

结合深度学习的机械臂视觉抓取控制

白成超¹, 晏卓², 宋俊霖¹

(1. 哈尔滨工业大学航天工程系, 哈尔滨 150001; 2. 西安航空动力试验技术研究所, 西安 710000)

摘要: 针对基于视觉的机械臂抓取精确抓取的需求, 考虑传统的视觉识别算法受环境、对象变化的制约且在识别正确率及快速性上存在的问题, 在现有研究的基础上, 提出了一种基于深度学习的目标精确检测与识别方法。首先基于深度学习改进了 YOLO 算法, 通过对数据集的训练, 基于英伟达 Jetson TX1 高性能处理单元实现了复杂环境下多目标的识别与定位, 给出了目标的类别与位置等信息; 以此为基础, 结合利用 MoveIt! 功能包完成的机械臂运动轨迹的求解与规划, 以及基于李群李代数建立的递推正逆动力学模型, 设计了机械臂抓取控制的滑模控制律。仿真及实物验证表明, 基于深度神经网络的方法学习到的特征对复杂背景具有较强的鲁棒性和稳定性; 所设计的滑模控制算法在 0.21 s 时跟踪误差在 2%, 取得了较高的控制精度。可为后续视觉抓取任务提供参考。

关键词: 机械臂; 李群李代数; 正逆动力学; 滑模控制; 深度学习

中图分类号: V529.1 **文献标识码:** A **文章编号:** 1674-5825(2018)03-0299-09

DOI:10.16329/j.cnki.zrht.2018.03.003

Visual grasp control of robotic arm based on deep learning

BAI Chengchao¹, YAN Zhuo², SONG Junlin¹

(1. Department of Aerospace Engineering, Harbin Institute of Technology, Harbin 150001, China;

2. Xi'an Aerospace Propulsion Test Technology Institute, Xi'an 710000, China)

Abstract: The precise capture of the robotic arm plays a very important role in actual military or industrial production. Vision is widely used because of its high efficiency, low price, etc. Therefore, robotic arm grabbing based on vision has gradually become a research hotspot in this field. However, the traditional visual recognition algorithms are restricted by the changes of the environment and the objects. At the same time, there exist recognition accuracy and rapidity problems. Based on the existing researches, an algorithm based on deep learning was proposed in this paper to accurately detect and identify the target so as to improve the recognition ability in a complex environment. First, the YOLO algorithm was improved by deep learning. Based on the high-performance processing unit of Nvidia Jetson TX1, the identification and location of multi-targets in a complex environment were achieved through data sets training, and the information such as the category and location of the target were given. Then, based on the solution and planning of the trajectory of the robot arm solved by the MoveIt! Function Package, and the efficient recursive and inverse dynamic model established according to Lie groups and Lie algebra theory, the sliding mode control algorithm for robotic arm grab control was designed. The simulation and physical verification showed that compared with the traditional image recognition algorithm, the features learned from the deep neural network method have stronger robustness and stability for the complex background. The tracking error of the designed sliding mode control algorithm was 2% at 0.21s and with high control accuracy which may provide relevant references for subsequent visual grabbing tasks.

Key words: manipulator; theory of Lie groups and Lie algebras; forward and inverse dynamics; sliding mode control; deep learning

收稿日期: 2017-11-15; 修回日期: 2018-03-05

基金项目: 载人航天预先研究项目(060101)

第一作者: 白成超, 男, 博士研究生, 研究方向为无人系统智能感知与自主规划技术。E-mail: baichengchao@hit.edu.cn

1 引言

人类航天活动产生了大量的空间碎片和太空垃圾,严重威胁航天器的正常运行,有必要对其进行监控管理和回收^[1]。机械臂是完成这个任务的重要工具,可以抓取接近航天器的太空垃圾,使其脱离轨道,从而有效地保护航天器^[1]。机械臂抓取目标首先需要知道目标的位置,这就需要利用传感器检测与识别技术识别出目标,同时计算出其准确的空间位置,因传感器的类型不同,识别原理也不同,本文主要基于相机来研究视觉精确识别。

机械臂视觉抓取在常规的机器人学、控制理论的基础上,融合了计算机视觉理论、深度学习和人工智能等学科,具有重要的科研和应用价值^[2]。其中涉及的目标检测技术在视觉领域是现在非常热门的研究方向,传统目标检测通常是静态检测,而且目标单一,目标检测受到外形、大小、视角的变化以及外部光照的变化等因素的影响,因而所提取的特征泛化能力不强,鲁棒性较差,通常需要对每一个特定目标都提出相适应的算法^[2],难以适应新物体。

基于深度学习的目标识别与定位是近年来的研究热点,尤其是深度学习中卷积神经网络的发展直接促进了目标识别和定位任务的进步。近几年的 CVPR、ECCV 等国际计算机视觉顶级会议不断有新的深度学习网络模型提出,并不断刷新各种大规模图像竞赛的精度。目标检测的精度和速度是两个主要的问题,基于区域选择的 R-CNN 系列网络与 End-to-End 的 YOLO、SSD 等网络在速度和精度上均有较大的进步^[3]。

机械臂视觉抓取系统中基于深度学习的视觉识别可通过对深度卷积神经网络的训练实现多目标快速识别而不必更改识别算法。这需要搜集被检测目标的大量图像制作标准的训练集,利用前述基于深度神经网络的目标识别算法对数据集进行离线训练,将得到的模型用于在线识别。R-CNN 系列算法^[3-4]是目前主流的用于机械臂抓取的深度学习目标检测算法,但是速度上并不能满足实时的要求。本文基于回归的思想搭建深度神经网络,并通过改进 YOLO 算法进行识别率及实时性的验证,即在小样本环境下,给定输入图像,

直接在图像上回归出目标位置和目标类别,并同时满足实时检测的要求。

2 基于深度学习的多目标识别算法

本文用于机械臂待抓取目标识别的深度学习多目标识别算法使用 YOLO 算法^[3],其流程如图 1 所示:

1) 采集摄像头原始图像,将原图划分为 $S \times S$ 的网格。

2) 每一个网格会预测 B 个 bounding box (Object 在图片所占的矩形框,下面记为 bbox) 和 confidence(置信度)得分,记录的五元素具体为 $(x, y, w, h, P(\text{Object}) * IOU)$, 其中 $P(\text{Object})$ 表示当前位置是一个 Object 的概率, IOU 预测的是重叠概率。其中 x, y 是中心坐标; w, h 是 bbox 的宽度和高度。

3) 计算每一个网格预测的类别概率 $C_i = P(\text{Class}_i | \text{Object})$ 。

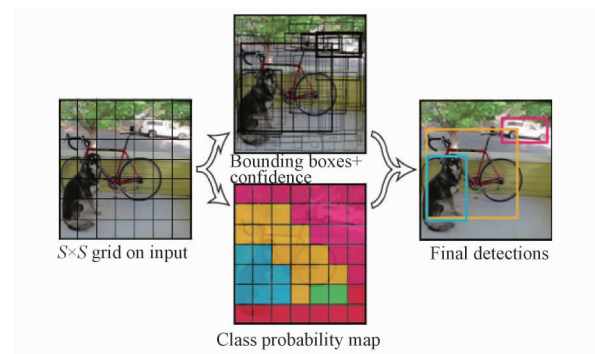


图1 YOLO 算法流程
Fig. 1 Flow chart of YOLO algorithm

YOLO 算法利用卷积神经网络实现目标识别模型,并在 PASCAL VOC 数据集上进行评估。网络的初始卷积层从图像中提取特征,而全连接层预测输出类别概率和相应图像坐标。网络结构参考同样应用于图像分类的 GoogLeNet 模型^[4],有 24 个卷积层,然后是 2 个全连接层,如图 2 所示。

在网络训练误差中,大 bbox 和小 bbox 权值相同。另外,网络训练误差的度量方式应该反映出大 bbox 的小偏差与小 bbox 的小偏差相比更不显著。为了解决这个问题,回归的是 bbox 宽度和高度的平方根,而不是直接的宽度和高度。YOLO 在每个网格中预测出多个 bbox。在训练网络时,对每个类别,只需要用一个 bbox 来负责该类别的

系下的三维坐标,通过双目立体视觉计算出末端执行器在摄像机坐标系下的三维坐标(通过第二章得到)。操作机械臂改变末端执行器位置我们便可以得到一组点在不同坐标系下的坐标值。

问题转化为已知空间中的同一组点在不同坐标系(摄像机坐标系、机械臂坐标系)下的坐标 $\mathbf{S} = \{(\mathbf{X}_i, \mathbf{X}'_i)\}$, 计算两个坐标之间的变换矩阵。 \mathbf{X}_i 是由齐次向量表示的第 i 个三维点, \mathbf{X} 是 $4 \times N$ 矩阵包含了 N 个点, 这些点均由齐次向量表示。得到一组对应点后, 问题变成在 3D 刚性场景的找最优 Γ_r 的点云配准问题, 可以为式 (2) 所示最小二乘问题:

$$\Gamma_r^* = (\mathbf{R}, \mathbf{T})^* = \underset{\Gamma_r \in SE(3)}{\operatorname{argmin}} \sum_{i=1}^N \|\Gamma_r \mathbf{X}_i - \mathbf{X}'_i\|_{12} \quad (2)$$

Γ_r 需要满足 $SE(3)$ 的所有约束。利用李群结构 $SE(3)$ 及其相关李代数 $se(3)$ 可以将约束问题可以转化为无约束优化问题。定义代价函数如式 (3):

$$(\boldsymbol{\omega}, \mathbf{v})^* = \underset{(\boldsymbol{\omega}, \mathbf{v}) \in se(3)}{\operatorname{argmin}} C(\boldsymbol{\omega}, \mathbf{v}) \quad (3)$$

其中 $C(\boldsymbol{\omega}, \mathbf{v}) = \sum_{i=1}^N d(\exp(\boldsymbol{\omega}, \mathbf{v}) \mathbf{X}_i, \mathbf{X}'_i)$, 为了后面计算 Jacobian 矩阵方便, 令 $d(\cdot, \cdot) = \|\cdot\|_{12}^2$ 。

在对代价函数化简之前, 有必要对输入数据进行规范化操作。这有利于防止收敛到错误的解, 且减少所需的迭代次数。分别对点集 \mathbf{X} 和 \mathbf{X}'

$$\text{计算如下转换矩阵: } \mathbf{w} = \begin{bmatrix} \mathbf{I}_{3 \times 3} & -\frac{\sum_{i=1}^N \mathbf{X}_i}{N} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix},$$

$$\mathbf{w}' = \begin{bmatrix} \mathbf{I}_{3 \times 3} & -\frac{\sum_{i=1}^N \mathbf{X}'_i}{N} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix}。$$

对所有点进行规范化操作 $\tilde{\mathbf{X}}_i = \mathbf{w} \mathbf{X}_i, \tilde{\mathbf{X}}'_i = \mathbf{w}' \mathbf{X}'_i$ 。对规范化后的点进行后续优化操作, 当找到最优的 $(\tilde{\mathbf{R}}, \tilde{\mathbf{T}})^*$ 时, 真实的最优解可通过式 (4) 计算:

$$(\mathbf{R}, \mathbf{T})^* = \mathbf{w}^{-1} (\tilde{\mathbf{R}}, \tilde{\mathbf{T}})^* \mathbf{w} \quad (4)$$

求代价函数对 $se(3)$ 的 Jacobian 矩阵如式 (5):

$$\begin{aligned} & \frac{\partial C}{\partial (\boldsymbol{\omega}, \mathbf{v})_{m,n}} \\ &= \operatorname{tr} \left(\left[\frac{\partial \exp(\boldsymbol{\omega}, \mathbf{v})}{\partial (\boldsymbol{\omega}, \mathbf{v})_{m,n}} \mid \mathbf{O}_{4 \times 4} \right] \right. \\ & \quad \left. (\mathbf{M}^T + \mathbf{M}) [\exp(\boldsymbol{\omega}, \mathbf{v}) \mid -\mathbf{I}_{4 \times 4} \mathbf{1}^T] \right) \quad (5) \end{aligned}$$

将上述算法记为算法 3, 其流程总结如下:

- 1) 将 \mathbf{X} 和 \mathbf{X}' 标准化, 并且计算 \mathbf{w} 和 \mathbf{w}' ;
- 2) 生成标准化后的点集, $\forall i \tilde{\mathbf{X}}_i = \mathbf{w} \mathbf{X}_i, \tilde{\mathbf{X}}'_i = \mathbf{w}' \mathbf{X}'_i$;
- 3) 利用标准化后的点集 \mathbf{X} 和 \mathbf{X}' 计算 \mathbf{M} ;
- 4) \mathbf{x} 随机取一组数 $\mathbf{x}_0 \in \mathbb{R}^6$, 满足 $\|\mathbf{x}_0\| < \pi$, 其中 \mathbf{x} 为优化变量 $(\boldsymbol{\omega}, \mathbf{v})$;

5) 利用梯度下降法优化代价函数, 不断迭代 \mathbf{x} 直到它收敛于 \mathbf{x}^* ;

6) 利用 \exp 将 \mathbf{x}^* 映射到 $SE(3)$, 得到 $(\tilde{\mathbf{R}}, \tilde{\mathbf{T}})^*$ 。

$(\mathbf{R}, \mathbf{T})^* = \mathbf{w}^{-1} (\tilde{\mathbf{R}}, \tilde{\mathbf{T}})^* \mathbf{w}$, 算法结束, 得到摄像机坐标系与机械臂坐标系之间的变换关系。在第二章通过深度学习算法可以获得目标在摄像机坐标系下的三维坐标, 利用这里的位姿变换关系可以转化为机械臂坐标系的三维坐标, 这样便可以用于后续机械臂的路径规划即期望轨迹的计算。

3.3 期望轨迹计算

利用前述系列算法已经可以获得目标在机械臂坐标系下的三维坐标, 现在将为机械臂从初始点到目标点规划一条合理的路径, 这里采用 MoveIt! 工具包^[6], 运动规划算法基于第三方运动学库 OMPL (Open Motion Plan Library)。

利用 Rviz 和 MoveIt! 使模型在 ROS 系统进行运动仿真时可视, 这样的话, 给定机械臂末端执行器的起点和终点位置及姿态, OMPL 库就可以在线完成运动规划。机械臂模型利用 MoveIt! 在从初始位姿到目标位姿的运动规划仿真如图 4 所示。

3.4 基于李群李代数的机械臂动力学模型

机械臂逆动力学通常用于计算使动力学系统沿着规定的关节轨迹运动所需的关节力矩, 假设机械臂各关节角及角速度 $(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}})$ 已知(期望轨迹计算)。通过给定的关节加速度 $\ddot{\boldsymbol{\theta}}$ 计算关节力矩 $\boldsymbol{\tau}$ 。另一方面, 机械臂正动力学通常用于积分加速

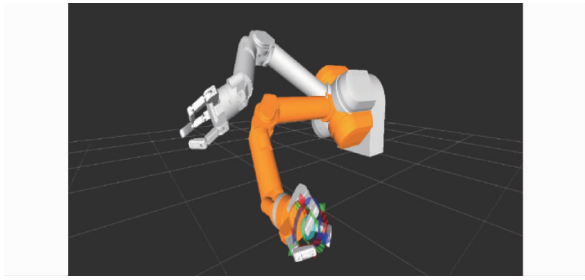


图4 利用 MoveIt! 完成运动规划
Fig.4 Motion planning made by MoveIt!

度得到系统下一个时间步长的状态 $(\theta, \dot{\theta})$ 仿真系统的运动^[7-9], 利用给定的关节力矩 τ 计算关节加速度 $\ddot{\theta}$ 。定义如下变量:

i : 第 i 个关节的索引;

$\lambda(i)$: 第 i 个关节父关节的索引;

$\mu(i)$: 第 i 个关节子关节的索引;

$\{i\}$: 第 i 个关节的固定坐标系;

T : 坐标变换矩阵;

I_i : 第 i 个关节的广义惯量;

V_i : $\{i\}$ 系下第 i 个关节的广义速度;

F_i : 第 i 个关节所受的广义力;

令 $S_i = T_{\lambda(i),i}^{-1} (\partial T_{\lambda(i),i} / \partial \theta_i)$, 将逆动力学算法记为算法 1, 计算逆动力学的算法包含下面两个递推过程, 总结如下:

1) 前向递推

$$\begin{cases} T_{\lambda(i),i} = f(\theta) \\ V_i = Ad_{T_{\lambda(i),i}} V_{\lambda(i)} + S_i \dot{\theta}_i \\ \dot{V}_i = Ad_{T_{\lambda(i),i}} \dot{V}_{\lambda(i)} + ad_{V_i} S_i \dot{\theta}_i + \dot{S}_i \dot{\theta}_i + S_i \ddot{\theta}_i \end{cases} \quad (6)$$

2) 反向递推

$$\begin{cases} F_i = I_i \dot{V}_i - ad_{V_i}^* I_i V_i - F_i^{\text{ext}} + \sum_{k \in \mu_i} Ad_{T_{i,k}}^* F_k \\ \tau_i = S_i^T F_i \end{cases} \quad (7)$$

将正动力学算法记为算法 2, 计算系统正动力学的算法包含三个递推过程。假设机械臂各关节角及角速度已知(期望轨迹计算)。通过给定的关节加速度计算关节力矩。

递推过程如下:

1) 前向递推

$$\begin{cases} T_{\lambda(i),i} = f(\theta) \\ V_i = Ad_{T_{\lambda(i),i}} V_{\lambda(i)} + S_i \dot{\theta}_i \\ \eta_i = ad_{V_i} S_i \dot{\theta}_i + \dot{S}_i \dot{\theta}_i \end{cases} \quad (8)$$

2) 反向递推

$$\begin{cases} \hat{I}_i = I_i + \sum_{k \in \mu(i)} Ad_{T_{i,k}}^* I_k Ad_{T_{i,k}} \\ \hat{B}_i = -ad_{V_i}^* I_i V_i - F_i^{\text{ext}} + \sum_{k \in \mu(i)} Ad_{T_{i,k}}^* \beta_k \\ \Psi_i = (S_i^T \hat{I}_i S_i)^{-1} \\ \Pi_i = \hat{I}_i - \hat{I}_i S_i \Psi_i S_i^T \hat{I}_i \\ \beta_i = \hat{B}_i + \hat{I}_i \{ \eta_i + S_i \Psi_i [\tau_i - S_i^T (\hat{I}_i \eta_i + \hat{B}_i)] \} \end{cases} \quad (9)$$

3) 前向递推

$$\begin{cases} \ddot{\theta}_i = \Psi_i [\tau_i - S_i^T \hat{I}_i (Ad_{T_{\lambda(i),i}} \dot{V}_{\lambda(i)} + \eta_i) - S_i^T \hat{B}_i] \\ \dot{V}_i = Ad_{T_{\lambda(i),i}} \dot{V}_{\lambda(i)} + S_i \ddot{\theta}_i + \eta_i \\ F_i = \hat{I}_i \dot{V}_i + \hat{B}_i \end{cases} \quad (10)$$

3.5 滑模控制律

相比于其他非线性控制方法, 滑模控制主要优点是对动力学建模误差和外部干扰具有良好的鲁棒性^[10]。将机械臂动力学方程(6)、(7)总结如式(11):

$$M(\theta) \ddot{\theta} + C(\theta, \dot{\theta}) + G(\theta) = \tau + d \quad (11)$$

其中 d 为干扰力矩向量, 表示机械臂真实动力学模型与名义动力学模型的偏差。下面给出两个假设^[11]:

1) $\|M^{-1}(\theta)\| \leq \alpha$, 这里 α 为有界正常数。

2) 干扰力矩向量 $d = [d_1 \ d_2 \ d_3 \ d_4]^T$ 也是有界的, 即 $\|d_i\| \leq \beta_i, i = 1, 2, 3, 4$ 。

这样机械臂动力学模型可以用式(12)表示:

$$\begin{cases} \dot{x} = z \\ \dot{z} = f(x, z) + g(x) u + g(x) d \end{cases} \quad (12)$$

定义跟踪误差: $e \triangleq x - x_d$, 设滑模面为式(13):

$$S = \ddot{e} + \lambda \dot{e} \quad (13)$$

这里 $\lambda = \text{diag}(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$, $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ 为正常数。首先需要设计控制律 u 使得 S 趋于 O 。设计 Lyapunov 函数如式(14):

$$\dot{\mathbf{V}} = \frac{1}{2} \mathbf{S}^T \dot{\mathbf{S}} \quad (14)$$

对时间求导得加速度矩阵如式(15):

$$\begin{aligned} \dot{\mathbf{V}} &= \mathbf{S}^T \dot{\mathbf{S}} \\ &= \mathbf{S}^T (\ddot{\mathbf{e}} + \lambda \dot{\mathbf{e}}) \\ &= \mathbf{S}^T [f(\mathbf{x}, \mathbf{z}) + g(\mathbf{x}) \mathbf{u} + g(\mathbf{x}) \mathbf{d} - \ddot{\mathbf{x}}_d + \lambda(\mathbf{z} - \dot{\mathbf{x}}_d)] \end{aligned} \quad (15)$$

从而控制律可设计为式(16):

$$\mathbf{u} = g(\mathbf{x})^{-1} [-f(\mathbf{x}, \mathbf{z}) + \ddot{\mathbf{x}}_d - \lambda(\mathbf{z} - \dot{\mathbf{x}}_d) - \mathbf{K} \text{sgn}(\mathbf{S})] \quad (16)$$

这里 $\mathbf{K} = \text{diag}(k_1, k_2, k_3, k_4)$, k_1, k_2, k_3, k_4 为正常数。将 \mathbf{u} 代入 $\dot{\mathbf{V}}$ 得式(17):

$$\begin{aligned} \dot{\mathbf{V}} &= \mathbf{S}^T [g(\mathbf{x}) \mathbf{d} - \mathbf{K} \text{sgn}(\mathbf{S})] \\ &= - \sum_{i=1}^4 k_i |s_i| + \mathbf{S}^T \mathbf{M}^{-1} \mathbf{d} \\ &\leq - \sum_{i=1}^4 |s_i| (k_i - \alpha \beta_i) \end{aligned} \quad (17)$$

如果满足 $k_i > \alpha \beta_i$, 将有 $\dot{\mathbf{V}} < 0$, 这样便保证了系统在 Lyapunov 意义下是稳定的, 即保证跟踪误差会收敛于 0。代入公式(16)可得式(18):

$$\mathbf{u} = \mathbf{C} + \mathbf{G} + \mathbf{M} [\ddot{\mathbf{x}}_d - \lambda(\mathbf{z} - \dot{\mathbf{x}}_d) - \mathbf{K} \text{sgn}(\mathbf{S})] \quad (18)$$

滑模控制有一个非连续项 $\mathbf{K} \text{sgn}(\mathbf{S})$, 当选取合适的 \mathbf{K} 时它将使得滑模控制对于干扰具有鲁棒性。但是, 另一方面, 该项也带来了抖震, 为了减小抖震, 使用边界层方法, 即用 $s_i / (|s_i| + \varepsilon)$ 代替 $\text{sgn}(s_i)$, 这里 ε 代表边界层宽度。

4 机械臂抓取试验系统的相关验证

4.1 抓取系统简介

本系统设计方案考虑到双目摄像机能够从不同角度实时获取目标物体的图像信息, 以便可以较为容易地得到目标的三维坐标, 故而采用双目视觉系统。从系统工作的实时性考虑, 采用固定摄像机的结构, 减少系统的计算压力, 提高相应速度, 并且将双目相机固定在机械臂工作台的一侧, 并且从安全稳定性考虑, 将其放在机械臂操作空间之外。搭建的基于视觉的机械臂控制系统如图 5 所示, 主要由两大核心部分组成: 视觉目标识别系统和机械臂控制系统。

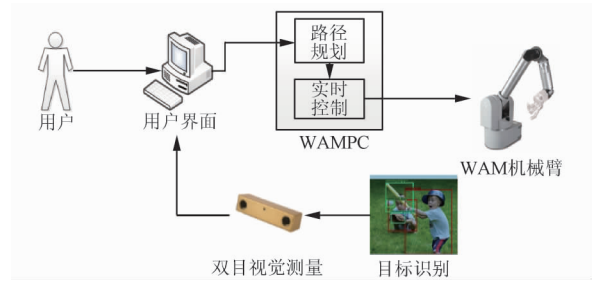


图 5 抓取系统方案
Fig. 5 Scheme of capture system

在操作机械臂前提前利用深度学习算法训练好多目标识别模型, 采集需要辨识的目标图像, 构建数据集, 利用数据集对深度卷积神经网络训练, 并通过正则化及非凸优化调整网络权值, 提高网络泛化能力。系统上电后, 双目摄像头采集视频信号将图像信息传送至上位机, 上位机对采集的图像信号分析, 利用已经训练好的深度卷积神经网络实现目标的识别, 通过双目视觉系统得到所需的目标物体的位置信息, 上位机在得到目标位置信息后通过机械臂运动学逆解或者轨迹优化算法计算得到机械手欲达到这个位置每个关节需要转动的角度, 根据控制算法实时计算控制器输出, 通过上位机软件发出相应的控制指令驱动机械臂控制系统的控制器使各关节转动相应的角度, 以使机械手准确抓取指定物体。

下面主要验证基于深度神经网络的多类目标识别与定位、机械臂的滑模控制以及避障抓取目标。

4.2 目标识别试验

考虑 NVIDIA 公司的 Jetson TX1(图 6) 具有 Teraflops 级的浮点计算能力和强大的软件堆栈以及超低能耗, 足以负载实时目标识别的应用和开发, 因此本文所用基于深度学习的 YOLO 算法使用此板卡实现。

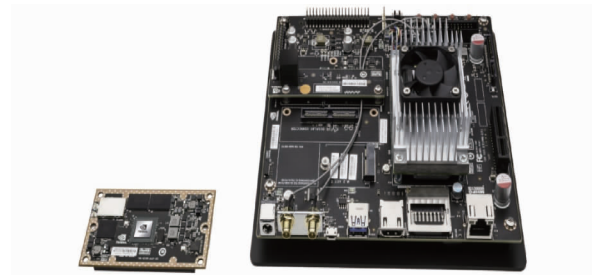


图 6 TX1 的 GPU 板卡及开发者套件^[12]
Fig. 6 TX1 GPU board and Developer Suite^[12]

YOLO 所依赖的深度学习框架为 darknet, 模型训练所需的超参数设置如下: 1) 基础学习率 learning_rate 为 0.001; 2) 学习率的衰减策略 policy 为 “steps”, 步长 stepsize 为 40000; 3) 学习率的变化比率 scales 为 0.1; 4) 最大迭代次数 max_batche 为 80200; 5) 动量 momentum 为 0.9; 6) 权重衰减项 decay 为 0.0005。

识别算法神经网络结构如表 1 所示。

表 1 网络结构
Table 1 Network structure

Type	Filters	Size / Stride	Output
Convolutional	32	3* 3	224* 224
Maxpool		2* 2/2	112* 112
Convolutional	64	3* 3	112* 112
Maxpool		2* 2/2	56* 56
Convolutional	128	3* 3	56* 56
Convolutional	64	1* 1	56* 56
Convolutional	128	3* 3	56* 56
Maxpool		2* 2/2	28* 28
Convolutional	256	3* 3	28* 28
Convolutional	128	1* 1	28* 28
Convolutional	256	3* 3	28* 28
Maxpool		2* 2/2	14* 14
Convolutional	512	3* 3	14* 14
Convolutional	256	1* 1	14* 14
Convolutional	512	3* 3	14* 14
Convolutional	256	1* 1	14* 14
Convolutional	512	3* 3	14* 14
Maxpool		2* 2/2	7* 7
Convolutional	1024	3* 3	7* 7
Convolutional	512	1* 1	7* 7
Convolutional	1024	3* 3	7* 7
Convolutional	512	1* 1	7* 7
Convolutional	1024	3* 3	7* 7
Convolutional	1000	1* 1	7* 7
Avgpool		Global	1000
Softmax			

训练过程如下:

1) 首先是数据的准备, 目前验证目标识别算法可以通过一些标准数据集, 如 Kitti 数据集。如果数据集不标准, 需要将数据集转化为标准数据

集的形式, 常见的有 VOC 格式。

2) 然后通过 YOLO 算法训练神经网络连接权值, 生成权值文件, 这个训练过程可以通过 TX1 的 GPU 加速。

通过图 7 所示结果可以看出, YOLO 算法识别出了包括显示屏、椅子、杯子、人等对象, 对多类目标实现了较好的识别与较精准的定位。



图 7 测试结果
Fig. 7 Test results

4.3 滑模控制验证

针对本文所用滑模控制律, 首先设计一个仿真算例。该示例用于跟踪一个参考轨迹 $\theta_d(t)$, 其表达式为: $\theta_d(t) = [0.1 + 0.5\sin 0.6t - 1.5708 \ 0 \ 3.14]^T$, 设初始条件为 $\theta(0) = [0 \ -1.5708 \ 0 \ 3.14]^T$, 控制律的相关参数为: $\lambda = 20I_{4 \times 4}$ 、 $K = 25I_{4 \times 4}$ 、 $\varepsilon = [0.5 \ 0.5 \ 0.5 \ 0.5]^T$, 仿真的采样时间为 0.001 s。得到图 8 所示结果。

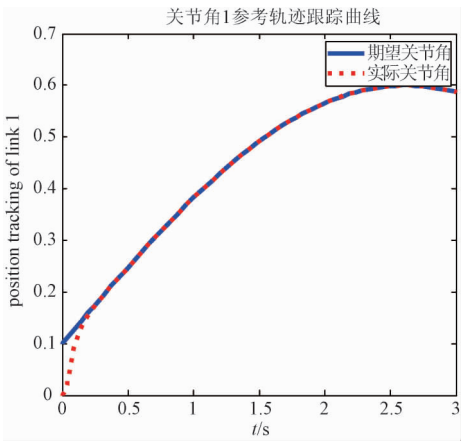


图 8 关节角 1 参考轨迹跟踪仿真结果
Fig. 8 Simulation results of joint angle 1 in trajectory tracking experiments

由上图可知滑模控制在跟踪正弦轨迹时响应速度很快, 在 0.21 s 时误差即在 2%。控制力矩在初始时刻较大, 随着实际关节轨迹趋近于期望

关节轨迹,控制力矩将趋于0。

通过该仿真试验可知该控制律控制精度满足控制要求,下面完成实物试验。

在 WAM 机械臂的 Ubuntu12.04 实时操作系统上用 C++ 编写控制程序实现上述控制算法。由 PC 机通过 Socket 协议向机械臂发出控制指令,从而实现对机械臂的实时控制。本试验用于跟踪一个参考轨迹 $\theta_d(t)$, 其表达式为:

$$\theta_d(t) = \begin{bmatrix} 0.5 \sin 0.6t \\ 0.5 \sin 0.6t - 0.7854 \\ 0.5 \sin 0.6t \\ 0.5 \sin 0.6t + 1.5708 \end{bmatrix}, \text{ 初始条件为}$$

$\theta(0) = [0.1 \quad -0.6854 \quad 0.1 \quad 1.6708]^T$, 控制律的相关参数设为 $\lambda = 70I_{4 \times 4}$ 、 $K = 50I_{4 \times 4}$ 、 $\varepsilon = [0.5 \quad 0.5 \quad 0.5 \quad 0.5]^T$ 。得到如图 9 所示结果。

由多关节联动试验可知,滑模控制可以对多关节同时进行高精度跟踪。

4.4 避障抓取试验

如图 10 ~ 11 所示,在完成目标识别与定位后,通过期望轨迹计算得到有效关节轨迹,控制机械臂跟踪该路径,同时还需要避免各关节与障碍物的碰撞,在空间足够的地方,张开 Bhand-280 机械手,然后在接近目标的位置处实施抓取动作。在目标位置完成抓取后,机械臂需要收回到基座,这时同样需要避免各关节与障碍物的碰撞。最终的抓取试验是成功的。

5 结论

本文通过仿真验证及实物抓取验证,实现了基于深度学习的的目标快速精确识别,能在小样本环境下快速给出目标的分类及定位信息,为抓取规划控制提供了前提。同时从实际机械臂运行精度可以得出所设计滑模控制律具有使用可靠性。但目前实验环境相对单一,验证任务相对简单,后续将在本文的基础上在高动态环境下复杂多目标识别,协同抓取等方面展开进一步研究。

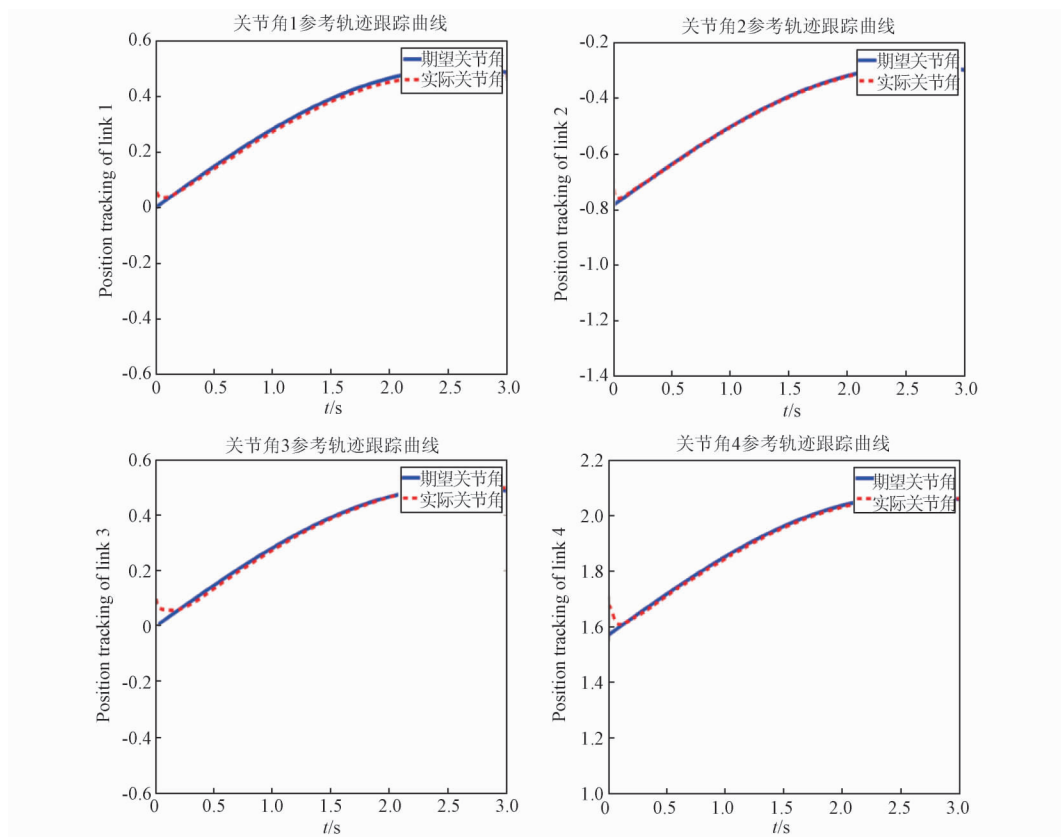


图9 多关节联动各关节角试验结果
Fig.9 Experimental results of various joint angles of multi joints

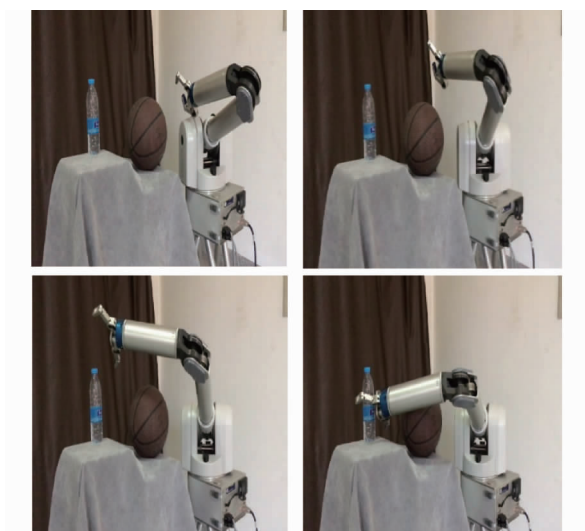


图 10 机械臂避障抓取目标过程
Fig. 10 The process of obstacle avoidance and target capture by the robotic arm

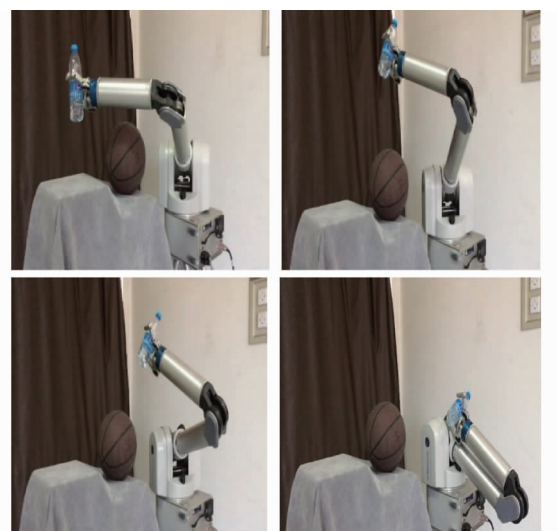


图 11 机械臂抓取目标后返回避障过程
Fig. 11 The process of obstacle avoidance and returning after target grasping of the robotic arm

参考文献(References)

- [1] 邓桦. 机械臂空间目标视觉抓取的研究[D]. 哈尔滨: 哈尔滨工业大学, 2013.
Deng Hua. Research on Visual Capture of Robot Arm with Spatial Target [D]. Harbin: Harbin Institute of Technology, 2013. (in Chinese)
- [2] 于灏洋. 基于双目视觉的伺服机械臂运动控制研究[D]. 大连: 大连理工大学, 2016.
Yu Haoyang. Research on Motion Control of Servo Arm based on Binocular Vision [D]. Dalian: Dalian University of Technology, 2016. (in Chinese)
- [3] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 779-788.
- [4] Redmon J, Farhadi A. YOLO9000: better, faster, stronger [J]. arXiv preprint arXiv: 1612.08242, 2016.
- [5] The WAMTM Arm [EB/OL]. <https://www.barrett.com/wam-arm/>.
- [6] 孟韶南, 梁雁冰, 师恒. 基于 ROS 平台的六自由度机械臂运动规划[J]. 上海交通大学学报, 2016(S1): 94-97.
Meng Shaonan, Liang Yanbing, Shi Heng. Motion planning of six degree of freedom robot arm based on ROS platform [J]. Journal of Shanghai Jiao Tong University, 2016(S1): 94-97. (in Chinese)
- [7] 刘鹏, 王强, 张伟. 机械臂动力学建模及控制仿真[J]. 自动化与仪表, 2017(03): 9-12.
- [8] 邵兵, 吴洪涛, 程世利, 等. 基于李群李代数的主被动关节机器人动力学及控制[J]. 中国机械工程, 2010(3): 253-257.
Shao Bing, Wu Hongtao, Cheng Shili, et al. The main passive joint robot dynamics and control based on the theory of Lie groups and Lie algebras [J]. China Mechanical Engineering, 2010(3): 253-257. (in Chinese)
- [9] 邵兵. 基于李群李代数的统一开闭环机械多体系统递推动力学研究[D]. 南京: 南京航空航天大学, 2010.
Shao Bing. Research on Recursive Mechanical Mechanics of Unified Open Closed Loop Mechanical Multibody Systems based on the Theory of Lie Groups and Lie Algebras [D]. Nanjing: Nanjing University of Aeronautics & Astronautics, 2010. (in Chinese)
- [10] Young K D, Utkin V I, Ozguner U. A control engineer's guide to sliding mode control [J]. IEEE Transactions on Control Systems Technology, 1999, 7(3): 328-342.
- [11] Yong F, Yu X, Man Z. Non-singular terminal sliding mode control of rigid manipulators [J]. Automatica, 2002, 38(12): 2159-2167.
- [12] Jetson TX1 Module [EB/OL]. <https://developer.nvidia.com/embedded/buy/jetson-tx1>.

(责任编辑: 龙晋伟)