

DOI: 10.13955/j.yzyj.2021.03.04.05

# 目标检测算法在集散中心的使用研究

胡世文

(中国邮政集团有限公司江苏省信息技术局, 江苏 南京 210019)

**摘 要:** 针对物流企业集散中心存在的业务痛点, 对问题进行抽象化并提出其在人工智能环境下的解决范畴。通过阐述目标检测算法的使用过程、测试条件以及算法在对痛点问题模拟测试中的效果, 为集散中心作业流程提供了优化方案。

**关键词:** 目标检测算法; 集散中心; 业务场景; 人工智能

**中图分类号:** F61      **文献标识码:** A

集散中心的标准化处理流程属于资源密集型操作, 传统的人工管理面临着工作量大、监管复杂和开销大的问题。利用人工智能技术解决作业流程中的共性问题对于集散中心意义重大。集散中心作业场景涉及生产活动多个流程, 任何一个流程出现问题都会影响整个流程。物流企业集散中心的工作流程主要分为卸车、分拣、装车三个部分, 其中分拣环节可以进一步细化为小件集包、大件路向选择等环节, 所有环节通过皮带机进行串联; 车辆、人员、皮带机三者有效协调, 才能确保整个作业流程的顺畅。

## 1 集散中心的业务痛点

### 1.1 垛口车辆检测

车辆的调度、监控目前由人工完成, 由于物流场地较大, 通过人工查看垛口状态来进行车辆调度, 工作量巨大, 会造成调度的滞后, 这种情况在业务高峰期更加突出。

### 1.2 卸车皮带机处人数统计

卸车皮带机处人数监控主要用于监控工作人员是否按时到岗、人数是否符合规定。

### 1.3 供件区人数统计

与卸车皮带机处人数统计的目标类似, 用于对供件区的工作人员进行监控。

### 1.4 皮带机堵点检测

皮带机的运转影响整个分拣流程, 皮带机上的货物堵塞不仅影响分拣的速度, 而且严重时还会加速皮带机的机械损耗, 因此快速定位和皮带机堵点问题亟需解决。

### 1.5 回笼口拥堵检测

回笼口位于大件皮带机的末端, 用于暂存没有及时分到路向的大件物品。回笼口的拥堵主要出现在业务高峰期, 会造成大件皮带机停止, 从而影响整个流程; 快速发现堵塞情况并通知相关人员进行处理是回笼口拥堵检测的目标。

上述五个场景可以归集为对人、物和状态的检测; 从技术实现角度来说, 人工智能检测算法需要被应用到人、物和状态的识别中, 以满足上述功能要求。在上述场景中, 对于人的识别, 检测算法只需要识别是否是人, 不需要关心其性别、动作等; 在物的识别上, 算法只需要识别是否是特定的物; 在状态的识别上, 算法只需要识别目前的状态是否

**作者简介:** 胡世文 (1983 ~), 男, 江苏南京人, 硕士, 工程师, 主要从事新技术与软件开发研究。

**收稿日期:** 2020-12-21

是预设的状态之一，不需要对状态进行量化。因此，人、物、状态的识别在上述场景中可以统一归类到同一类目标检测算法：基于特征，判断“是”与“否”。目标检测算法的引入是为了减少上述检测中人工检查的工作量，对符合目标特征的场景报警，及时发现潜在危险。

## 2 目标检测算法

近年提出的 SSD 目标检测算法因其性能优势被广泛使用。SSD 算法基于前向反馈卷积网络，识别过程比较简单。SSD 算法通过此网络产生规模固定的集合，包含了图像中的各目标边界框和目标评价值；对于每个预测框，首先要根据类别置信度确定其类别（置信度最大者）与置信度值，并过滤掉属于背景的预测框，然后根据置信度阈值（如 0.6）过滤掉阈值较低的预测框；对于留下的预测框进行解码，根据先验框得到其真实的位置参数；解码后，一般需要根据置信度进行降序排列，仅保留前 k（如 200）个预测框；最后使用非最大值抑制法算法，过滤掉那些重叠度较大的预测框，最后剩余的预测框就是检测结果（见图 1）。它的核心理念有三个方面。

### 2.1 采用多尺寸特征图用于检测

SSD 网络分为两个部分：基础层和卷积层。基础层是 VGG-16 的前 4 层网络；卷积层是由特征图逐渐变小的简单卷积网络组成。使用不同大小特征图的好处是可以适应不同大小的目标检测，大的特征图用来检测比较小的目标，小的特征图用来

检测大的目标。

### 2.2 使用卷积运算

在图 1 中，卷积层中每层都有  $3 \times 3$  的卷积运算进行预测运算，运算结果可以是某一类评估值，或是相对于先验框的偏移量。

### 2.3 设置默认框和多纵横比

在特征图的每个位置预测 k 个不同纵横比的框，每个框的预测结果包括 c 个分类得分和 4 个相对于默认框的位置偏移量，那么在  $m \times n$  的特征图中，就可以得到  $(c+4) \times k \times m \times n$  个预测值，如图 2 所示。

## 3 方案选择

### 3.1 计算模式选择

计算模式是指算法计算运行的位置。目前存在两种计算模式：边缘计算和集中计算。

边缘计算的特点是其计算量分布在各前端的数据收集节点，无需发送大量的原始数据到后台，优点是网络带宽消耗小、后台服务器压力小，并且反应快，缺点是软件版本更新难、前端的硬件代价较高。

集中计算的特点是所有的计算发生在后台服务器上，优点是软件版本可以根据计算的要求和能力做到及时更新，缺点是网络带宽消耗大、后台服务器的计算压力大，与边缘计算模式相比，对监控事件的反应较慢。

结合目前安防产品的现状，上述两种计算模式衍生出三类针对智能监控的解决方案。

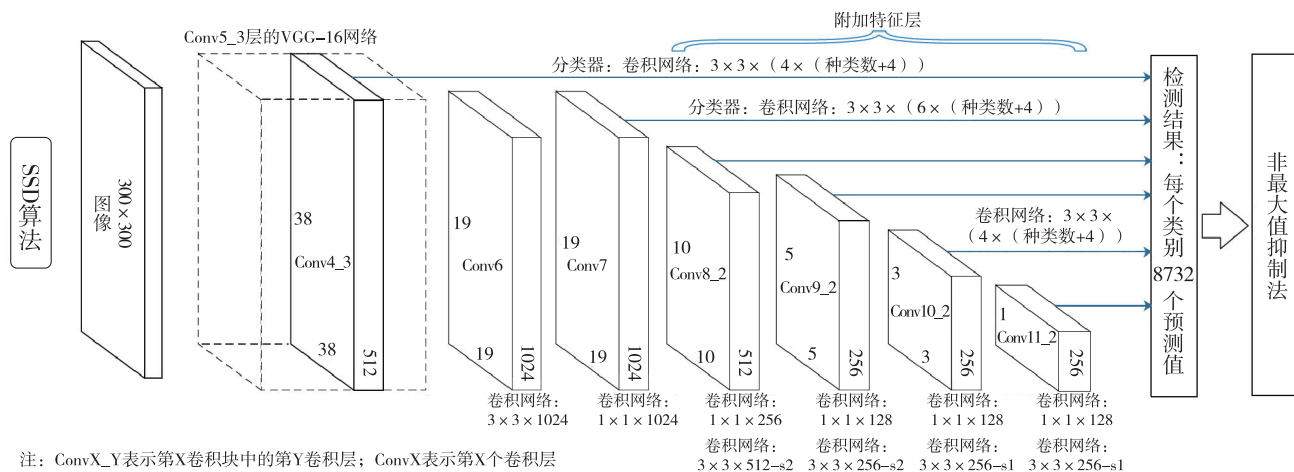


图 1 SSD 网络结构图

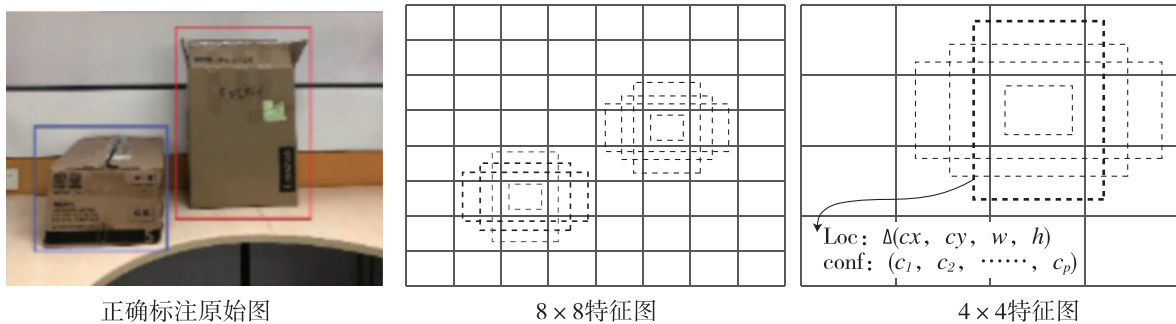


图2 SSD特征示意图

纯硬件方案：主要由安防企业提供的智能摄像头、NVR以及视频分析服务器组成，其算法主要由第三方提供，安防企业提供智能硬件的接口，购买用户可以灌输自己的算法，算法只需要符合安防企业的开发规范即可。

纯软件方案：此类产品主要以近几年出现的人工智能算法公司为代表，例如云从、旷视、百度等，这些公司虽然也有零星的智能硬件，但主要以销售算法为主。

软硬件结合方案：此类产品以传统的安防企业为代表，人工智能识别算法在硬件产品出厂时已经被灌输到硬件设备上。

纯硬件方案和软硬件结合方案从计算模式判断，属于边缘计算模式；纯软件方案偏向集中计算模式，其识别算法部署在后台服务器上，对采集到的数据进行集中处理。

在物流中心监控场景中，前端数据收集工具是已经部署好的摄像头；从调研数据来看，目前部署的摄像头大部分是普通摄像头，小部分区域部署了高清摄像头。近年来安防企业推出的已安装人工智能算法的智能摄像头和可配置自定义算法的“软件定义硬件”型摄像头没有在物流企业进行部署。对于一个大型作业中心，更新基础监控设备的时间和经济代价是必须考虑的因素；智能摄像头的单价结合目前主流物流中心的规模，其总代价较高，因此上述纯硬件方案和软硬件结合方案，并不适合目前物流企业针对集散中心场地的智能改造需求。

纯软件方案从服务器硬件成本、算法软件成本两个方面进行评估。服务器硬件成本主要是GPU成本，根据后期的算法测试，中低端的英伟达显卡可以带4到6路视频流，硬件成本属于可控范围。算法软件成本主要来源于采购成熟的商业算法，但

商用算法的缺点也很明显：识别的目标对象和使用场景明确，对各物流企业的特定使用场景不支持或者识别准确率较低，而且目前商用算法的价格较高。

因此，基于上述各计算模式的特点和实际的成本因素，集中计算模式和公开的开源算法是合理选择。

## 3.2 算法选择

### 3.2.1 传统算法

传统算法以Harr、HOG和DPM为代表，通过手工特征进行构建，其识别目标取决于输入的正负样本，在缺乏有效的图像表示的背景下，只能设计复杂的特征，基于有限的计算资源，利用各种加速技术进行复杂计算。

### 3.2.2 两阶段算法

目前最先进的目标检测算法是基于深度学习的。这种算法首先通过对输入图片进行处理，找到可能包含目标物体的候选区域，然后使用分类器在这些候选区域对目标物体进行分类。两阶段的目标检测算法源于R-CNN。由于R-CNN的算法使用Selective Search算法对每张输入图片检测候选区域，再用CNN分类模型在候选区域中进行分类，所以运行效率低是它最大的缺点。

### 3.2.3 单阶段算法

单阶段目标检测算法能实现完整的单次训练共享特征，且在保证一定准确率的前提下，速度得到极大提升。单阶段目标检测算法不用产生候选区域，可以直接从图片中获得目标检测结果。最早的单阶段目标检测算法从YOLO开始，YOLO只需要处理一次图片就可以同时得到目标物体的位置和类别。YOLO最大的优点在于运行速度非常快，相对于两阶段的目标检测算法来说，YOLO的速度

有明显优势。针对 YOLO 系列算法定位精度的问题，2016 年提出的 SSD 算法将 YOLO 的回归思想和 Faster R-CNN 的 anchor box 机制结合在一起，保证了 YOLO 的快速运行效率，也保证了与 Faster R-CNN 相近的边界框定位精度。

综上所述，相比于传统算法，基于深度学习的 SSD 算法在特征表达、计算资源利用以及计算速度方面有了质的飞跃；和两阶段算法相比，SSD 算法省去了候选区域识别阶段，减少了识别时间；SSD 算法在 YOLO 算法的基础上进一步提高了识别精度。因此在当前阶段，SSD 算法在目标检测场景中是较优选择。

## 4 SSD 算法实施

### 4.1 实施条件

数据来源：部署在生产场地的海康威视摄像头视频流

GPU 型号：英伟达 RTX2080Ti

操作系统：CentOS

实现语言：Python3

深度学习开发平台：PyTorch

### 4.2 实施步骤

#### 4.2.1 样本采集

样本数据来源于两个场景：某物流集散中心的卸车垛口视频（约 2.4G，170 分钟）和实验室模拟回笼口拥堵场景视频。这两类数据分别用于车辆的检测训练和回笼口拥堵模拟场景训练。

视频数据不能直接用于机器学习，需要对视频流数据进行取帧操作。对于目前生产环境中的非高清视频流，每秒采集的帧数为 24，考虑到帧之间的相关性，采样程序对每秒的 24 帧只取 3 帧（每 7 帧之后取一帧）。

#### 4.2.2 样本预处理

此阶段主要是对上一步骤的数据进行检查，剔除异常值，剔除过度曝光的图片。

#### 4.2.3 标记样本数据

根据预定义的识别结果范围，对预处理后的每张图片进行标记，标记值用于告诉算法模型的真实值，算法的目的是通过减少预测值与真实值之间的误差，不断调整算法中的各类参数，从而使算法达到真实环境的最优化模拟。

#### 4.2.4 数据预处理

此阶段对样本进行处理，使其符合计算框架和算法的需求。筛选后的样本需要经过高通滤波器勾勒出感兴趣的区域（在标记阶段标注），经过低通滤波器去掉噪声；同时根据识别目的对图片进行灰度处理（即换掉 RGB 格式的背景色，用 0 ~ 255 之间的单值表示其灰度）。预处理是一个必须的步骤，经过此步骤，开发人员可以对所选的模型结果有更好的认识，同时剔除可能会影响预测结果的因素，例如逆光、背光、暗光等，这些作为噪声的因素会直接影响识别结果。预处理完成后，对样本进行分组：训练数据和测试数据。

#### 4.2.5 模型训练（包含交叉验证）

在训练之前，训练数据被分成 10 个组，训练的时候使用其中 1 组数据，训练完成后用其余 9 组数据进行验证，因此需要训练 10 次，选出 10 组模型参数中最优的模型数据。

#### 4.2.6 测试样本验证

根据上一个步骤的模型，对测试数据进行验证。

## 4.3 模型验证结果

### 4.3.1 车辆状态检测

在此次测试模型训练中，车辆的状态分为车尾 (car\_tail) 和车厢 (car\_content) 两类：车尾状态是车后门没完全打开时的状态；车厢状态是车后门完全打开的状态。

在测试过程中，算法对于倒车过程和车辆静止状态下的车尾和车厢状态的识别准确率达到 95%。检测错误的情况多出自于装满车的车厢：在货物满载的情况下，即使车门完全打开，货物仍会被错误识别为车门。这类错误目前无法解决，因为通过标记告诉算法实际结果的手段并不能告诉算法关闭的车门和满载货物的区别。

### 4.3.2 回笼口拥堵模拟环境检测

在模拟实验中，回笼口的拥堵场景通过在桌面上堆积纸箱进行呈现；根据纸箱的堆积高度，分为三类：Category-1、Category-2 和 Category-3。

在测试过程中，随着叠加的纸箱高度逐渐增大，拥堵状态逐渐从 Category-1 进化到 Category-2 和 Category-3。

经过上述测试场景可以得出，对于固定场所

下的有限范围目标, 识别的准确率较高, 这和学习样本的范围、数量有关。在学习样本中, 因为场景特定, 因此识别准确率较高。作为特定场景的目标检测, 上述样本的选取方式是可行的, 但如果识别的场景范围较大, 样本选取的范围和数量都必须相应扩大。

上述两个测试分别模拟了垛口车辆检测和回笼口拥堵。测试结果验证了 SSD 算法在非刚性特征的目标检测上有较好的性能, 同时也证明了通过增加学习样本的种类和数量, SSD 算法可以满足第一部分提出的人、物和状态的识别。

## 5 结语

SSD 算法的速度和精度优势在上述模拟测试中得到了验证。通过在生产和管理系统中集成 SSD 算法, 可以有效降低人工识别的工作量, 进一步减少人力成本。SSD 算法和其他算法一样, 也存在不足: 首先, SSD 的速度优势需通过 GPU 的加持得到体现, 在大规模部署时, GPU 的费用是很大的考量因素; 其次, 算法的学习需要大量的时间成本和人力成本, 因为样本的多样性、数量多等属性保证以及样本筛选、标注都需要人工介入。在后续工作中, SSD 的网络优化和参数调整是重点, 通过优化 SSD 算法, 提高 GPU 的利用率, 间接减少 GPU 的数量。

人工智能系统的建设有其特殊性, 不同场景有不同的实现方式, 这在传统的系统建设中是没有的, 还需要开发者结合业务场景和实际生产条件, 规划切实可行的实现路径。随着理论研究的深入, 目标检测算法在计算速度和精度方面不断提升, 使得其在现实场景的可落地性越来越大。

### 参 考 文 献

- [1] WEI LIU, ANGUELOV. SSD: Single shot multibox detector [C]. European Conference on Cvpr, 2016
- [2] IAN GOODFELLOW, YOSHUA BENGIO, AARON COURVILLE. 深度学习 [M]. 北京: 人民邮电出版社, 2017
- [3] 施巍松, 孙辉, 曹杰. 边缘计算: 万物互联时代新型计算模型 [J]. 计算机研究与发展, 2017 (5)

## 乔达集团在美洲推出 可视化数据平台

乔达集团在美洲地区推出一个数据平台——乔达可视化平台, 旨在为客户提供涵盖全供应链的 360 度数据视图。

乔达可视化平台提供了集成多个复杂操作系统的控制塔视图, 包括仓库管理系统、运输管理系统、人力管理系统和场地管理系统, 创建了跨供应链所有环节的运输简化视图。

可视化平台为客户提供了多种选项, 可以方便地与业务系统集成, 并提供供应链的统一视图。除了将所有详细数据整合在一起, 乔达可视化平台还与包裹运营商, 零担货运和整车货运商, 以及海运承运商合作, 为客户提供货物在途视图。

除了货运跟踪之外, 乔达可视化平台还为用户提供库存战略视图, 包括运费分析、服务成本分析和承运商绩效分析等供企业规划所用。用户可以快速查看、导出和分析数据。

(王旭 译)

[4] PAUL VIOLA, MICHAEL JONES. Rapid object detection using a boosted cascade of simple features [C]. Proc of CVPR, 2001

[5] NAVNEET DALAL, BILL TRIGGS. Histograms of oriented gradients for human detection [C]. Proc of CVPR, 2005

[6] PEDRO FELZENSZWALB. Object detection with discriminatively trained partbased models [C]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010

[7] ROSS GIRSHICK, JEFF DONAHUE, TREVOR DARRELL, JITENDRA MALIK. Rich feature hierarchies for accurate object detection and semantic segmentation [C]. Proc of CVPR, 2014

[8] J.R.R. UIJLINGS, K.E.A. VAN DE SANDE, T. GEVERS. Selective search for object recognition [J]. International Journal of Computer Vision, 2013 (2)