于图像配准的食品包装印刷缺陷检测与实现[J]. 计算机科学，2015，42(8)：319—322.
YANG Zu-bin, DAI Xiao-hong. Printing Defects Detection and Realization in Food Packaging Based on Image Registration[J]. Computer Science, 2015, 42(8): 319—322.

[2] 刘学福. 基于机器视觉的药卷多缺陷在线监测[D]. 广州：广东工业大学，2015：17—33.
LIU Xue-fu. On-line Inspection for Cartridge Defects Based on Machine Vision[D]. Guangzhou: Guangdong University of Technology, 2015: 17—33.

[3] 方文星，王野. 一种铝塑泡罩药品包装缺陷检测方法[J]. 包装工程，2019，40(1)：133—139.
FANG Wen-xing, WANG Ye. Defect Detection Method for Drug Packaging with Aluminum Plastic Bubble Cap[J]. Packaging Engineering, 2019, 40(1): 133—139.

[4] 张志豪, 杨文忠, 袁婷婷, 等. 基于 LSTM 神经网络模型的交通事故预测[EB/OL]. (2019-05-14)[2019-07-20]. http://kns.cnki.net/kcms/detail/11.2127.tp.20190511.1229.002.html.
ZHANG Zhi-hao, YANG Wen-zhong, YUAN Ting-ting, et al. Traffic Accident Prediction Based on LSTM Neural Network Model[EB/OL]. (2019-05-14)[2019-07-20]. http://kns.cnki.net/kcms/detail/11. 2127.tp. 20190511.1229.002.html.

[5] 李云鹏, 侯凌燕, 王超. 基于YOLOv3的自动驾驶中运动目标检测[J]. 计算机工程与设计, 2019, 40(4): 1139—1144.
LI Yun-peng, HOU Ling-yan, WANG Chao. Moving Objects Detection in Automatic Driving Based on YOLOv3[J]. Computer Engineering and Design, 2019, 40(4): 1139—1144.

[6] 余绍德. 卷积神经网络和迁移学习在癌症影像分析中的研究[D]. 深圳: 中国科学院深圳先进技术研究院, 2018: 7—21.
YU Shao-de. Convolutional Neural Network and Transfer Learning in Medical Image Analysis[D]. Shenzhen: Shenzhen Institutes of Advanced Technology, University of Chinese Academy of Sciences, 2018: 7—21.

[7] 李琼, 柏正尧, 刘莹芳. 糖尿病性视网膜图像的深度学习分类方法[J]. 中国图像图形学报, 2018, 23(10): 1594—1603.
LI Qiong, BAI Zheng-yao, LIU Ying-fang. Automated Classification of Diabetic Retinal Images by Using Deep Learning Method[J]. Journal of Image and Graphics, 2018, 23(10): 1594—1603.

[8] 赵洪山, 刘辉海. 基于深度学习网络的风电机组主轴承故障检测[J]. 太阳能学报, 2018, 39(3): 588—595.
ZHAO Hong-shan, LIU Hui-hai. Fault Detection of Main Bearing of Wind Turbine Based on Deep Learning Network[J]. Acta Energiae Solaris Sinica, 2018, 39(3): 588—595.

[9] 周福娜, 高育林, 王佳瑜, 等. 基于深度学习的缓变故障早期诊断及寿命预测[J]. 山东大学学报(工学版), 2017, 47(5): 30—37.
ZHOU Fu-na, GAO Yu-lin, WANG Jia-yu, et al. Early Diagnosis and Life Prognosis for Slowly Varying Fault Based on Deep Learning[J]. Journal of Shandong University(Engineering Science), 2017, 47(5): 30—37.

[10] 毛欣翔, 刘志, 任静茹, 等. 基于深度学习的连铸板坯表面缺陷检测系统[J]. 工业控制计算机, 2019, 32(3): 66—68.
MAO Xin-xiang, LIU Zhi, REN Jing-ru, et al. Slab Surface Defect Detection System Based on Deep Learning[J]. Industrial Control Cmputer, 2019, 32(3): 66—68.

[11] 路志英, 任一墨, 孙晓磊, 等. 基于深度学习的短时强降水天气识别[J]. 天津大学学报(自然科学与工程技术版), 2018, 51(2): 111—119.
LU Zhi-ying, REN Yi-mo, SUN Xiao-lei, et al. Recognition of Short-time Heavy Rainfall Based on Deep Learning[J]. Journal of Tianjin University(Science and Technology), 2018, 51(2): 111—119.

[12] 舒悦, 谢传东, 何明, 等. 往复压缩机环状气阀振动信号的 LMD 故障特征提取方法研究[J]. 流体机械, 2019(11): 13—18.
SHU Yue, XIE Chuan-dong, HE Ming, et al. Research on Extraction Method of LMD Fault Features for Vibration Signal of Annular Valve of Reciprocating Compressor[J]. Fluid Machinery, 2019(11): 13—18.

[13] 黄文坚, 唐源. TensorFlow 实战[M]. 北京: 电子工业出版社, 2017: 74—80.
HUANG Wen-jian, TANG Yuan. TensorFlow Practice[M]. Beijing: Publishing House of Electronics Industry, 2017: 74—80.

[14] 高杨, 卫峥. 白话深度学习与 TensorFlow[M]. 北京: 机械工业出版社, 2017: 103—104.
GAO Yang, WEI Zheng. Vernacular In-depth Learning and Tensor Flow[M]. Beijing: China Machine Press, 2017: 103—104.

[15] 李嘉璇. TensorFlow 技术解析与实战[M]. 北京: 人民邮电出版社, 2017: 109—120.
LI Jia-xuan. Technical Analysis and Practice of TensorFlow[M]. Beijing: Posts and Telecom Press, 2017: 109—120.

[16] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-time Object Detection[C]// The IEEE Conference on Computer Vision and Pattern Recognition. Computer Science. Las Vegas, NV, USA: IEEE, 2016: 779—788.

[17] REDMON J, FARHADI A. YOLO9000: Better, Faster, Stronger[C]// Proceedings of the Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA, 2017: 6517—6525.

[18] REDMON J, FARHADI A. YOLOv3: An Incremental Improvement[C]// Proceedings of the Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA, 2018: 1—4.

[19] GIRSHICK R. Fast R-CNN[C]// The IEEE Inter-national Conference on Computer Vision, Santiago, Chile: IEEE, 2015: 1440—1448.

[20] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards Real-time Object Detection with Region Proposal Networks[J]. Pattern Analysis and Machine Intelligence, 2017, 39(60): 1137—1149.

[21] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking The Inception Architecture for Computer Vision[C]// The IEEE Conference on Computer Vision and Pattern Recognition. Computer Science, Las vegas, NV, USA, IEEE, 2016: 2818—2826.

[22] ROSEBROCK A. Deep Learning for Computer Vision with Python Practitioner Bundle[M]. Maryland: PyImage Search, 2017: 57—69.