

[文章编号] 1003—4684(2021)04-0013-04

改进 Faster R-CNN 的快件搬运机器人视觉定位策略

田星星, 涂玲英, 周意入, 秦 宇
(湖北工业大学电气与电子工程学院, 湖北 武汉 430068)

[摘 要] 针对快件搬运机器人视觉导航定位精确度低的缺点,在其预设轨道上设计了一种含有 4 种不同颜色小矩形的导航图案,再利用改进的 Faster R-CNN 检测导航图案里的目标区域。然后,通过 Hams 角点检测算法提取目标区域的参考角点继而使用 PnP 算法计算出该机器人坐标和偏转角。改进的 Faster RCNN 是在 Faster RCNN 的模型基础上增加卷积层,之后在多层卷积层的 feature map 上生成候选框且用两个 3×3 卷积核分别进行卷积运算,直接进行分类和回归。最后,设计的仿真测试结果表明:两种算法检测图案中目标类别的正确率均为 100%,但改进后算法的精确度提高了 3%。

[关键词] 搬运机器人; 视觉导航定位; 导航图案; Faster R-CNN

[中图分类号] TP120.30 [文献标识码] A

随着电子商务的迅猛发展,现代化物流快件搬运已成为智能化转型的热点研究问题^[1-2]。国内第一台 AGV 由北京起重运输机械研究所研制,采用电磁导航的方式,该定位方式保养维修困难且易受外界干扰。2013 年,沈阳新松公司向市场展示了一种新型惯性导航方式的 AGV,其绝对定位的方式产生的误差具有累积性,精确度比较低^[3-5]。2015 年,阿里天猫 Geek+ 团队和国外亚马逊公司相继研发的仓储 AGV 与“Kiva”机器人,均采用二维码进行定位与导航,成功将机器视觉技术应用于 AGV,但同时二维码类型图案在工业场地上铺设保养不方便^[6-7]。

在机器视觉领域,Ross B.Girshick^[8]相继提出了 R-CNN、Fast R-CNN 和 Faster R-CNN 这一系列目标检测算法,其中 Faster R-CNN 可以看做是 RPN 和 Fast R-CNN 的结合,且用 RPN 完全替代了 Selective Search 算法生成候选框,但只在单层卷积层的 feature map 上生成候选框且 RPN 和最后的分类与回归层都各自用目标函数进行分类与回归,导致运算量还是偏大,精度仍不高。

为改进上述问题,在 Faster-RCNN 的网络模型基础上增加卷积层,在多层卷积层的 feature map 上生成候选框且用两个 3×3 卷积核分别进行卷积运算,直接进行分类和回归。该改进的 Faster-RCNN 检测到所设计导航图案中目标的区域后,再通

过 Harris 角点检测算法提取目标区域的参考角点,继而使用 PnP 算法计算出快件搬运机器人的坐标位置和偏转角。最后,通过仿真设计的实验对其有效性进行验证。

1 快件搬运机器人模型

图 1 为快件搬运机器人模型三维图,基本结构主要分为驱动结构和举升结构两个部分。其中,驱动结构左右两边选用相同的体小量轻和响应速度快的伺服电机,且对称安装。当两台伺服电机连线中点位置的工业相机定位到位置信息后,每台伺服电机分别带动一个驱动轮,两个驱动轮再通过前后四个相同规格且同样对称安装的小万向轮实现该机器人的差速行驶转向及位置纠偏。

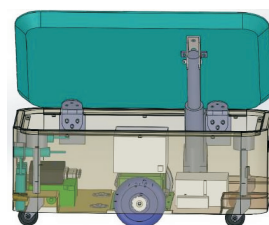


图 1 快件搬运机器人的三维模型

快件搬运机器人一般在工业场地上沿着预设好的轨道中线运动,图 2 为其差速驱动数学模型,在 XOY 坐标系中左右两侧驱动轮的速度分别设为 V_r 和 V_l ,机器人移动线速度设为 V ,移动角速度为

[收稿日期] 2021—03—04
[基金项目] 国家自然科学基金青年科学基金项目(41601399)
[第一作者] 田星星(1993—),男,湖北蕲春人,湖北工业大学硕士研究生,研究方向为机器视觉及电机运动控制
[通信作者] 涂玲英(1963—),女,湖北武汉人,湖北工业大学教授,研究方向为电子信息及嵌入式技术

ω ,左右轮之间的距离设为 D ,则在当前坐标系中该机器人的移动线速度和移动角速度关系可表示为:

$$V = (V_r + V_l)/2$$

$$\omega = (V_r - V_l)/D$$

假设该机器人的转弯半径为 R ,则

$$R = V/\omega = D(V_r + V_l)/(V_r - V_l)$$

由此,通过对该机器人两轮速度作不同的规划,能够得到相对应的不同的运动状态如下:

1) $V_r = V_l$ 时,该机器人作直线运动,并且 $V = V_r = V_l$, R 为无穷远。

2) $V_r > V_l$ 时,该机器人向右作圆弧运动,并且 $V = \frac{V_r + V_l}{2}$, R 为 $\frac{D(V_r + V_l)}{2(V_r - V_l)}$ 。

3) $V_r < V_l$ 时,该机器人向左作圆弧运动,并且 $V = \frac{V_r + V_l}{2}$, R 为 $\frac{D(V_r + V_l)}{2(V_r - V_l)}$ 。

4) $V_r = -V_l$ 时, R 为 $\frac{D}{2}$, $V = 0$,即该机器人以自身中心点旋转掉头。

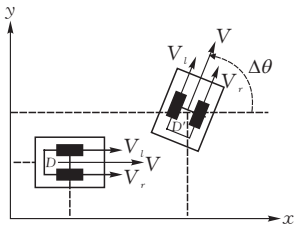


图2 快件搬运机器人的差速运动模型

进一步地进行该机器人行驶时速度与位置的数学关系分析,选择其工业相机中心点作为载体质心,任意时刻轴连线的中垂线和 X 轴的夹角记为 $\theta(t)$,则在 Δt 时间内该机器人所转过的角度 $\Delta\theta$ 与转动角速度 ω 的关系可表示为:

$$\frac{\Delta\theta}{\Delta t} \approx \frac{d\theta(t)}{dt} = \omega(t) = \frac{V_r(t) - V_l(t)}{D}$$

积分可得到 $\theta(t)$ 与左右车轮运动速度 $V_l(t)$ 和 $V_r(t)$ 关系如下:

$$\theta(t) = \theta(t_0) + \int_{t_0}^t \frac{1}{D} [V_r(t) - V_l(t)] dt$$

从上式可知:该机器人上的工业相机定位到偏转角等位置信息后,其主控制器处理后通过调节两台伺服电机的差速输出即可得到新的位置。

2 导航定位策略

2.1 导航图案设计及其目标检测算法概述

当快件搬运机器人沿着预设好的轨道中线行驶时,笔者拟在其轨道中央等间距地铺设一些导航图案,行驶路径导引图见图 3。这些图案间距小于工

业相机的视觉直径,并且每张都是以白色为衬底,里面所包含的小矩形大小不一,颜色均为红色、黄色、绿色和蓝色。这种导航图案不仅设计简单,而且在工业场地上铺设保养极为方便。

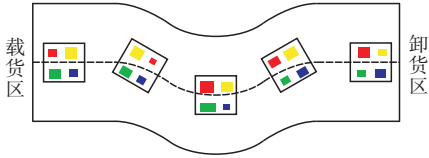


图3 行驶路径导引图

近些年来,计算机视觉技术在各类机器人导航定位中得到发展,R-CNN, Fast R-CNN 和 Faster R-CNN 这一系列目标检测算法被相继提出,其可以检测上述导航图案中的目标区域。

R-CNN 首先通过 Selective Search 算法得到大约 1000—2000 个候选框,再将每个候选框尺寸固定到 227×227 后送入 CNN 中提取特征,然后将这些 feature map 集合放入 21 个 SVM 中进行分类,之后利用非极大值抑制算法得到每个目标类别中得分最高的生成框,最后用回归器对生成框回归操作以精修其位置。由于每个候选框都需要送入 CNN 中逐一提取,这种目标检测算法图像处理速度过慢,且分类与回归没有统一到一个网络模型中,占用大量的内存空间。

Fast R-CNN 先将原图送入 CNN 中提取特征,再将 Selective Search 算法得到的近 2000 个候选框映射到刚提取的 feature map 上,生成对应的图像块,然后通过 ROI-Pooling 层统一到相同的尺度上,最后经过两个全连接层的特征连在各自的损失函数上,一个是用于分类的 softmax 函数,另一个是用于生成框回归的 smooth 的 L1 范式。这种目标检测算法只在末尾处理每个候选框,并且将分类与回归统一到一个网络模型中,既节省了图像处理时间又节省了内存空间。但是,候选框用 Selective Search 算法来获取,仍存在速度慢的问题。

Faster R-CNN 的模型框图见图 4,其可以看作是 RPN 和 Fast RCNN 的结合。

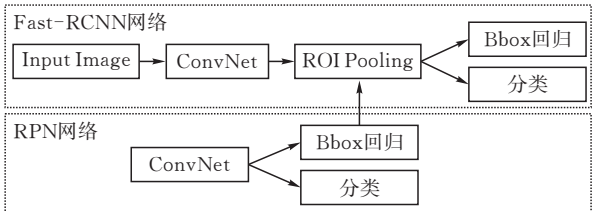


图4 Faster-RCNN 模型框图

Faster R-CNN 首先利用 Fast R-CNN 将原图送入 CNN 中进行特征提取,接着 RPN 网络在某一卷积层提取的 feature map 上利用锚框生成法生成

锚框。锚框生成法在 feature map 所划分的单元格中心点生成 k 个不同尺寸和长宽比的候选框,初步回归和分类后传入 Fast RCNN 的 ROI Pooling 层,统一到相同的尺寸,然后进行精确的分类和回归。锚框生成模型见图 5。

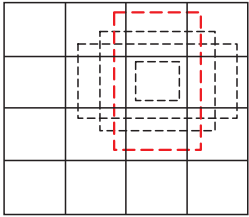


图 5 锚框生成模型

这种目标检测算法中的 RPN 虽然将 Selective Search 算法产生候选框的环节有效地替代并整合到了深度卷积神经网络中,但是 RPN 只在单层的 feature map 上生成候选框, 这样会遗漏一些小目

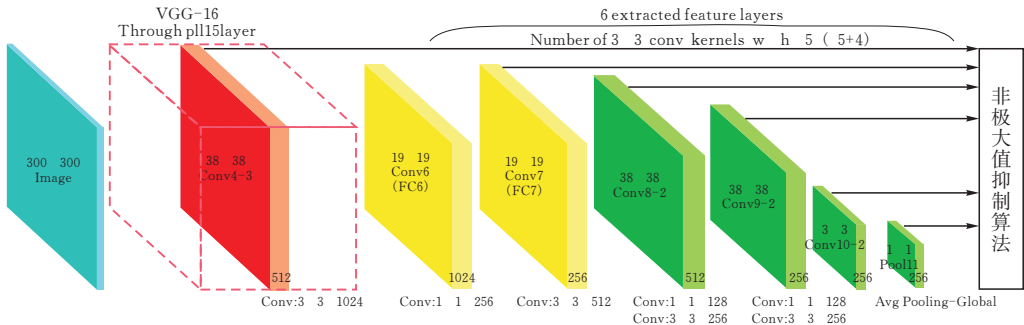


图 6 改进的 VGG-16 网络模型

对于提取为 $W \times H$ 的 feature map,锚框生成法生成候选框的个数为 $5 \times W \times H$ 个,尺度集设为 $[1:1,1:2,1:3,2:1,3:1]$,尺寸的大小由下式定义:

$$\text{size}_k = \text{size}_{\min} + \frac{\text{size}_{\max} - \text{size}_{\min}}{m - 1} (k - 1); k \in [1, m]$$

可知 m 为 5; size_{\min} 为尺寸最小值,取 0.2; size_{\max} 为尺寸最大值,取 0.96。尺度集用 $p_r = \{1, 2, 3, 1/2, 1/3\}$ 表示,则每一个候选框的长宽可由下式进行计算:

$$\begin{cases} w_k^a = s_k \sqrt{p_r} \\ h_k^a = s_k / \sqrt{p_r} \end{cases}$$

当用锚框生成法生成这些候选框后,将其映射回原图基本上覆盖了整张图片,然后计算每个候选框和真实框区域位置的交并比,即重叠程度 IOU。当 $\text{IOU} > 0.5$ 时标记为正样本,否则标记为负样本。由于负样本数量过多会造成网络模型难以收敛,故通过正负样本数量 1:3 的比例和 IOU 值排序从高到低两个原则剔除一些 IOU 值特别小的负样本。

2.3 Harris 角点检测算法与 PnP 定位算法介绍

Harris 角点检测算法的核心是利用局部窗口

标。其次,RPN 和最后的分类与回归层都各自用目标函数分类与回归,导致运算量还是偏大。

2.2 Faster-RCNN 算法的改进

针对 Faster RCNN 目标检测算法存在的问题作出进一步改进,即在多层卷积层的 featuremap 上生成候选框后,直接进行分类与回归。具体设计方案是:首先选用 VGG-16 网络模型(CNN 常用的网络模型),将其最后两层全连接层改成卷积层,并在后面增加四层卷积层(图 6),再把 Conv4-3、Fc7、Conv8-2、Conv9-2、conv10-2 这 5 层不同卷积层提取到的 feature map 划分成 $8 \times 8, 4 \times 4$ 两种尺寸的单元格,然后用锚框生成法在每个划分成的单元格中心点生成候选框,之后用两个 3×3 卷积核分别进行卷积运算,最后一个用 softmax 函数进行多分类,另一个用 smooth 的 L1 范式进行回归,不进行二次分类与回归。

在图像上进行移动判断灰度是否发生较大的变化。若从各个方向上移动小窗口,此窗口区域内图像的灰度发生了较大变化,那么就认为遇到了角点。该算法可以有效地提取上述改进后目标检测算法定位目标区域的角点。

PnP 算法最后解决的是:已知相机内部参数(焦距参数和畸变参数)以及 n 个世界坐标系和图像坐标系参考点的情况下,利用这几个参考点之间的对应关系求解相机在拍摄图像时所处位姿。该算法可根据 Harris 角点检测算法提取的角点计算出快件搬运机器人沿预设路径时的偏转角以及坐标位置。

3 改进的 Faster RCNN 算法仿真实验

3.1 仿真实验图像样本训练与测试

针对 2.3 提出的目标检测算法优化模型的训练与测试:在第一阶段:利用 Pascal Voc 数据集包含的 21 个类别近 30000 张图片进行迁移学习,满足其大数据量学习的需要;在第二阶段,先通过绘图软件绘制了 125 张 2.1 方案所设计的图片,之后对这些图片进行了数据增强处理,其步骤是先通过镜像翻

转处理,图片数量增加至4倍变为500张,再对每张图片进行模糊处理,图片数量增加至1000张且增强了该算法在模糊背景下的目标检测能力。最后将这1000张图像样本分为两份,80%的样本用于第二次训练,剩余20%的样本用于最终的测试。

本实验硬件采用 gtx1070 显卡以及 inter corei7 处理器,软件环境为 Windows 10 操作系统,在 Tensorflow 框架下 python 语言进行开发,图 7 和图 8 分别为 Faster RCNN 算法和改进的 Faster RCNN 算法在相同实验条件下部分测试样本的结果显示,其中 Rr、Yr、Gr、Br 分别代表红色矩形、黄色矩形、绿色矩形和蓝色矩形。

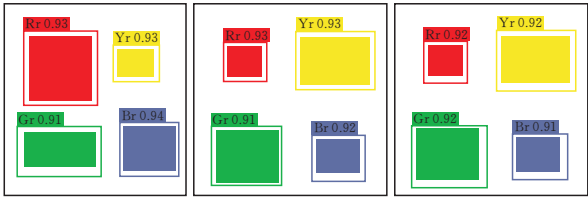


图 7 Faster RCNN 算法两组测试结果

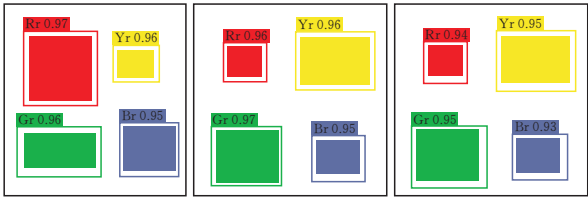


图 8 改进的 Faster RCNN 算法两组测试结果

3.2 仿真实验测试结果分析

首先定义改进前和改进后两种算法的图像样本测试正确率和精确度分别为 cor1、cor2 和 acc1、acc2,则改进前算法图像每个样本里红、黄、绿、蓝四钟颜色小矩形的识别正确值(正确区分出四色小矩形值为 1,反之为 0)和定位精确值(测试图像结果显示的值)分别为 $R - cor1_i$, $Y - cor1_i$, $G - cor1_i$, $B - cor1_i$ 和 $R - acc1_i$, $Y - acc1_i$, $G - acc1_i$, $B - acc1_i$;改进后算法的图像每个样本里红、黄、绿、蓝四钟颜色小矩形的识别正确率和定位精确度分别为 $R - cor2_i$, $Y - cor2_i$, $G - cor2_i$, $B - cor2_i$ 和 $R - acc2_i$, $Y - acc2_i$, $G - acc2_i$, $B - acc2_i$ 。cor1、cor2 和 acc1、acc2 的计算公式分别如下:

$$\begin{cases} cor1 = \frac{\sum_{i=1}^{200} R - cor1_i + Y - cor1_i + G - cor1_i + B - cor1_i}{4 \times 200} \\ cor2 = \frac{\sum_{i=1}^{200} R - cor2_i + Y - cor2_i + G - cor2_i + B - cor2_i}{4 \times 200} \end{cases}, i \in (1, 2, 3, \dots, 200)$$
$$\begin{cases} acc1 = \frac{\sum_{i=1}^{200} R - acc1_i + Y - acc1_i + G - acc1_i + B - acc1_i}{4 \times 200} \\ acc2 = \frac{\sum_{i=1}^{200} R - acc2_i + Y - acc2_i + G - acc2_i + B - acc2_i}{4 \times 200} \end{cases},$$

$i \in (1, 2, 3, \dots, 200)$

根据以上公式计算改进前后两种算法测试 200 个样本的统计情况见表 1。

表 1 改进前、后算法正确率和准确度统计表

	cor1	cor2	acc1	acc2
value	1	1	0.93	0.96

4 结论

对快件搬运机器人驱动结构的导航定位技术进行研究,在 Faster R-CNN 的网络模型基础上增加卷积层,在多层卷积层的 feature map 上生成候选框且用两个 3×3 卷积核分别进行卷积运算,直接进行分类和回归。测试结果显示相比 Faster-RCNN,改进的目标检测算法定位图像中目标类别的正确率同为 100%,精确度提高了 3%。之后用 Harris 角点检测算法提取参考角点,继而使用 PnP 算法计算出坐标位置,也即提高了该机器人的导航定位精度。

[参 考 文 献]

[1] 陈宏勋,张翔. 物流技术与装备[M]. 北京:国防工业出版社,2005,26-27.

[2] 张文利,周友良. 我国智能制造装备产业发展的调控策略[J]. 经济纵横,2016,26(12):87-90.

[3] 陈谦益. 磁带导引 AGV 小车中央控制系统的设计[D]. 武汉:武汉理工大学,2014.

[4] 江杰,彭帅,李刚. 基于多传感器信息融合的惯性导引 AGV 设计[J]. 自动化应用,2015,(10):58-61.

[5] 杜玉立. 智能搬运机器人的导航策略研究及应用[D]. 武汉:武汉工程大学,2017.

[6] 李思源. 机器人化的仓储物流系统动态调度研究[D]. 天津:南开大学,2015.

[7] LI J, YANG W M, ZHOU D Z. Modeling study on the effect of piston bowl geometries in a gasoline biodiesel fueled RCCI engine at high speed[J]. Energy Conversion and Management, 2016,112(15):359-368.

[8] 廖鹏. 深度学习实践:计算机视觉[M]. 北京:清华大学出版社,2019.

[9] 周建浩. 用于物流中心的快件自动分拣机器人的研究[D]. 广州:华北理工大学,2018.

[10] 田星星,夏耀威,张洪涛. 一种基于图像处理的运输机器人:中国.CN201911352371.5[P].2020-03-31.

[11] 黄晓浪. 基于灰度变化的角点检测算法研究[D]. 上海:东华理工大学,2019.

[12] 席志红,李爽,甘兴利. PnP 算法在室内定位中的应用[J]. 无线电工程,2017,47(10):39-44.