

# YOLO 模型结合图像处理方法 实现铝棒准确识别与计数

毕晓琳, 侯卓轩, 张瀛天

(广东东软学院, 广东 佛山 528225)

**摘要** 传统的形态学的棒材图像识别计数方法无法在复杂紊乱的环境下对目标样品进行准确识别计数, 需要严格的拍摄环境和拍摄角度。对拍摄条件要求较低的卷积神经网络图像识别算法通常运算量大、训练时间长、难以自适应更新模型。为了解决棒材识别计数以上痛点问题, 文章以铝棒计数为研究背景, 提出将铝棒端面图像通过系列图像处理方法压缩图像样本的信息量, 再把处理后图像输入 YOLO 模型进行训练、测试。结果表明, 模型的训练时长相比不处理图像、直接训练 YOLO 模型缩短了 84.84%, 证明了模型自适应更新的可行性, 并且通过图像处理方法, 将有效信息进行提取后再训练 YOLO 模型, 识别准确率相比直接使用铝棒端面图像训练 YOLO 模型提高了 14%, 验证了该方法在棒材计数应用中的优势和可行性。

**关键词** 铝棒计数; 图像处理; 卷积神经网络图像识别; YOLO

**中图分类号**: TP317.4

**文献标识码**: A

**文章编号**: 2097-3365(2023)12-0016-03

## 1 前言

铝型材产品被普遍应用到人们实际生活的许多环节, 比如铝合金门窗、灯饰等, 与人们的生活联系密切。铝棒是铝型材成品的加工的原料, 目前, 多数铝材厂商采用吨作为包装单位的方式按重量进行铝棒原料的采购, 但是实际生产中常常要了解掌握的是具体的铝棒条数, 因而要求铝材厂商对采购的铝棒实现更快速准确的复核和验收计数。目前大多数铝材厂对铝棒的数量清点工作主要依靠人工清点完成。这种方法不仅造成人力物力的浪费, 而且计数的准确率通常不高, 计数结果难以复核。

随着计算机科学与技术的发展日益成熟, 铝材厂商都倾向于使用更加智能的方式来代替一些重复、乏味的人力工作。20世纪70年代便有棒材自动识别计数的实践。第一种方法是使用机械装置来分离棒材, 然后在机械装置的基础上利用光电管实现自动计数<sup>[2]</sup>, 第二种方法是将数字图像技术运用到棒材图像识别和处理中, 比如采用模板匹配的方法进行计数<sup>[3]</sup>, 或连通区域匹配实现棒材计数的方法<sup>[4]</sup>。尽管使用图像处理方法进行计数的相关技术已经比较成熟, 但是该类实践对图像拍摄背景、图像采集设备有较高的要求。在上述所提的研究中, 棒材都是垂直放置, 且拍摄背景是纯色, 非常容易将背景与棒材分离, 但是铝棒在实

际的加工生产过程中, 由未经培训非专业的工人拍摄到的铝棒图片大概率不是垂直放置的, 且拍摄背景杂乱, 因此仅依靠一般的数字图像处理方法进行计数是不符合实际的。

结合卷积神经网络图像识别算法的优点, 能够在复杂环境中准确识别特征图像。YOLO 是由 Redmon et al.<sup>[5]</sup> 于 2016 年提出的一种只完整查看一次图像的识别模型, YOLO 与其他卷积神经网络图像识别算法相比, YOLO 模型将对象检测视为简单的回归问题, 具有快速、准确的优点<sup>[6]</sup>。本研究使用 YOLO 算法作为图像识别模型, 并在使用模型识别之前, 采用系列图像处理算法压缩输入模型的信息量, 提升训练和识别速度。

## 2 实验样本

### 2.1 样本采集

在本研究中, 拍摄的铝棒端面样本图像都是来自铝材生产车间, 在自然灯光状态下拍摄, 样本图像背景中有设备、墙体、屋顶等干扰因素。

用手机拍摄共图像 1309 张, 其中我们将样本又分为训练集、验证集以及测试集。训练集是用于模型拟合的数据样本, 验证集是用于调整模型的超参数和用于对模型的能力进行初步评估。通常用来在模型迭代训练时, 用以验证当前模型泛化能力, 以决定是否停

表 1 模型和 M+Y 模型训练时长对比

序号	图片数量			训练时间 (s)		时间比值
	训练集	验证集	总数	t (Y)	t (M+Y)	t (M+Y) / t (Y)
1	453	128	581	26026	4329	16.63%
2	349	101	450	19393	2723	14.04%
3	141	37	178	9414	1395	14.82%

平均时间比值为 15.16%

止继续训练。测试集是用来评估最终模型的泛化能力。

## 2.2 图像标注

使用 YOLOv5 框架进行目标检测必须先对图片进行标注,生成对应的 txt 文档。本研究样本标注使用在线图像标注软件 Roboflow (<https://app.roboflow.com/>)。用 Roboflow 生成 YOLOv5 能够读取的标注信息。

由于识别目标为单独对铝棒检测和计数,图片中没有其他与铝棒特征相似的对象,因此我们只标注 1 个类,名为“Aluminum bar”。图片总共 1309 张,被随机化分为 3 组:943 张为训练集,266 张为验证集,100 张为测试集。所有的铝棒都进行了标注,总的标注 6492 个。图片标注好之后进行导出,导出时将图片尺寸调整为 640\*640 像素,以适应 YOLOv5 模型。

## 2.3 图像处理

标注工作完成后,将铝棒端面彩色图片进行灰度化处理<sup>[7]</sup>,去除掉图片原本的色彩信息,简化了输入训练模型的信息量,以达到提高训练效率的结果。

在真实生产环境中,拍摄到的铝棒端面图像背景相对复杂,铝棒端面相较复杂背景亮度更高,因此为了使铝棒端面在背景中更加突出,对铝棒端面灰度图作对比度增强<sup>[8]</sup>的处理,使得铝棒端面和背景亮度差异更大。为避免复杂背景产生噪点影响模型准确率,将对比度增强后的图片进行了高斯滤波<sup>[9]</sup>去噪。

为了进一步缩减图像信息量,仅将铝棒端面的形态信息输入图像识别模型进行训练,我们将去噪后的图片采用 Sobel 算子边缘提取算法<sup>[10]</sup>获取铝棒形貌信息,至此完成模型训练前的图像处理工作。

## 3 图像识别模型与分析

### 3.1 YOLO 模型结构

基于 YOLO 模型在图像识别领域的有着模型简单、识别准确的优点,综合 YOLO 几种模型的优缺点,本研

究采用 YOLOv5 模型作为图像识别模型。YOLOv5 的总体架构由四部分构成:输入端、Backbone、Neck 和 Head。

Backbone 部分作为特征提取网络,主要是由 Focus 和 BottleneckCSP 组成。

Focus 在减少计算量的同时实现下采样过程,BottleneckCSP 是此特征提取网络的核心。

Neck 部分采用了 FPN 与 Pan 相结合的结构。Head 部分实现输出的功能,包括检测的 Probability、Score 以及 Bounding-box。

### 3.2 模型训练与结果分析

在本研究的实验中使用的训练权重文件为 YOLOv5m,使用 Google Colab 提供的云 GPU 进行训练。共训练了三组铝棒图片进行测试,581 张图片需要约 7 小时 13 分,450 张图片需要 5 小时 23 分,178 张图片需要 2 小时 36 分。

为了验证使用图像处理方法能够减少模型训练信息,提高模型训练速率,本研究做了两组对比实验。一组实验是直接将标注过的图片与和标注文件输入 YOLOv5 模型进行训练,以下简称 Y 模型。另一组实验则是在模型对图片进行训练前,对所有图片进行上述图像处理,再将图片和标注信息输入 YOLOv5 模型进行训练以下简称 M+Y 模型。

对于 M+Y 模型来说,仅将铝棒形态信息输入模型,大幅压缩了图像样本信息量,因此训练时间得到大幅缩短。

本研究中,为排除偶然性的结果,采取不同数量的训练集和验证集图像,输入 Y 模型和 M+Y 模型进行训练,总共训练 3 次 YOLOv5 模型,综合比较训练时长。每次训练模型输入的训练集、验证集图像数量,以及 Y 模型与 M+Y 模型的训练时间对比如表 1 所示。结果表明,不论训练集和验证集图像数量多少,采用图像处理方法先提取有效信息再训练模型,都能大幅提升训练速度。M+Y 模型的平均训练时间仅为 Y 模型的 15.16%,

平均训练时长大约缩短了85%。

此外,在本研究中,使用准确率作为模型评估的标准,其中准确率分为单张图片(简称单图)的准确率和总体准确率。单图准确率定义为:

$$AI_i = \frac{AN_i - |AN_i - DN_i|}{AN_i} \times 100.0\%$$

其中,  $i$  为本张图的序号, AI (Accuracy of one Image) 为单图准确率, AN (Actual Number of aluminum bars) 为单图中实际铝棒数, DN (Detected Number of aluminum bars) 为单图中检测出的铝棒数量。  $n$  张铝棒识别图的总体准确率定义为:

$$A(n) = \frac{\sum_{i=1}^{100} [AI_i]}{n} \times 100\%$$

其中,  $A(n)$  为  $n$  张图的总体准确率, AI (Accuracy of one image) 为单图准确率。

运用 Y 模型与 M+Y 模型通过 YOLOv5 训练, 分别得到两个权重文件, 即 Y.pt 与 MY.pt。分别使用这两个权重文件对 100 张铝棒图进行识别与计数, 其中 100 张铝棒图包含 54 张来自测试集与 46 张来自验证集的图像。将识别计数结果统计整理, Y 模型与 M+Y 模型的总体准确率如表 2 所示。若不采用图像处理方法, 直接将铝棒彩图输入 YOLOv5 得到的 Y 模型的总体准确率为 70%, 而采用图像处理方法得到的 M+Y 模型的总体准确率为 84%。

表 2 Y 模型与 M+Y 模型的总体准确率

	Y 模型	M+Y 模型
A(100)	70%	84%

#### 4 结论

本研究表明, 在对铝棒端面图像进行铝棒识别计数的实践中, 先采用图像处理方法提取有效信息, 压缩图片信息, 大幅减少输入模型训练的信息量, 能够有效地提高训练速率, 并且由于排除了无关信息造成的误差, 识别计数的准确率也得到了提升。在本研究中, 压缩铝棒图像信息量后, 训练时长仅为未压缩前的 15.16%, 训练速度的大幅提升为实现模型自适应提供了实践基础。在真实铝材加工工厂的生产流程中, 各生产线上有大量工人会在不同流程阶段在一天内多次使用该铝棒识别计数软件, 一天累计新拍摄的铝棒端面图像有可能上千张, 因此为了模型的优化, 模型

需要自适应地将新拍摄的图像作为补充训练样本, 使模型能够“认识”更多各样格式排列的铝棒, 从而提高模型的准确率, 而在真实生产应用中, 如果要频繁地自适应重新训练模型, 且不耽误软件的日常使用, 则对模型训练速度有极高要求。因此, 本研究提出的先提取图像有效信息再进行模型训练的识别计数方法, 不仅使训练速度得到大幅提升, 而且能提高识别计数准确率, 为棒材识别计数以及模型自适应发展提供了新的研究方向。

#### 参考文献:

- [1] Sadao K, Satoru U, Tomoji O. AUTOMATIC SEPARATING DEVICE FOR BAR STEEL CONTINUOUSLY TRANSPORTED BY CHAIN CONVEYER. JP19810011139[P]. JPS57126315A[2023-10-19].
- [2] 罗祯伟. 棒材在线自动计数分钢系统应用[J]. 南方金属, 2021(04):56-58.
- [3] 刘娜娜, 孙文赞. 基于纹理特征和模板匹配的棒材自适应计数方法[J]. 信息通信, 2011(06):25-26.
- [4] XING YAN, XIANAIQIAO CHEN. Research on the Counting Algorithm of Bundled Steel Bars Based on the Features Matching of Connected Regions[C]//2018 3rd International Conference on Image, Vision and Computing: ICIVC 2018, Chongqing, China, 27-29 June 2018.:Institute of Electrical and Electronics Engineers, 2018.
- [5] JOSEPH REDMON, SANTOSH DIVVALA, ROSS GIRSHICK, et al. You Only Look Once: Unified, Real-Time Object Detection[C]//29th IEEE Conference on Computer Vision and Pattern Recognition: 29th IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 26 June -1 July 2016, Las Vegas, Nevada.:Institute of Electrical and Electronics Engineers, 2016.
- [6] 王琳毅, 白静, 李文静, 等. YOLO 系列目标检测算法研究进展[J]. 计算机工程与应用, 2023, 59(14):15-29.
- [7] 凌双明. 智能视觉机器人图像灰度化处理方法研究[J]. 山西电子技术, 2023(02):88-91.
- [8] ERIC WHARTON, SOS AGALIAN, KAREN PANETTA, et al. A Logarithmic Measure of Image Enhancement[C]//Mobile multimedia/image processing for military and security applications, 2006.
- [9] MAHMOUD SAEIDI, MOHAMMAD HESAM TA DAYON, REZA HASANZADEH. IMPROVING ATTENUATING NOISE IN IMAGE SEQUENCES USING AVERAGING AS THE CENTRAL VALUE OF FILTER[C]//The 11th IASTED International Conference on Artificial Intelligence and Soft Computing, 2007.
- [10] 沈德海, 侯建, 鄂旭. 基于改进的 Sobel 算子边缘检测算法[J]. 计算机技术与发展, 2013(11):22-25.