

## 适用于非合作目标捕获的轻量级位姿估计网络

蒋 明<sup>1</sup>, 陈 雨<sup>1</sup>, 周青华<sup>2</sup>, 袁 媛<sup>1</sup>, 何世琼<sup>1</sup>

(1. 四川大学 电子信息学院, 成都 610065; 2. 四川大学 空天科学与工程学院, 成都 610065)

**摘 要:** 空间非合作目标的增多导致太空安全受到严重威胁, 对非合作目标进行捕获回收具有维护空间安全、节约资源等现实意义。非合作目标捕获回收需要进行位姿估计, 而目前在硬件资源有限的航天器平台上, 现有的大多数非合作目标位姿估计算法无法同时满足及时性和准确性的要求。设计一种超轻量级目标检测网络 YOLO-GhostECA, 利用 GhostBottleneck 网络减少特征图冗余, 并使用高效注意力机制提取核心特征图, 以降低模型参数, 在提升运算速度的同时保证精度水平几乎不下降。根据 YOLO-GhostECA 网络的检测结果粗略估计姿态, 以协助机械臂更加合理地执行智能捕获任务, 解决 2D 识别算法无法检测出物体姿态的问题。在 7 自由度冗余机械臂上开展的空间非合作目标捕获地面模拟的实验结果表明, 与 YOLOv5s 网络相比, 该网络模型大小减小了 80.4%, 运算复杂度降低了 78.9%, 而精度基本保持不变, 可准确快速地对非合作目标进行位姿估计, 能够引导机器人成功捕获非合作物体。

**关键词:** 非合作目标; 神经网络; 目标检测; 智能捕获; 轻量级; 位姿估计

开放科学(资源服务)标志码(OSID):



**中文引用格式:** 蒋明, 陈雨, 周青华, 等. 适用于非合作目标捕获的轻量级位姿估计网络[J]. 计算机工程, 2022, 48(6): 235-242.

**英文引用格式:** JIANG M, CHEN Y, ZHOU Q H, et al. Lightweight pose estimation network for non-cooperative target acquisition[J]. Computer Engineering, 2022, 48(6): 235-242.

## Lightweight Pose Estimation Network for Non-Cooperative Target Acquisition

JIANG Ming<sup>1</sup>, CHEN Yu<sup>1</sup>, ZHOU Qinghua<sup>2</sup>, YUAN Yuan<sup>1</sup>, HE Shiqiong<sup>1</sup>

(1. School of Electronic Information, Sichuan University, Chengdu 610065, China;

2. School of Aeronautics and Astronautics, Sichuan University, Chengdu 610065, China)

**[Abstract]** The increase in non-cooperative targets in space has led to a serious threat to space security. Therefore, the capture and recovery of non-cooperative targets have practical significance in maintaining space security and resource-saving. Non-cooperative target acquisition and recovery require pose estimation. On a spacecraft platform with limited hardware resources, most existing non-cooperative target pose estimation algorithms cannot simultaneously meet the requirements of timeliness and accuracy. An ultra-lightweight target detection network YOLO-GhostECA is proposed. The Ghostbottleneck network reduces the redundancy of the feature map, and an efficient attention mechanism extracted the core feature map to reduce the model parameters and ensure that the accuracy level was not compromised while improving the operation speed. Using the detection results of the YOLO-GhostECA network to estimate the attitude to assist the manipulator in a sensibly intelligent acquisition task and solve the problem that the 2D recognition algorithm cannot detect the object attitude. The experimental results of the ground simulation of space non-cooperative target acquisition on a 7-DOF redundant manipulator show that compared with the algorithm based on the YOLOv5s network, the model size was reduced by 80.4%, the computational complexity was reduced by 78.9%, and the accuracy remained constant. It accurately and quickly estimates the pose of non-cooperative targets and guides the robot to successfully capture non-cooperative objects.

**[Key words]** non-cooperative target; neural network; target detection; intelligent capture; lightweight; pose estimation

DOI: 10.19678/j.issn.1000-3428.0060882

**基金项目:** 国家自然科学基金面上项目(51875373); 四川省科技计划项目(2019YJ0093)。

**作者简介:** 蒋 明(1997—), 男, 硕士研究生, 主研方向为机器视觉、智能机器人; 陈 雨(通信作者), 副教授、博士; 周青华, 教授、博士; 袁 媛、何世琼, 硕士研究生。

**收稿日期:** 2021-02-18 **修回日期:** 2021-06-14 **E-mail:** jmllearn@163.com

## 0 概述

随着人类开发太空进程的加快,大量卫星、火箭等航天器被发射到太空中执行各项任务,但由于航天器可能会面临意外故障、燃料耗尽等问题,导致这些航天器由合作目标转化成非合作目标,威胁着太空安全。因此,回收这些非合作目标具有维护空间安全及节约资源的现实意义<sup>[1]</sup>。研究人员提出在轨服务技术<sup>[2]</sup>,利用空间特种机器人来完成在轨维修、清理空间碎片等高危险性空间任务,其中,非合作目标捕获是在轨服务技术的一项关键技术。在特种机器人对非合作目标实施在轨抓捕之前,需要对目标航天器进行识别与定位,为下一步捕获任务做准备。但是在太空环境中,相机对太阳光照等外界信息非常不敏感,且机器人面临的是未知的非结构化环境,传统的目标检测算法需要人为提取特征,泛化能力较差,在非合作目标检测任务上表现不佳。

自2012年卷积神经网络CNN被应用于图像处理领域并一举夺得ImageNet图像识别比赛冠军之后<sup>[3]</sup>,基于CNN的目标检测算法进入了快速发展阶段。由于其具有自动提取特征、泛化能力强、稳定性好等优点,因此在非结构化环境中表现了巨大的优势。基于CNN的目标检测算法大致可以分成2大类型:第1类是以区域推荐法获取不同尺寸的二阶段网络,其精度较高,但由于需要产生大量候选框,所以速度慢、实时性较差,主要代表算法有R-CNN<sup>[4]</sup>、Fast-RCNN<sup>[5]</sup>、Mask-RCNN<sup>[6]</sup>等;第2类是一阶段网络,直接端到端进行预测,不需要区域推荐,直接回归产生物体的类别概率和位置信息,因此速度快,模型较小,但精度相对较低,代表性的算法有YOLO<sup>[7]</sup>、SSD<sup>[8]</sup>。

大部分空间非合作目标的运动规律是未知的,甚至很大可能是高速运动的,且均具有一定的非合作性。受机载处理器计算能力的限制,非合作目标的复杂识别算法无法在航天器上独立完成<sup>[9]</sup>,图像处理算法的设计和实现面临着准确性、及时性和小型化的矛盾,因此对空间机器人非合作目标抓捕任务的目标检测算法准确性和实时性均提出了较高要求<sup>[10]</sup>。此外,传统的2D目标检测只能得到空间位置,无法估计物体的姿态,不太适用于机械臂捕获操作。针对这些问题,YU等<sup>[11]</sup>提出利用检测目标矩形框长宽比进行抓取姿态的粗略估计,但没有考虑目标物体位置对捕获姿态的影响,且估算姿态太过简单,很多估计情况不符合现实场景。

本文基于YOLO设计一种YOLO-GhostECA网络,通过GhostBottleneck网络减少特征图冗余,并利用高效通道注意力模块ECA提取核心特征图,从而在保持检测精度不下降的同时,提升检测速度及降低模型大小。此外,提出一种定性估计非合作目标

抓取姿态的方法,以弥补传统2D目标检测算法无法估计抓取姿态的缺陷。

## 1 轻量级目标检测识别

### 1.1 YOLOv5s目标检测网络

由于YOLO网络兼顾实时性和准确性,因此成为了工业界落地最好的目标检测网络之一,在这之后相继提出了YOLOv2<sup>[12]</sup>、YOLOv3<sup>[13]</sup>、YOLOv4<sup>[14]</sup>以及各种YOLO网络的变种。GLENN等<sup>[15]</sup>开发出的YOLOv5,成为YOLO网络实现版本中最优秀的网络之一。YOLOv5s网络由骨干网络、瓶颈层、检测层3大模块组成,如图1所示。

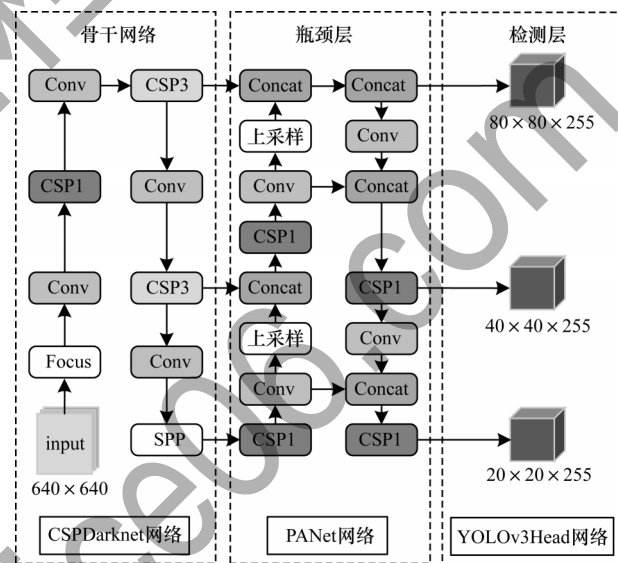


图1 YOLOv5s网络模型结构

Fig.1 Network model structure of YOLOv5s

骨干网络的主要作用是提取特征,由Focus<sup>[16]</sup>、CSPDarknet53<sup>[14]</sup>、SPP<sup>[17]</sup>组成。其中Focus为切片操作,能够替换卷积操作,减少卷积所造成的特征信息损失。CSPDarknet53主要进行网络特征提取,该网络由跨层局部连接CSP<sup>[18]</sup>构建而成,可以增加网络提取特征的能力,降低计算瓶颈和内存成本。SPP是空间特征金字塔池化结构,此结构能扩大感受野,实现局部和全局特征融合,丰富特征图的信息。

瓶颈层用来融合不同尺寸特征图及提取高层语义特征,由PANet网络组成。PANet网络可以将特征图自上而下以及自下而上进行特征融合,以减少信息损失<sup>[19]</sup>。

检测层将瓶颈层产生的高层语义特征用于分类和回归出物体的类别和位置,由YOLOv3Head<sup>[13]</sup>组成。YOLOv3Head有3个检测头,分别可以检测大、中、小物体,很好地克服了一阶段检测小物体精度低的缺点。

### 1.2 YOLO-GhostECA轻量级目标检测网络

传统的深度卷积网络模型一般使用大量常规卷积对图像进行特征提取,但这需要消耗大量的算力。2020年1月,华为诺亚方舟实验室HAN等<sup>[20]</sup>在观测常规卷积中间特征图时发现很多通道的特征图十分

相似,这说明利用常规卷积提取出来的特征图在一定程度上是存在冗余的。他们猜想并不需要进行完整的卷积运算得到需要的特征图,而只需要用常规卷积去获取一部分特征图即可,因此提出了 GhostNet 网络。该网络可以用更少的参数来生成更

多特征,能够减少特征图冗余,从而达到减少网络模型大小和计算量的目的。GhostNet 网络的相关结构原理如图 2 所示,其中,图 2(b)的 GhostBottleneck 由图 2(a)的 GhostModule 网络构成, GhostNet 网络由图 2(b)的 GhostBottleneck 网络构成。

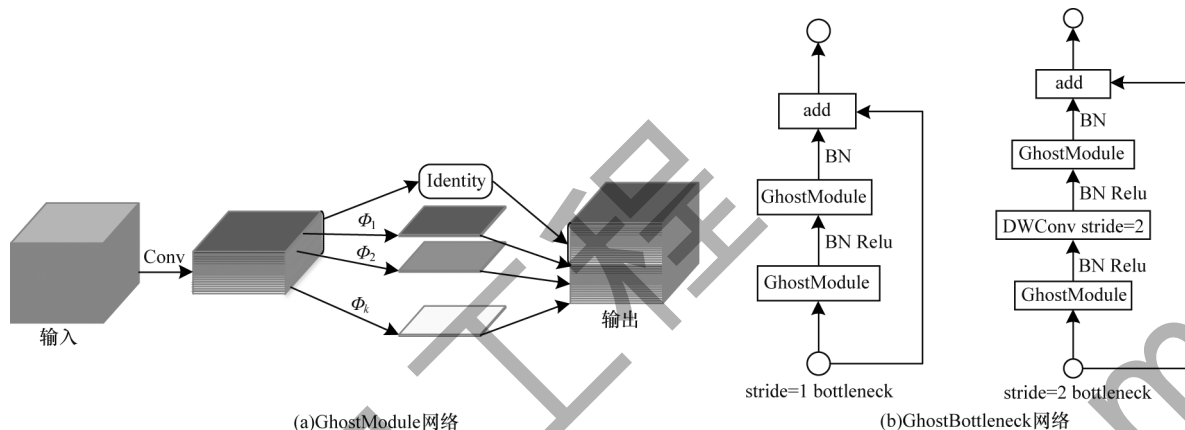


图2 GhostModule 网络及 GhostBottleneck 网络的结构

Fig.2 Network structure of GhostModule and GhostBottleneck

GhostModule 网络的具体实现原理如图 2(a)所示,与常规卷积不同的是它只需要利用常规卷积运算得到半数通道的特征图,使用一系列廉价的线性运算对这半数通道的特征图进行处理,并生成幻影特征图,最后将生成的幻影特征图与原始卷积得到的半数通道特征图合并,构成与常规卷积相同通道数的特征图,此特征图可以近似于原始的完整特征图,但是计算量和参数量却减少了很多。此外, HAN 等利用 GhostModule 网络设计了图 2(b)中 2 种 GhostBottleneck 网络结构,其中 stride=1 的 GhostBottleneck 网络与步长为 1 的 CSPConv 网络效果相同, stride=2 的 GhostBottleneck 与步长为 2 的 CSPConv 网络效果相同,因此 GhostBottle 网络可以替换常规卷积在网络模型中进行特征提取。本文利用 GhostBottleneck 网络替换 CSPBottleneck 网络,并重新构建出了 GhostDarknet 网络作为 YOLO-Ghost 的主干网络。为进一步提升精度,本文引入了高效通道注意力(Efficient Channel Attention, ECA)<sup>[21]</sup>模块,并将 ECA 应用于 PANet 网络中,形成具有注意力机制的 PANet 网络,即 ECA-PANet 结构。该结构能够加强特征融合,在获得更多想要的图像特征的同时,忽略不需要的特征,与其他通道注意力模块相比, ECA 模块可以在不引入多余参数进行计算的情况下,提升模型的精度。

以 YOLOv5s 网络为基础,结合上述提出的几个新型模块构建出了新型轻量级目标检测网络 YOLO-GhostECA,该网络可以利用少量的参数获取更多需要的特征图,其结构如图 3 所示。

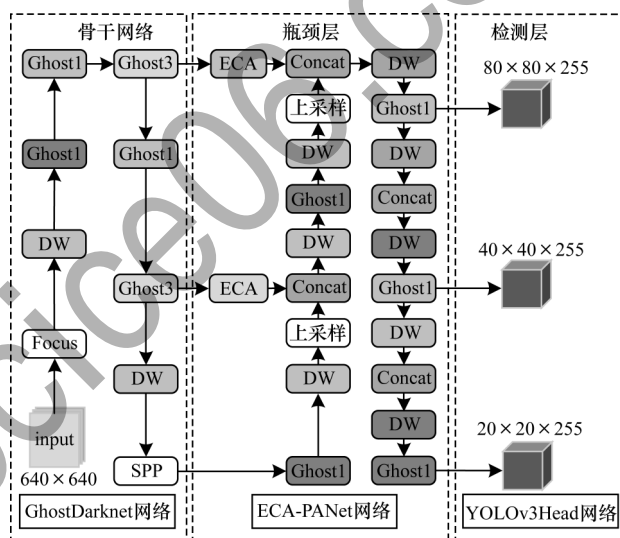


图3 YOLO-GhostECA 网络结构

Fig.3 Network structure of YOLO-GhostECA

### 1.3 YOLO-GhostECA 网络的损失函数

原 YOLOv5 中 IoU 损失函数使用的是 GIoU-loss,但其只考虑到了重叠区域和非重叠区域,没有考虑矩形框对 IoU 的影响。因此 YOLO-GhostECA 使用 Complete-IoU(CIoU)<sup>[22]</sup>损失函数,该损失函数同时考虑到了矩形框的长宽比、目标与 anchor 之间的距离、重叠率、尺度等因素,这可以进一步提高模型的性能,加快损失函数收敛。CIoU 损失函数如式(1)所示:

$$L_{CIoU} = 1 - I_{IoU} + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \quad (1)$$

其中:  $\alpha$  是权重函数。

计算方法如式(2)所示:



$$\alpha = \frac{v}{(1 - I_{\text{IoU}}) + v} \quad (2)$$

其中: $v$ 是度量长宽比的相似性。

$v$ 的计算公式如式(3)所示:

$$v = \frac{4}{\pi^2} \left( \arctan \frac{w^{\text{gt}}}{h^{\text{gt}}} - \arctan \frac{w}{h} \right)^2 \quad (3)$$

其中: $\rho(b, b^{\text{gt}})$ 代表计算预测框中心点 $b$ 和真实框中心点 $b^{\text{gt}}$ 的欧氏距离; $c$ 表示能够同时包含预测框和真实框的最小闭包区间的对角线距离。

## 2 空间定位与姿态估计

### 2.1 空间定位

上述目标检测算法只能获得非合作目标在图像坐标系中的位置,但是机械臂抓取目标需要基于自身基座坐标系下的三维点信息。为完成抓取任务,还需建立相机成像模型,如图4所示,其中: $O_c X_c Y_c Z_c$ 是相机坐标系; $O_p X_p Y_p$ 是图像坐标系; $O_w X_w Y_w Z_w$ 是世界坐标系; $P(x_w, y_w, z_w)$ 是物体在世界坐标系下的坐标; $P(u, v)$ 是物体在像素坐标系下的坐标。

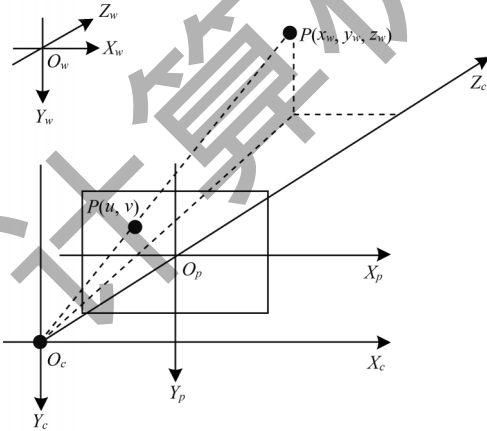


图4 相机成像模型

Fig.4 Camera imaging model

借助相机成像模型可建立目标检测得到的二维图像坐标和机械臂抓取的三维世界坐标之间的映射关系,如式(4)所示:

$$z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{f}{dx} & 0 & u_0 & 0 \\ 0 & \frac{f}{dy} & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = M_1 M_2 \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (4)$$

其中: $M_1$ 是摄像机内部参数矩阵,用来将像素坐标转化到相机坐标系下,是一个由二维转三维的过程,可由相机标定张正友棋盘格标定法<sup>[23]</sup>得到; $M_2$ 是摄像机外部参数矩阵,能够将相机坐标转化到世界坐标系下(本文的世界坐标系定义在机械臂的基础坐标系处),一般由手眼标定<sup>[24]</sup>得到; $z_c$ 是深度值,由 Intel Rense D435i 深度相机根据结构光与双目立体视觉融合的混合测距技术实时计算得到。当知道以上相关参数和

目标检测得到的物体像素坐标后,即可使用式(4)完成对目标物体的空间定位。

### 2.2 姿态估计

空间定位可以得到目标物体的三维空间坐标,但为了合理引导机器人捕获目标物体,姿态估计是一个重要的过程。事实上,目标检测得到的结果数据中已经包含了一些物体的姿态信息,例如当物体呈现不同姿态时,其检测得到的矩形框尺寸比例将呈现不同的数值。为进一步利用目标检测的结果,本文提出一种通过目标检测得到的非合作目标信息( $u, v, w, h, \text{class}$ )来粗略预估机械臂抓取姿态的新型方法。该方法将抓取姿态 $\Phi$ 大致分成了6种,即中正抓姿态 $\Phi_A$ 、中偏抓姿态 $\Phi_B$ 、上抓姿态 $\Phi_C$ 、下抓姿态 $\Phi_D$ 、左抓姿态 $\Phi_E$ 和右抓姿态 $\Phi_F$ 。估算的6种抓取姿态的姿态参数如表1所示。

表1 6种抓取姿态的相关参数(绕固定轴旋转)

Table 1 6 parameters related to grasping attitude (rotating around a fixed axis) (°)

姿态名	绕X轴旋转 姿态角W	绕Y轴旋转 姿态角P	绕Z轴旋转 姿态角R
中正抓姿态 $\Phi_A$	180	-90	0
中偏抓姿态 $\Phi_B$	90	0	90
上抓姿态 $\Phi_C$	0	0	0
下抓姿态 $\Phi_D$	180	0	0
左抓姿态 $\Phi_E$	90	0	0
右抓姿态 $\Phi_F$	-90	0	0

目标物体的抓取姿态 $\Phi$ 可以由式(5)和式(6)估计确定:

$$r = \frac{O_w}{O_h} \quad (5)$$

$$\Phi = \begin{cases} \Phi_A, r > r_1 \\ \Phi_B, r < \frac{1}{r_1} \\ \Phi_C, r_2 < r < r_1, v > V_s \\ \Phi_D, r_2 < r < r_1, v \leq V_s \\ \Phi_E, \frac{1}{r_1} < r < r_2, u > U_s \\ \Phi_F, \frac{1}{r_1} < r < r_2, u \leq U_s \end{cases} \quad (6)$$

式(5)是计算矩形宽高比例的公式,其中: $O_w$ 为矩形框的宽; $O_h$ 为矩形框的高。式(6)是计算目标物体抓取姿态的公式,其中: $u, v$ 为目标检测得到的物体像素坐标; $r_1, r_2$ 为常数,可以根据实际物体class的长宽比例进行粗略计算,多个物体即有多个值; $V_s$ 用来确定抓取的上下姿态; $U_s$ 用来确定抓取的水平方向的姿态,这2个参数可以将笛卡尔空间划分成4个区域,根据目标检测像素点的位置可判定物体处在4个区域中的位置, $U_s, V_s$ 的值取决于相机与机械臂末端初始位置的空间几何关系,即机械臂在初始位置时在相机像素坐标系下的映射点坐标。

6种抓取姿态的估计示意图如图5所示。

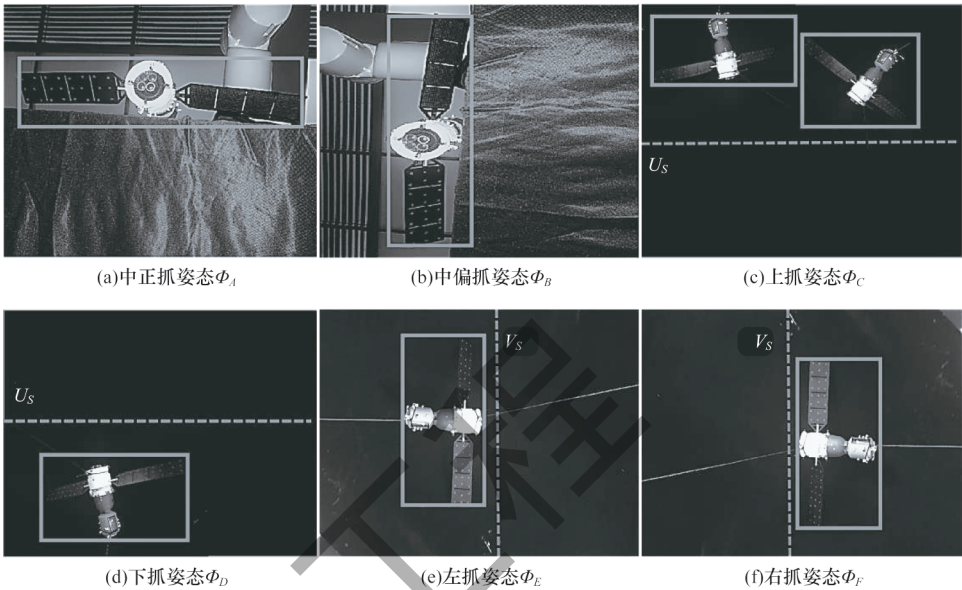


图5  6种抓取姿态的估计示意图

Fig.5  Schematic diagram of estimation of six grasping postures

3  实验结果与分析

3.1  实验仿真系统

本文利用7自由度机械臂等设备搭建了地面仿真实验平台。实验中先将航天器模型用细绳系紧,悬挂于空中,并将实验室的灯熄灭,只留1个补光灯,给非合作目标补光,以此来模拟太空环境暗黑失

重环境。实验仿真流程如图6所示,先使用YOLO-GhostECA进行非合作目标识别,再根据目标检测结果使用第2节的方法估计物体的位姿,最后将估计的目标物体位姿信息 $(x,y,z,W,P,R)$ 传送给机械臂的控制器,机械臂控制器控制电机进行相应驱动,即可对目标物体进行捕获操作。本文所有实验均在表2所示实验条件下进行。

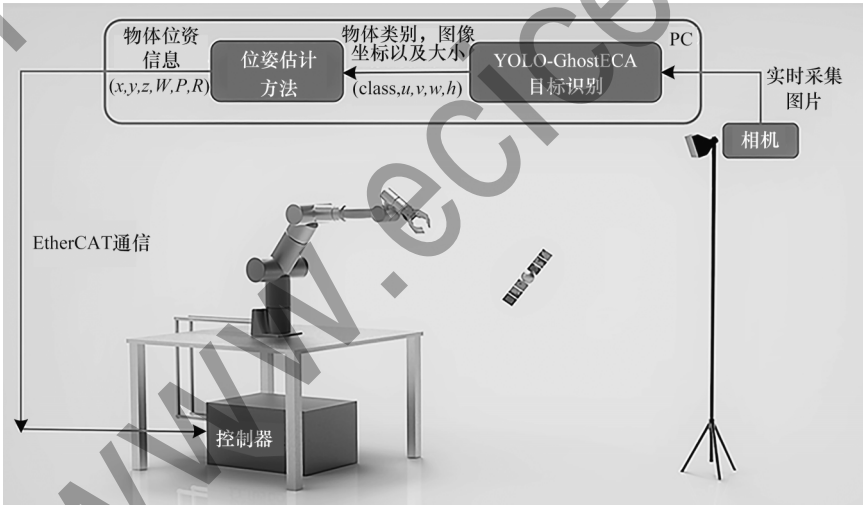


图6  实验仿真流程

Fig.6  Procedure of experimental simulation

表2  实验参数

Table 2  Experimental parameters	
实验设备	型号参数
CPU	Intel i7 7800
GPU	RTX2080super
深度相机	Intel RealSense D435i
编译器	pycharm2020.2
深度学习环境	pytorch1.6、cuda10.1
图片大小/像素	640×640
机械臂	7自由度冗余机械臂(同川TCR织女号)

3.2  非合作目标识别

非合作目标识别的具体步骤如下:

**步骤1  数据集的制作。**由于目前没有非合作目标航天器的公有数据集,因此在实际中对太空环境黑暗且失重环境进行现实模拟,采集拍摄387张多姿态多场景的航天器模型图片,并利用采集的图像数据进行联合训练,部分数据集如图7所示。



图7 非合作目标数据集示例

Fig.7 Examples of non-cooperative target dataset

**步骤2 训练策略。**实验中使用了多种训练策略来提升网络的性能。

1) anchor 大小。利用 K-mean 聚类算法找出数据集中较好的 9 个 anchor 尺寸,有助于提高平均精度均值(mean Average Precision, mAP)。

2) 超参数演进。利用基因遗传算法训练 200 代,找到较好的超参数后,利用超参数演进得来的超参数对网络进行训练,使网络能够迅速得到全局最优解,让模型达到较好的效果。

3) 学习率  $l_r$ 。利用类余弦退火进行学习率学习,减少陷入局部最优解的情况发生。

4) 数据增强操作。在训练过程中对数据进行随机裁剪、放大等操作来增大数据集,使每一个 epoch 训练的图都不是同一张图,在数据集有限的情况下有效提高模型的泛化能力。

**步骤3 性能测试。**通过上面的策略进行模型训练,完成后在模型大小、运行速度、模型参数量(即

过程中梯度更新数量)、模型运算复杂度(即每秒浮点数运算次数)等多方面对网络模型性能进行分析评估,其结果如图 8 所示。由图 8(a)可知,YOLOv5-GhostECA 模型大小仅 2.9 MB,相较于 YOLOv5s 减少了 80.4%。图 8(b)显示 YOLO-GhostECA 在 GPU 上运行时间仅 2.7 ms,即帧率高达 370 frame/s, GPU 运行速度相对于 YOLOv5s 提升了 49.1%。图 8(c)显示 YOLOv5-GhostECA 模型在 CPU 上的运行时间为 320.6 ms,相对于 YOLOv5s 的 460 ms 来讲,运行时间减少了 30.3%。图 8(d)显示 YOLO-GhostECA 的 mAP 为 0.896,依然保持在较高值,相对于 YOLOv5s 的 mAP 仅下降了 0.2%。由图 8(e)可知 YOLOv5s 的模型参数大小为 7.468 MB,而 YOLO-Ghost 和 YOLO-GhostECA 均只有 1.313 MB,新提出的 YOLO-Ghost 相较于 YOLOv5s 参数量减少了 82.4%,而融合 ECA 注意力机制的 YOLO-GhostECA 参数量相比于 YOLO-Ghost 保持不变。图 8(f)所示为模型的运算复杂度,YOLOv5s 为 17.5 GFLOPs, YOLO-Ghost 和 YOLO-GhostECA 均为 3.7 GFLOPs,可以看出 YOLO-Ghost 相比于 YOLOv5s 运算复杂度降低了 78.9%,且融合的 ECA 模块并没有增加运算复杂度。综上分析,本文提出的 YOLO-GhostECA 模型在 GhostDarknet 基础上结合高效通道注意力模块 ECA 的方法在保持精度的同时,使模型的参数量和运算复杂度均大幅降低,且减小了模型尺寸,提升了运行速度。

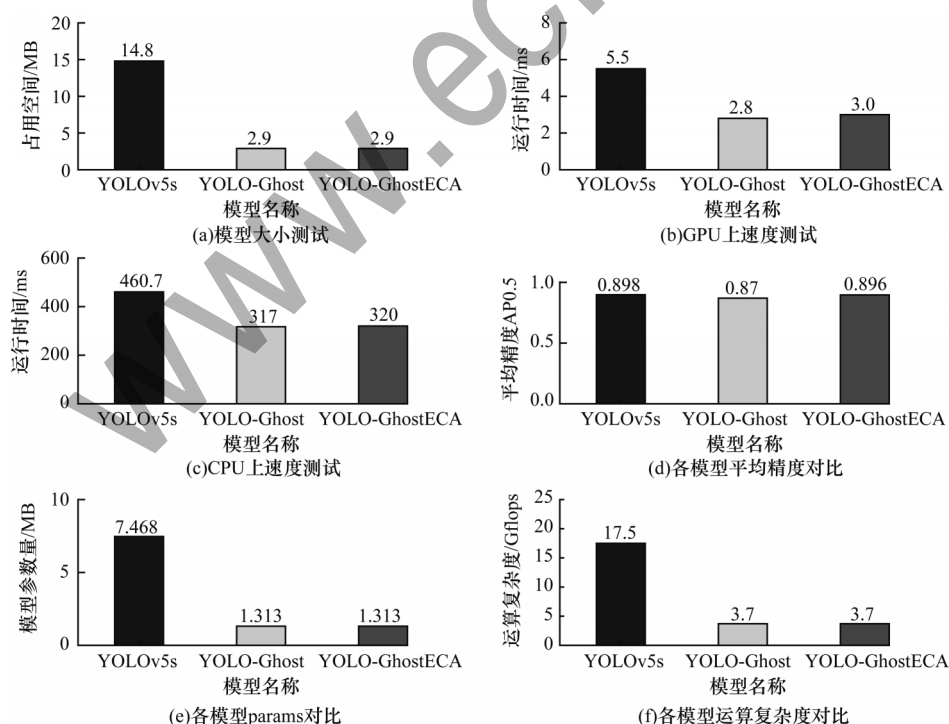


图8 不同模型的性能参数对比

Fig.8 Comparison of performance parameters of different models



步骤4 模型效果测试。利用训练好的YOLO-GhostECA进行非合作目标识别实验,实验结果如图9所示。可以看出,该模型能够准确地识别和定位出非合作目标物体,且在多姿态多场景下,其预测准度分数均保持较高水平。

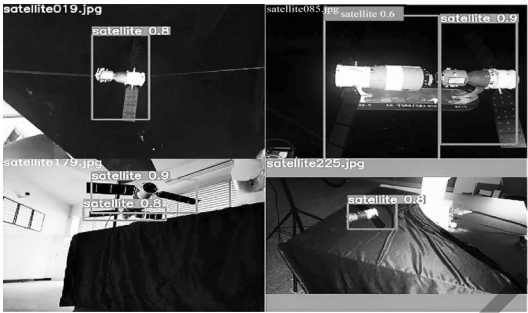


图9 本文网络模型测试结果  
Fig.9 Test results of network model in this paper

表3 非合作目标位姿估计的结果  
Table 3 Results of pose estimation of non-cooperative targets

序号	像素坐标(u,v)/像素	笛卡尔坐标(x,y,z)/mm	矩形宽高(w,h)/像素	宽高比例r	旋转姿态角(W,P,R)/(°)
1	(322,418)	(551.97,10.06,644.32)	(458,139)	3.295	$\Phi_A(180,-90,0)$
2	(325,426)	(548.24,20.06,610.67)	(71,302)	0.235	$\Phi_B(90,0,90)$
3	(330,129)	(660.32,32.23,810.91)	(148,114)	1.298	$\Phi_C(0,0,0)$
4	(326,513)	(555.73,22.12,294.65)	(319,152)	2.098	$\Phi_D(180,0,0)$
5	(243,403)	(574.49,-148.58,431.63)	(234,286)	0.818	$\Phi_E(90,0,0)$
6	(412,373)	(574.49,148.58,411.63)	(207,376)	0.551	$\Phi_F(-90,0,0)$

3.4 非合作目标捕获

将上述得到的位姿估计结果通过EtherCAT发送给织女机械臂控制器,控制机械臂进行非合作目标捕获,捕获实验效果如图10所示。

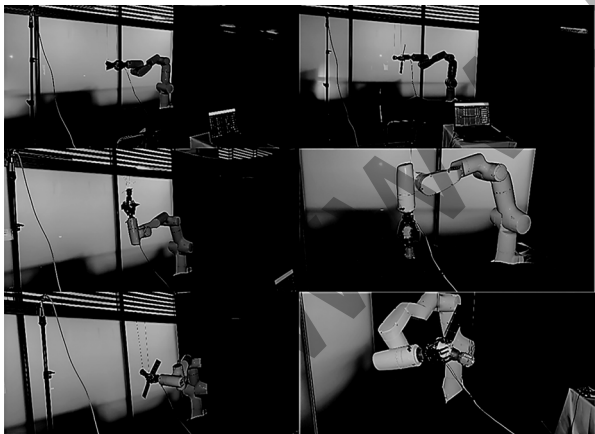


图10 非合作目标捕获的实验效果  
Fig.10 Experimental effect of capturing non-cooperative targets

由图10可知,本文提出的位姿估计算法可以对非合作目标进行准确的空间定位和粗略的姿态估计,其位姿估计结果可以合理地引导机械臂对非合作目标进行捕获。

3.3 非合作目标位姿估计

先采集-20张以上不同姿态的棋盘格图片,然后利用张正友棋盘格标定法<sup>[23]</sup>在Matlab的cameracalib中进行标定得到相机内参。相应地采集40张以上的RGB图以及深度图,利用Tsai方法<sup>[24]</sup>进行手眼标定得到相机外参。

利用标定的结果和YOLO-GhostECA检测网络,通过第2节提出的位姿估计方法进行非合作目标位姿估计,其中姿态估计公式(6)中的 $r_1$ 、 $r_2$ 通过分析非合作目标数据集得到,分别为3.0和1.0, $P_s$ 、 $U_s$ 根据相机空间几何位置分别设定为320和360。根据该位姿估计算法估计的位姿结果见表3。

4 结束语

本文设计一种适用于非合作目标捕获的轻量级位姿估计网络,将深度学习YOLOv5网络应用于空间非合作目标智能识别与捕获任务中,并基于GhostNet网络设计GhostDarknet骨干网络,减少冗余特征图的提取,在基本保持精度不变的情况下,减少网络模型大小,提高模型运算速度。此外,提出一种利用YOLO-GhostECA检测结果来估计物体抓取位姿的方法,并通过深度相机Intel RealSense D435i、非合作目标缩小模型和7自由度冗余机械臂搭建地面仿真实验平台。实验结果表明,该网络在非结构化环境中能很好地完成非合作目标捕获任务。下一步将通过扩充数据集提升网络的精度,并将训练好的网络部署到嵌入式设备中,对实时性和准确度进行验证。

参考文献

[ 1 ] 孙康. 空间机械臂抓捕非合作目标的辨识、规划及控制策略研究[D]. 哈尔滨:哈尔滨工业大学,2020.  
SUN K. Research on identification, kinematic planning and control strategy in capture of non-cooperative target by space manipulator [D]. Harbin: Harbin Institute of Technology, 2020. (in Chinese)  
[ 2 ] 梁斌,杜晓东,李成,等. 空间机器人非合作航天器在轨

- 服务研究进展[J]. 机器人, 2012, 34(2): 242-256.
- LIANG B, DU X D, LI C, et al. Advances in space robot on-orbit servicing for non-cooperative spacecraft[J]. Robot, 2012, 34(2): 242-256. (in Chinese)
- [3] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [4] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C. , USA: IEEE Press, 2014: 580-587.
- [5] REN S Q, HE K M, 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.
- [6] HE K M, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN [C]//Proceedings of IEEE International Conference on Computer Vision. Washington D. C. , USA: IEEE Press, 2017: 2980-2988.
- [7] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C. , USA: IEEE Press, 2016: 779-788.
- [8] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector [C]//Proceedings of European Conference on Computer Vision. Berlin, Germany: Springer, 2016: 21-37.
- [9] 刘帅. 空间机器人抓捕非合作目标的智能控制与识别[D]. 大连: 大连理工大学, 2019.
- LIU S. Intelligent control and recognition of space robot capturing non-cooperative targets [D]. Dalian: Dalian University of Technology, 2019. (in Chinese)
- [10] LI Y W, BO Y M, ZHAO G P. Survey of measurement of position and pose for space non-cooperative target [C]//Proceedings of the 34th Chinese Control Conference. Washington D. C. , USA: IEEE Press, 2015: 5101-5106.
- [11] 余玉琴, 魏国亮, 王永雄. 基于改进YOLOv2的无标定3D机械臂自主抓取方法[J]. 计算机应用研究, 2020, 37(5): 1450-1455.
- YU Y Q, WEI G L, WANG Y X. 3D uncalibrated robotic grasping method based on improved YOLOv2 [J]. Application Research of Computers, 2020, 37(5): 1450-1455. (in Chinese)
- [12] REDMON J, FARHADI A. YOLO9000: better, faster, stronger [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C. , USA: IEEE Press, 2017: 6517-6525.
- [13] REDMON J, FARHADI A. YOLOv3: an incremental improvement [EB/OL]. [2021-01-02]. <https://arxiv.org/abs/1804.02767>.
- [14] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: optimal speed and accuracy of object detection [EB/OL]. [2021-01-02]. <https://arxiv.org/abs/2004.10934>.
- [15] JOCHER G, STOKEN A, BOROVEC J. Ultralytics yolov5: v3.0 (Version v3.0). [EB/OL]. [2021-01-02]. <https://github.com/ultralytics/yolov5/releases/tag/v3.0>.
- [16] ZHANG S, WU Y X, MEN C G, et al. Tiny YOLO optimization oriented bus passenger object detection [J]. Chinese Journal of Electronics, 2020, 29(1): 132-138.
- [17] PURKAIT P, ZHAO C, ZACH C. SPP-net: deep absolute pose regression with synthetic views [EB/OL]. [2021-01-02]. <https://arxiv.org/abs/1712.03452>.
- [18] WANG C Y, MARK LIAO H Y, WU Y H, et al. CSPNet: a new backbone that can enhance learning capability of CNN [C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D. C. , USA: IEEE Press, 2020: 1571-1580.
- [19] LIU S, QI L, QIN H F, et al. Path aggregation network for instance segmentation [C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D. C. , USA: IEEE Press, 2018: 8759-8768.
- [20] HAN K, WANG Y H, TIAN Q, et al. GhostNet: more features from cheap operations [C]//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D. C. , USA: IEEE Press, 2020: 1580-1589.
- [21] WANG Q L, WU B G, ZHU P F, et al. ECA-net: efficient channel attention for deep convolutional neural networks [C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D. C. , USA: IEEE Press, 2020: 11531-11539.
- [22] ZHENG Z H, WANG P, LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression [J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993-13000.
- [23] ZHANG Z Y. Flexible camera calibration by viewing a plane from unknown orientations [C]//Proceedings of the 7th IEEE International Conference on Computer Vision. Washington D. C. , USA: IEEE Press, 1999: 666-673.
- [24] TSAI R. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses [J]. IEEE Journal on Robotics and Automation, 1987, 3(4): 323-344.