

机械与过程控制

基于 Faster R-CNN 改进的数粒机系统

胡安翔, 李振华

(山东大学控制科学与工程学院, 济南 250061)

摘要: **目的** 解决目前数粒机只能计数不能同时分拣残损药粒的问题。**方法** 设计以 Faster R-CNN 深度神经网络为核心的药粒数粒机系统。在原有的数粒机基础之上, 更换 CCD 线阵相机为面阵相机, 以满足图像采集的需求, 进一步使用图像分割和多线程技术加快图像处理速度。最终通过训练好的 Faster R-CNN 网络检测出目标并分拣。**结果** 经过测试集的验证, 正常药粒识别率达到了 95.47%, 残损药粒识别率达到了 97.94%, 单幅图像处理达到了 65 ms 的实时速度。**结论** 该方法在传统的计数基础上很好地融合了先进的深度学习技术, 实现了目标的自动分拣。

关键词: 数粒机; 面阵相机; Faster R-CNN; 药粒分拣

中图分类号: TB486; TP391 **文献标识码:** A **文章编号:** 1001-3563(2018)09-0141-05

DOI: 10.19554/j.cnki.1001-3563.2018.09.025

Improved Capsule Counting Machine Based on Faster R-CNN

HU An-xiang, LI Zhen-hua

(School of Control Science and Engineering, Shandong University, Jinan 250061, China)

ABSTRACT: The work aims to solve the problem that the current capsule counting machine can only count capsules and cannot sort damaged capsules at the same time. A capsule counting machine with the Faster R-CNN deep neural network as the core was designed. On the basis of the original capsule counting machine, the CCD line-array camera was replaced by area-array camera to meet the demand of image acquisition, and the image segmentation and multi-thread technology were further used to speed up the image processing speed. Finally, the target was detected and sorted through the well trained Faster R-CNN network. After verification of the test set, the identification rate of normal capsule reached 95.47%, the identification rate of damaged capsule reached 97.94%, and the single image processing reached the real-time speed of 65 ms. The proposed method properly combines the advanced in-depth learning technology based on the traditional counting and realizes the automatic sorting of the target.

KEY WORDS: capsule counting machine; area-array camera; Faster R-CNN; capsule sorting

数粒机在各个行业都有着广泛的应用, 不同行业对数粒机的要求也不尽相同。随着技术的不断进步, 数粒机也从早期简单低速的PLC控制方式^[1], 逐步过渡到基于CCD线阵相机与DSP结合的方式^[2-3]。后者在计数的速度和精度上都得到了很大的提升, 但缺憾是不能同时区分正常药粒和残损药粒, 灌装时混入的残损药粒没法自动剔除, 进而影响了药品包装的质量。为了解决这一问题, 文中尝试引入了当前先进的图像处理技术, 即将基于深度学习的Faster R-CNN^[4]

图像处理技术与数粒机相结合, 增加数粒机自动检测并剔除残损药粒的功能, 极大地提高药品包装的质量。

1 数粒机的工作原理与系统流程

药粒数粒机的功能是完成药粒计数与分拣, 主要由上位机控制系统、图像采集系统、药粒输送系统、灌装系统等组成。其基本的机械结构部分已经趋于稳定, 见图1。

收稿日期: 2017-11-12

作者简介: 胡安翔(1992—), 男, 山东大学硕士生, 主攻机器视觉, 图像处理等。

通信作者: 李振华(1976—), 男, 山东大学副教授、硕导, 主要研究方向为多传感器信息融合、图像处理及模式识别、目标跟踪、以及嵌入式系统应用等。

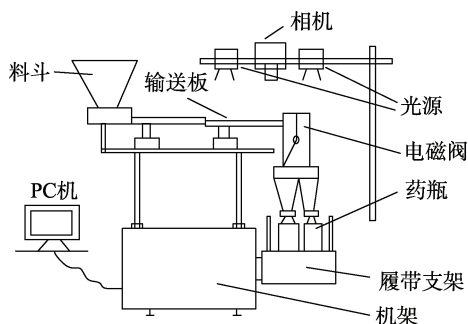


图1 数粒机机械结构

Fig.1 Mechanical structure of the counting machine

该系统改进的重点在于采用的是面阵相机,与采用线阵相机的数粒机所不同的是线阵相机需要安置在药粒下落阶段,容易被粉尘污染镜头,且需要高频背光源。前者的视野只需要拍摄到药粒输送区域,手动调节方便,无需精准对焦,由于相机与光源处于同一方位,对输送区域的要求是需要对输送板进行打磨,避免了反光和重影。分拣和计数在图像处理阶段同步进行,通过训练好的Faster R-CNN神经网络模型从图像中提取目标并完成分类,目标提取后计数工作便可同时进行,其数粒机工作流程见图2。

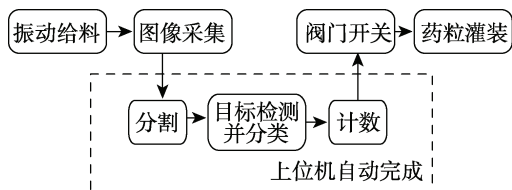


图2 数粒工作流程

Fig.2 Working process of counting machine

2 Faster R-CNN 在药粒分拣中的应用

药粒分拣是数粒机除计数功能以外的重点,如何区分药粒正常与否则是关键。以往在目标提取的方法中,有进行手动设计特征的^[5],难度较高。也有通过给定目标样本进行匹配的方式^[6-7],精度不够。基于深度学习的特征表达具有强大的分层学习能力^[8],特别是卷积神经网络在图像领域具有强大的特征提取能力,使得原来目标提取的方式发生了很大变化。Faster R-CNN是目前比较优秀的深度神经网络,具有良好特征提取和目标分类的功能。Faster R-CNN被借鉴应用于3D目标检测^[9]、图像字幕识别^[10]以及车辆识别^[11-12]等领域。

2.1 Faster R-CNN 原理

在目标检测领域,基于区域的卷积神经网络取得了很大的成功,当忽略花费在区域建议上的时间时,Fast R-CNN基本实现了实时速率^[13]。此时区域建议成为实时目标检测的瓶颈,而在Fast R-CNN的基础上,Faster R-CNN提出了共享卷积特征层的RPN (Region

Proposal Network), 将其特征提取、区域建议和目标分类融合到一个网络当中,使得综合性能得到了较大的提高,特别是在检测速度方面^[4]。

在网络结构上,特征提取方面使用VGG16卷积神经网络层(conv layers)。由于VGG16拓展性很强,结构简洁,在迁移到其他图像数据上的泛化性能非常好,因而被广泛用来提取图像特征。经典的检测方法生成建议框都非常耗时,而Faster R-CNN则相对于传统的EdgeBoxes^[14]和Selective Search^[15]方法,直接使用RPN生成建议框(proposals)。由于共享了卷积网络生成的特征映射(feature maps),因而能极大地提升建议框的生成速度,这也是Faster R-CNN的巨大优势所在。分类器(classifier)部分则利用已经生成的区域建议特征,通过全连接层计算每一个建议框所属的目标类别。基本结构见图3。

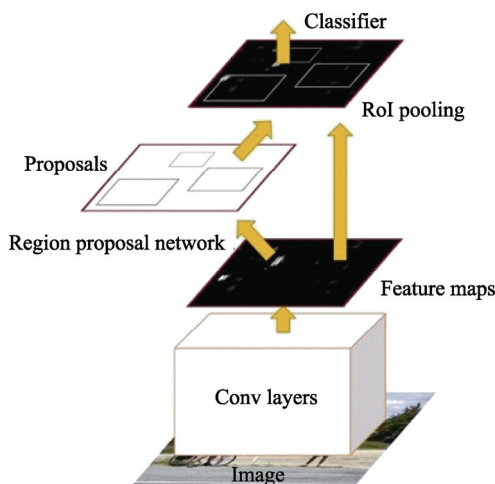


图3 Faster R-CNN 基本结构[4]

Fig.3 Basic structure[4] of Faster R-CNN

对于边框,一般使用四维向量 (x, y, w, h) 分别表示中心坐标、宽和高。RPN在训练时,目的在于寻找一种关系,使得给定锚定 $A = (A_x, A_y, A_w, A_h)$ 与真实目标 $G = (G_x, G_y, G_w, G_h)$ 之间存在一种变换 F ,使得 $F(A_x, A_y, A_w, A_h) = (G'_x, G'_y, G'_w, G'_h)$,其中 $(G'_x, G'_y, G'_w, G'_h) \approx (G_x, G_y, G_w, G_h)$,见图4。

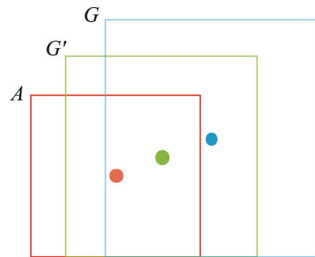


图4 RPN 边框回归

Fig.4 RPN border regression

当 A 与 G 相差较小时,采用的平移量和尺度因子见式(1)。其中 x, x_a, x^* 分别为 A, G', G 的中心坐标的横坐标(同理可知 y, y_a, y^* 等)。

$$\begin{aligned}
t_x &= \frac{(x - x_a)}{w_a}, t_y = \frac{(y - y_a)}{h_a}, \\
t_w &= \log_2 \left(\frac{w}{w_a} \right), t_h = \log_2 \left(\frac{h}{h_a} \right), \\
t_x^* &= \frac{(x^* - x_a)}{w_a}, t_y^* = \frac{(y^* - y_a)}{h_a}, \\
t_w^* &= \log_2 \left(\frac{w^*}{w_a} \right), t_h^* = \log_2 \left(\frac{h^*}{h_a} \right),
\end{aligned} \quad (1)$$

分类器通过全连接层实现最终的目标分类,其基本结构见图5。

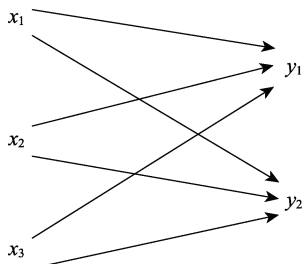


图5 全连接层基本结构

Fig.5 Basic structure of full connection layer

计算公式见式(2),由于权重 W 和偏置 b 是固定的,因而使用了图3所示的RoI pooling层将所有区域建议框映射到同样大小。

$$\begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \\ w_{31} & w_{32} \end{pmatrix} + \begin{pmatrix} b_1 & b_2 \end{pmatrix} = \begin{pmatrix} y_1 & y_2 \end{pmatrix} \quad (2)$$

神经网络训练最重要的是围绕最小化代价函数来修改权值,整个网络使用的代价函数见式(3)。其中下标 i 为建议框的索引; p_i 为目标预测概率; p_i^* 为相对应的真实目标概率; t_i 为预测的边框; t_i^* 为相对应的真实边框; N_{cls} 为最小批量的大小,取值256; N_{reg} 为建议框的数量,取值2400,为平衡两者的权重,参数 γ 取值10。 L_{cls} 和 L_{reg} 分别采用的是神经网络中常用的softmax函数和smooth L1函数。

$$\begin{aligned}
L(\{p_i\}, \{t_i\}) &= \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \\
&\quad \gamma \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*)
\end{aligned} \quad (3)$$

2.2 检测与计数相结合

以往需要使用光电开关等,在目标经过时给出触发信号,经过PLC等控制器控制电磁阀进行计数^[1]。该系统设计了一个简单有效的计数方式,见图6。相机在进行每一帧的图像采集后,都会经过目标检测,对所有检测到的目标位置以及当前采样时间都会进行保存。由于输送板是划分成固定几个通道的,因此每个通道检测到的目标之间相互不受影响。由于药粒

在每个通道中是单向移动的,所以在采样时,目标都会向同一方向移动一定的距离,在对比前后帧时,只需对比基准线药粒移动反方向的目标(即图6中基准线左边的目标)位置。在采样时,目标与基准线之间的距离会逐渐缩小,所以当某一采样时刻,若离基准线最近的目标与基准线之间的距离出现增大,说明此时离基准线最近的目标是前一帧中第二近的目标,原先最近的目标在此刻跨越了基准线,称此刻为目标经过基准线。每当目标经过基准线A时,会被依次记录它们的时间戳和位置信息,根据药粒在通道中单向移动的特性,当目标经过基准线B时,同样会再被依次记录,此刻在记录的同时并根据FIFO(First In First Out)原则提取之前保存的目标信息,对比2次记录的目标位置和时间信息可以估算出目标移动速度。由于基准线B到电磁阀的距离是固定的,结合自由落体公式可以预测目标下落所需时间。经实验检测,当相机拍摄速率超过12帧/s时,基本满足需求。

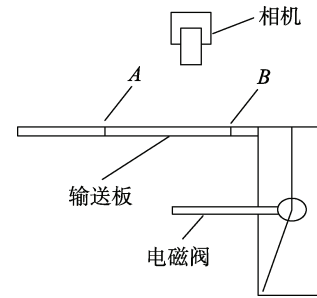


图6 计数示意

Fig.6 Schematic diagram of counting

2.3 实验测试与数据

将采集到的图像分割成约为375×200的尺寸,目的是减少单次迭代的训练时间。丢弃掉没有目标以及目标不清晰的图像完成数据清洗,在剩余的图像中用矩形框标注目标,并记录目标类别以及目标在图像中的实际坐标。在完成以上的数据集制作过程之后,就可以使用Faster R-CNN网络进行训练。由于该网络参数量庞大,直接用于大尺寸的图像,会导致检测速度下降,因此该系统采用了图像区域分割结合多线程技术,解决了这一问题。示例采用的是分割后的图像见图7。含有正常与缺陷目标的检测效果见图7a—b。

根据目标需求,对测试数据集进行了实验,保存了测试目标类别以及实际坐标值和测试坐标值,坐标采用 $(x_{min}, y_{min}, x_{max}, y_{max})$ 形式,分别表示目标矩形框左上角横坐标值,左上角纵坐标值,右下角横坐标值,右下角纵坐标值。测试结果见表1—2。根据结果进行分析后发现,部分缺陷不明显的药粒被标注为残缺,但测试时会被归于正常,较小的残缺目标出现漏检等,都会导致识别失败,检测精度出现下降,因而在后期系统升级的过程中仍有改进空间。

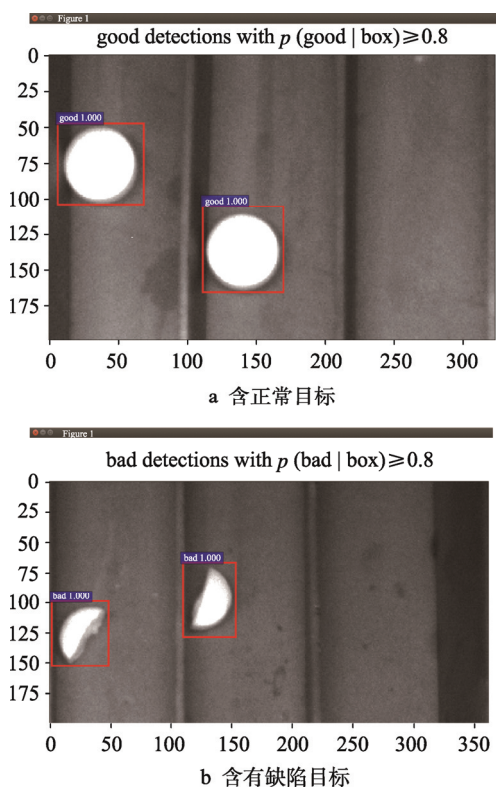


图7 检测结果
Fig.7 Testing result

3 创新点与存在的问题

3.1 创新点

利用面阵相机和基于深度学习的图像处理技术,将原本只能用来计数的数粒机变成了既可以计数,又可以区分正常药粒和残缺药粒的多功能数粒机。同样的技术可以用于其他种类的药粒乃至其它目标物体的计数和分拣上。

结合图像分割和多线程处理的技术,解决了引入Faster R-CNN神经网络后,图像处理速度变慢的问题。保证了计数和分拣的实时性。

3.2 存在的问题

由于引入了深度学习技术,前期需要人工标注大量的高质量训练样本,这一点上比较费时费力。样本训练对设备要求较高,训练时间较长,该系统是在酷睿 i7-7700 型 CPU+ 英伟达 1080TI 型 GPU, 结合 Tensorflow 深度学习框架上进行的,单次训练长达4 h。后期努力探索的方向是尽可能做到在线训练,减少训练时间成本。

为了减少设备成本投入采用了单摄像头,结果是每个目标所实际分配到的像素数较少,在目标缺陷不明显的情況下会有误分类发生。

表1 测试结果对比

Tab.1 Comparison of test results

实验序号	实际类别	实际坐标	测试类别	测试坐标	识别结果
1	正常	(115, 82, 175, 143)	正常	(114.0, 80.5, 173.5, 142.3)	成功
2	正常	(5, 86, 66, 148)	正常	(5.9, 87.6, 66.6, 150.8)	成功
3	正常	(110, 70, 170, 131)	正常	(111.0, 74.1, 168.7, 132.7)	成功
4	正常	(220, 91, 281, 152)	正常	(223.1, 92.8, 284.8, 153.5)	成功
5	正常	(108, 37, 169, 98)	正常	(109.8, 39.3, 170.6, 102.3)	成功
6	正常	(4, 37, 65, 97)	正常	(7.7, 36.2, 66.9, 101.7)	成功
...	...	...	...	...	...
7852	缺陷	(4, 23, 46, 77)	缺陷	(5.5, 23.9, 49.1, 79.4)	成功
7853	缺陷	(61, 25, 122, 84)	缺陷	(64.1, 25.8, 121.8, 88.3)	成功
7854	缺陷	(65, 18, 110, 75)	缺陷	(66.3, 20.2, 112.7, 77.6)	成功
7855	缺陷	(161, 58, 195, 119)	缺陷	(163.0, 57.9, 199.8, 122.5)	成功
7856	缺陷	(69, 41, 117, 93)	缺陷	(68.8, 42.7, 120.8, 97.8)	成功
7857	缺陷	(115, 41, 163, 92)	缺陷	(115.4, 41.4, 165.7, 94.5)	成功

表2 汇总准确率

Tab.2 Summary of accuracy

测试数据组	识别成功数量	识别失败数量	准确率/%
正常	3137	149	95.47
缺陷	4477	94	97.94

4 结语

以面阵相机和Faster R-CNN深度神经网络为核

心,改进了现有的药粒数粒机系统。该系统的机械设备部分的主体采用的都是趋于稳定的结构,包括三级振动输送模块、全自动灌装流水线模块、支架模块等。改进后的系统通过图像采集模块将图像输送到上位机部分进行处理,获取药粒的数量和缺陷信息,通过控制板卡对灌装系统进行控制,整个过程共同构建了药粒数粒机的检测和计数功能。操作简单,不需要复杂的专业技能,设备成本控制得较好,不需要额外设计和采购嵌入式上位机设备,只需要利用已有的个人

台式机电脑安装该系统的相应软件即可。此外采用Faster R-CNN深度神经网络系统,可智能地控制药粒的检测和计数,保证了产品生产的质量和速度。

参考文献:

- [1] 李东. 基于 PLC 的自动数粒机系统设计[J]. 电子制作, 2010, 4(4): 26—27.
LI Dong. The Design of Automatic Granulator System Based on PLC[J]. Hands on Projects, 2010, 4(4): 26—27.
- [2] 齐建虹, 蔡锦达, 李祥伟. 基于 CCD 和小波变换的数粒机控制系统设计[J]. 包装工程, 2014, 35(13): 116—122.
QI Jian-hong, CAI Jin-da, LI Xiang-wei. Design of Control System for Counting-machine Based on CCD and Wavelet Transform[J]. Packaging Engineering, 2014, 35(13): 116—122.
- [3] 蔡锦达, 唐静, 齐建虹, 等. 基于 DSP 和自组织竞争神经网络的数粒机控制系统[J]. 控制工程, 2017, 24(3): 487—494.
CAI Jin-da, TANG Jing, QI Jian-hong, et al. The Control System of Counting-machine Based on DSP and Self-organizing Competitive Neural Network[J]. Control Engineering of China, 2017, 24(3): 487—494.
- [4] REN S, HE K, 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.
- [5] 于来行, 冯林, 张晶, 等. 自适应融合目标和背景的图像特征提取方法[J]. 计算机辅助设计与图形学学报, 2016, 28(8): 1250—1259.
YU Lai-hang, FENG Lin, ZHANG Jing, et al. An Image Feature Extraction Method Based on Adaptive Fusion of Object and Background[J]. Journal of Computer-Aided Design & Computer Graphics, 2016, 28(8): 1250—1259.
- [6] 朱奇光, 张朋珍, 李昊立, 等. 基于全局和局部特征融合的图像匹配算法研究[J]. 仪器仪表学报, 2016, 37(1): 170—176.
ZHU Qi-guang, ZHANG Peng-zhen, LI Hao-li, et al. Investigation on the Image Matching Algorithm Based on Global and Local Feature Fusion[J]. Chinese Journal of Scientific Instrument, 2016, 37(1): 170—176.
- [7] 向程谕, 王冬丽, 李建勋, 等. 基于改进 SIFT 特征的深度图像匹配[J]. 计算机应用, 2016, 36(S2): 135—138.
XIANG Cheng-yu, WANG Dong-li, LI Jian-xun, et al. Depth Image Matching Based on Improved SIFT Feature[J]. Journal of Computer Applications, 2016, 36(S2): 135—138.
- [8] 尹宏鹏, 陈波, 柴毅, 等. 基于视觉的目标检测与跟踪综述[J]. 自动化学报, 2016, 42(10): 1466—1489.
YIN Hong-peng, CHEN Bo, CHAI Yi, et al. Vision-based Object Detection and Tracking: A Review [J]. Acta Automatica Sinica, 2016, 42(10): 1466—1489.
- [9] SONG S, XIAO J. Deep Sliding Shapes for Amodal 3D Object Detection in RGB-D Images[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition, 2016: 808—816.
- [10] JUSTIN JOHNSON N A K N. Dense Cap: Fully Convolutional Localization Networks for Dense Captioning[C]// IEEE, 2016, 4565—4574.
- [11] 桑军, 郭沛, 项志立, 等. Faster-RCNN 的车型识别分析[J]. 重庆大学学报, 2017, 40(7): 32—36.
SANG Jun, GUO Pei, XIANG Zhi-li, et al. Vehicle Detection Based on Faster-RCNN[J]. Journal of Chongqing University, 2017, 40(7): 32—36.
- [12] 宋焕生, 张向清, 郑宝峰, 等. 基于深度学习方法的复杂场景下车辆目标检测[J]. 计算机应用研究, 2018, 35(4): 1270—1273.
SONG Huan-sheng, ZHANG Xiang-qing, ZHENG Bao-feng, et al. Vehicle Detection Based on Deep Learning in Complex Scene[J]. Application Research of Computers, 2018, 35(4): 1270—1273.
- [13] GIRSHICK R. Fast R-CNN[J]. Computer Science, 2015: 1440—1448.
- [14] ZITNICK C L, DOLLÁR P. Edge Boxes: Locating Object Proposals from Edges[C]// Computer Vision-ECCV 2014: 13th European Conference, Zurich, Switzerland, Proceedings, Springer International Publishing, 2014: 391—405.
- [15] UIJLINGS J R R, SANDE K E A, GEVERS T, et al. Selective Search for Object Recognition[J]. International Journal of Computer Vision, 2013, 104(2): 154—171.