

# 基于GR-CNN改进的抓取检测算法及仿真

娄晓恒, 毛鹏军, 牛鑫宇, 段云涛

河南科技大学机电工程学院, 河南 洛阳

收稿日期: 2023年8月31日; 录用日期: 2023年11月3日; 发布日期: 2023年11月13日

## 摘要

随着计算机视觉的发展, 抓取检测在机器人视觉领域取得了广泛应用, 以往的抓取检测方法存在对抓取相关特征学习不够重视、检测精度低、泛化能力差等问题。GR-CNN算法解决了之前这些检测算法存在的一些问题, 但是该算法预测的抓取框所处位置和角度有时会在不准确的情况。针对这一问题, 对该算法进行改进优化。通过加入CBAM注意力机制模块使网络主要集中于学习可抓取的信息, 从而提高抓取预测的准确性。并在ROS环境下对改进后的算法进行验证, 通过Gazebo与MoveIt!联合仿真, 使用UR5机械臂对桌面上随机放置的四类物体进行抓取放置任务, 平均抓取成功率达到86%, 实验结果表明改进后的算法抓取框适应性更强, 对规则物体与形状结构复杂的物体都可以有效抓取。

## 关键词

抓取检测, 注意力机制模块, ROS, Gazebo, MoveIt!

# Improved Grasping Detection Algorithm and Simulation Based on GR-CNN

Xiaoheng Lou, Pengjun Mao, Xinyu Niu, Yuntao Duan

School of Mechanical and Electrical Engineering, Henan University of Science and Technology, Luoyang Henan

Received: Aug. 31<sup>st</sup>, 2023; accepted: Nov. 3<sup>rd</sup>, 2023; published: Nov. 13<sup>th</sup>, 2023

## Abstract

With the development of computer vision, grasp detection has been widely used in the field of robot vision. The previous grasp detection methods have some problems, such as not paying enough attention to grasp-related feature learning, low detection accuracy and poor generalization ability. The GR-CNN algorithm solves some problems existing in the previous detection algorithms, but

the position and Angle of the grab frame predicted by the algorithm are sometimes inaccurate. To solve this problem, the algorithm is improved and optimized. By adding CBAM attention mechanism module, the network mainly focuses on learning the information that can be captured, so as to improve the accuracy of grasping prediction. The improved algorithm was verified in ROS environment through Gazebo and MoveIt! Jointly, the UR5 robot arm is used to carry out the grasping simulation task of four types of objects randomly placed on the table, and the average grasping success rate reaches 86%. The experimental results show that the improved algorithm has stronger adaptability to the grasping frame, and can effectively grasp both regular objects and objects with complex shapes and structures.

## Keywords

Grasp Detection, Attention Mechanism Module, ROS, Gazebo, MoveIt!

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

近年来,机器人的跨学科特性显著增长,人工智能和计算机视觉[1]在机器人领域有着日益增长的趋势。基于深度学习的检测方法主要用于人脸识别、车牌识别、工件缺陷检测[2]等,机器人抓取也涉及到检测问题,将深度学习检测算法应用于机器人视觉领域具有丰富的研究价值。

在非结构化环境下,物体的大小、形状结构不同,所处的背景复杂时,传统抓取检测方法在特征提取时容易产生噪声和影像辨识失败的问题,泛化能力及鲁棒性较差[3]。而采用深度学习检测算法用于抓取姿态检测具有良好的抗噪能力[4],具有更高的准确率,通过进行端到端的训练,能够得到深层的抽象特征,泛化能力更好。机器人抓取系统主要由检测系统、规划系统和控制系统组成[5]。抓取检测的目的是为了通过视觉信息得到合适的抓取位姿,得到位姿信息后进行定位,最后进行机械臂路径规划和抓取执行。抓取的本质是一个检测问题,一个有效的检测算法有利于提高机械臂抓取的成功率,本文以抓取检测为主,主要对抓取检测算法进行展开研究。

深度学习在机器人抓取检测中取得了优异的成绩[6]。视觉检测是机器人伺服控制的关键技术之一,深度学习相对于传统机器视觉有着检测速度快、准确性强、鲁棒性高的优点[7]。Lenz 等人[8]提出将深度学习算法应用于抓取检测问题,利用深度学习算法推断出物体抓取位置。通过接收对象的图像作为输入,并预测抓取矩形作为输出。Gariépy 等人[9]提出一种抓取空间变换网络 GQ-STN (Grasp Quality Spatial Transformer Network)以解决实时抓取检测问题。算法输入为深度图,输出为平面抓取框,抓取表示中只包含深度图的坐标和抓取角度,通过采样得到候选抓取框。Morrison 等人[10]提出了一种用于闭环抓取的实时抓取综合方法,称为生成式抓取卷积神经网络 GG-CNN (Generative Grasping Convolutional Neural Network)。该系统预测每个像素的抓取质量和姿态,这种算法避免了抓取候选对象离散采样和候选时间长的问题,克服了当前深度学习抓取技术的局限性。Yu 等人[11]使用 GG-CNN 构建了一个级联的深度学习框架,利用回归和分类的思想保证了在非结构化环境中机器人操作的抓取鲁棒性。Wang 等人[12]修改了 GG-CNN 模型,提出了一种高效的神经网络模型来生成具有高分辨率图像的抓取表示。为了充分利用 RGB 和深度图像的信息,Kumra 等人[13]在 GG-CNN 的基础上提出了生成残差卷积神经网络 GR-CNN

(Generative Residual Convolutional Neural Network), 这种算法利用不可知论方法来解决抓取未知物体的问题, GR-CNN 是一种功能强大的与对象无关的抓取综合算法, 本文接下来将针对 GR-CNN 算法展开研究。

## 2. 抓取姿态表示

按照机器人抓取方式不同, 分为平面抓取和 6 自由度抓取, 平面抓取是指夹爪垂直于工作台桌面进行抓取的, 位姿由平面坐标和绕  $z$  轴旋转角组成, 6 自由度抓取的位姿由空间坐标值和四元数值组成; 平面抓取在工业领域应用广泛, 因此本文采用平面抓取方式。对物体进行平面抓取时, 由于物体种类、形状不同, 采用沿着物体表面的中心点进行抓取, 具有较高的通用性。通常在图像平面中用一个矩形框来表示夹爪抓取的位置, 表示方法如图 1 所示。这种抓取表示取自 Lenz, 夹爪抓取表示函数由五维矩形向量  $g = (x, y, w, h, \theta)$  表示, 其中  $(x, y)$  表示夹握的中心点坐标,  $w$  表示夹爪所需宽度,  $h$  表示夹爪  $PQ$  两点的开合长度,  $\theta$  表示夹爪相对于图像平面水平轴的旋转角度。

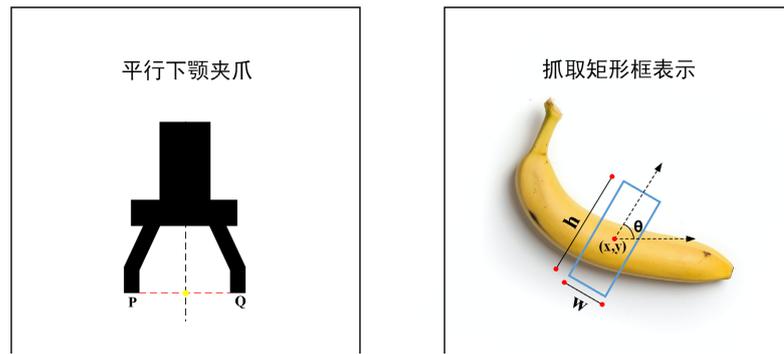


Figure 1. Gripper grasp representation  
图 1. 夹爪抓取表示

与五维抓取向量不同的是, GRCNN 抓取向量定义为:

$$g = (p, \phi, w, q) \tag{1}$$

其中  $p = (x, y, z)$  表示夹爪抓取的中心位置,  $\phi$  表示夹爪绕  $z$  轴的旋转角度,  $w$  为夹爪的开口宽度,  $q$  为抓取质量分数。GRCNN 网络可以从  $H \times W$  的深度图像  $I$  中获取平面抓取姿态:

$$\tilde{g} = (s, \tilde{\phi}, \tilde{w}, q) \tag{2}$$

其中  $s = (u, v)$  表示中心抓取点的像素坐标,  $\tilde{\phi}$  是在相机参照系下沿  $z$  轴的旋转角度,  $\tilde{w}$  为夹爪的开口宽度, 通过坐标变换  $\tilde{g}$  可转化为机器人坐标系下的抓取姿态  $g$ :

$$g = {}^R T ({}^C T (\tilde{g})) \tag{3}$$

其中  ${}^C T$  是利用相机内参将图像平面坐标系变换到相机坐标系的变换矩阵,  ${}^R T$  为相机坐标系到机器人坐标系的变换矩阵。

在图像空间中所有抓取点的集合可以简化表示为:

$$G = (\Phi, W, Q) \tag{4}$$

式中  $\Phi, W, Q$  分别为像素图像的抓取角度、抓取宽度和抓取质量。 $\Phi$  的取值范围在  $[-\pi/2, \pi/2]$  之间,  $W$  的宽度范围在 0 和夹爪宽度之间,  $Q$  的取值范围为  $[0, 1]$ , 值越接近 1 表示在像素图像下抓取质量越好, 抓取成功率越高。

### 3. GR-CNN 算法研究

#### 3.1. GR-CNN 算法概述

GR-CNN 是由多个残差块堆叠构建成的卷积神经网络，相较于之前的检测算法，GR-CNN 不需要提前对物体进行采样。它将深度图像作为输入通过像素信息对抓取姿势进行预测，使用全卷积神经网络以端到端的方式直接输出每个像素的抓取姿势，使其可以抓取未知物体。该网络可以有效地对潜在特征进行解码提取，使其关注有效特征信息，抑制不相关特征信息。

网络使用深度图像  $I$  作为输入，输出为抓取图像  $G$ 。该网络需要图像预处理以匹配 GR-ConvNet 的输入格式，首先需要将图像裁剪成正方形，并缩放到  $224 \times 224$  像素大小，利用 OpenCV 对无效深度值进行补绘，然后通过高斯核对抓取位姿进行滤波处理，降低图像噪声去除异常值。从输入图像中提取特征，将特征信息输送到 3 个卷积层中然后经过 5 个残差层和卷积转置层的传播，输出得到抓取矩形框，检测模块如图 2 所示：

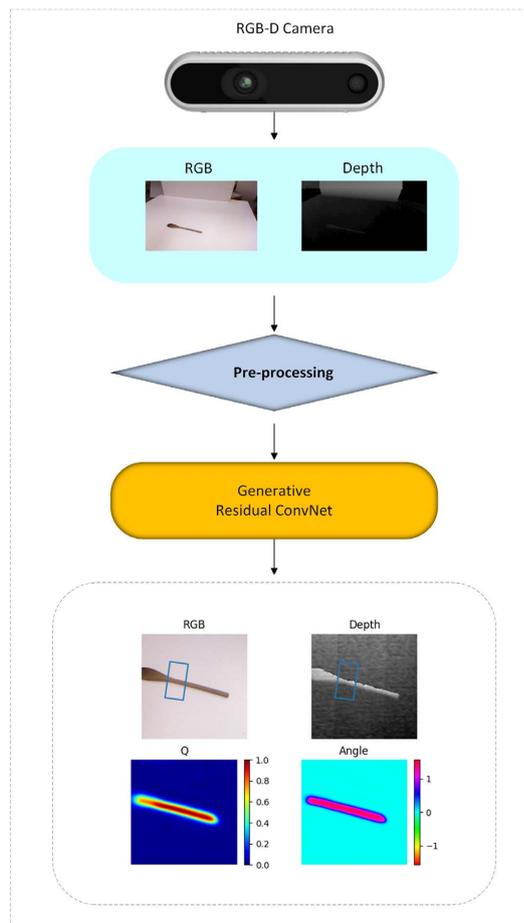


Figure 2. Detection module  
图 2. 检测模块

但这种算法得到的预测结果并不是很准确，预测框所处位置与预测框旋转角度存在偏差，该算法的关注点在一些小区域，抓取区域并不是最优，物体中心抓取区域的质量分数低，存在无法将目标物体与背景区域分离等问题。因此，针对以上存在的问题，对该算法进行改进优化。

### 3.2. GR-CNN 算法改进

针对 GR-CNN 上述所存在的问题，通过在网络中加入注意力机制模块来提高网络预测的性能，注意力机制可以快速有效地分析复杂的场景信息以提高系统性能。通过特征自适应加权来实现对图像中重要信息进行动态选择，过滤掉深度学习样本中不相关信息，筛选出对当前任务需要的关键信息。

本文为了在 GR-CNN 网络中更好的区分特征，使其得到有效的消息，抑制无效特征权重，加入卷积注意力机制模块 CBAM (Convolutional Block Attention Module)。CBAM 是一种轻量级的简单而有效即插即用的通用注意力模块，加入 CBAM 注意力机制模块后对复杂架构神经网络影响小[14]。该模块可以有有效的集成到神经网络中，与基础卷积神经网络(CNN)一起进行端到端训练，并且不会增加额外的开销，可以提高卷积神经网络的精度，有助于改进特征映射和提高性能。CBAM 网络结构图如图 3 所示：

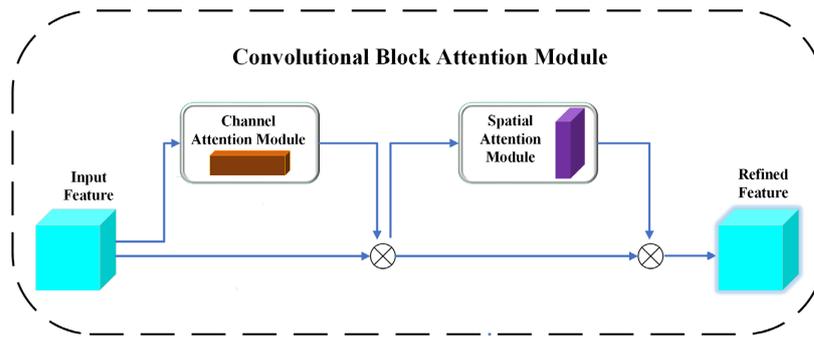


Figure 3. CBAM network structure diagram  
图 3. CBAM 网络结构图

CBAM 注意力机制模块由通道注意力模块(CAM)和空间注意力模块(SAM)两个子模块组成，CAM 与 SAM 网络结构如图 4 所示：

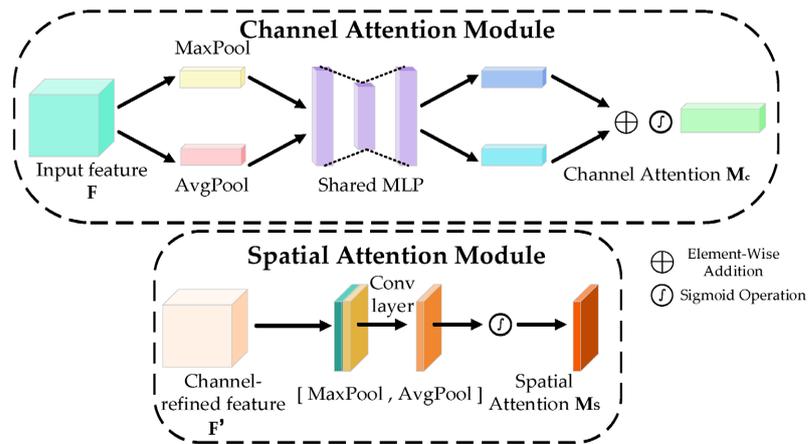


Figure 4. Channel attention module (CAM) and Spatial attention module (SAM) network structure  
图 4. 通道注意力模块(CAM)和空间注意力模块(SAM)网络结构

通道注意力机制(CAM)是 CBAM 模块的核心组成部分，其通过卷积层提取图像特征，卷积层根据图像的尺寸、深度参数，生成矩形形式的特征映射。空间注意力模块(SAM)将特征图中的每个像素编码到空间注意图中，并通过训练 SAM 为每个像素分配权重。经过平均池化与最大池化操作，在通道维度拼接得到特征映射  $F'$ ，由卷积网络提取特征，最后通过 Sigmoid 函数为每个像素分配新的权重。公式如下：

$$F_{new} = \sigma\left(f\left(\text{Concat}\left(\text{AP}(F), \text{MP}(F)\right)\right)\right) \quad (5)$$

其中  $f$  表示卷积核， $\text{AP}$  表示平均池化， $\text{MP}$  表示最大池化， $\text{Concat}$  表示连接操作， $\sigma$  表示 Sigmoid 函数。

CBAM 注意力模块在深度网络卷积层上通过自适应定义中间特征模块，使用平均池化方法与最大池化方法进行提取通道注意力图，在这种情况下，提取的目标特征更全面、精度更高、性能更好、鲁棒性更强。能够有效提高网络性能，增强通道信息之间的特征融合能力。通过最大池化可以进一步获取得到特征图中深度特征，使神经网络将关注放在重要信息的目标区域，抑制不相关信息，提高抓取检测的整体精度。

加入 CBAM 注意力模块后的网络共有 6.5M 个参数，与其它使用包含数百万个参数的复杂架构预测网络相比该网络更加轻量化，网络相对简短，资源耗费低。加入注意力机制模块后的生成式残差卷积神经网络模型如图 5 所示：

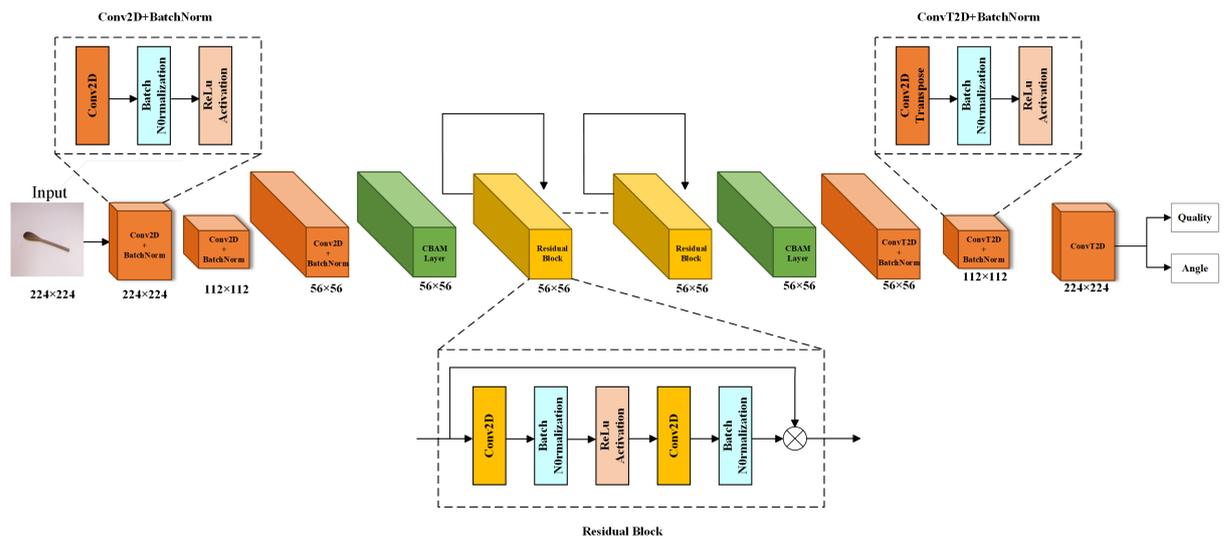


Figure 5. The residual convolutional neural network model is added to the attention mechanism module

图 5. 加入注意力机制模块的神经网络模型

### 3.3. 预测评估指标

抓取向量的准确性评估指标是指预测抓取框的位置是否接近理论抓取位置。如果满足以下两个条件，则预测抓取向量是可取的：

- 1) 预测抓握矩形框偏移方向与理论抓取夹角差绝对值在  $30^\circ$  以内。
- 2) 预测抓握框与理论抓握框的 Jaccard 相交系数大于 25%。评价指标如下：

$$J(g_p, g_t) = \frac{|g_p \cap g_t|}{|g_p \cup g_t|} > 0.25 \quad (6)$$

向量  $g_p \cap g_t$  是检测到的抓取矩形与理论抓取之间的重叠面积， $g_p \cup g_t$  是两者的并集面积，如果  $J(g_p, g_t)$  高于 0.25，且预测框偏移方向与实际抓取旋转角度差在  $30^\circ$  以内，则表示预测的抓取矩形是可取的[7]。

### 3.4. 算法验证

在 Pycharm2021 平台下基于深度学习 Pytorch1.7.1、Cuda11.0、Cudnn8.0.4 框架进行训练测试。采用

Cornell 抓取数据集对算法进行验证, Cornell 数据集由 240 个生活中常见的物体组成, 该数据集包含内容丰富, 由 885 张 RGB 图像和与其对应的深度图组成, 数据集中 80% 用作训练集, 其余 20% 用于测试集[15]。通过 Cornell 抓取数据集来评估改进前后网络的抓取效果, 以验证该模型对不同物体的泛化能力。

对于数据集  $D = \{D_1 \cdots D_n\}$ , 通过输入场景图像信息  $I = \{I^1 \cdots I^n\}$  和在图像信息上成功抓取的  $G_i = \{g_1^1 \cdots g_{m_1}^1 \cdots g_1^n \cdots g_{m_n}^n\}$ , 端到端的训练此模型, 通过输入图像信息  $I$  学习以  $G_i$  为条件的最小负对数似然函数来得到映射函数  $\gamma(I, D) = G_i$ , 式如下:

$$-\frac{1}{n} \sum_{i=1}^n \frac{1}{m_i} \sum_{j=1}^{m_i} \log \gamma(g_j^i | I^i) \quad (7)$$

模型优化器使用 Adam, 利用标准反向传播和小批量随机梯度下降(SGD)技术训练模型, 批量大小为 8, 学习率设置为  $10^{-3}$ , 迭代次数为 150 次。

对 GR-CNN 算法训练后进行评估, 通过输入图像对抓取框进行评估预测, 最终得到在 RGB 图和深度图上的抓取预测框。可视化效果如图 6 所示, 可视化图中包括 RGB 图、深度图、质量图 and 角度图。

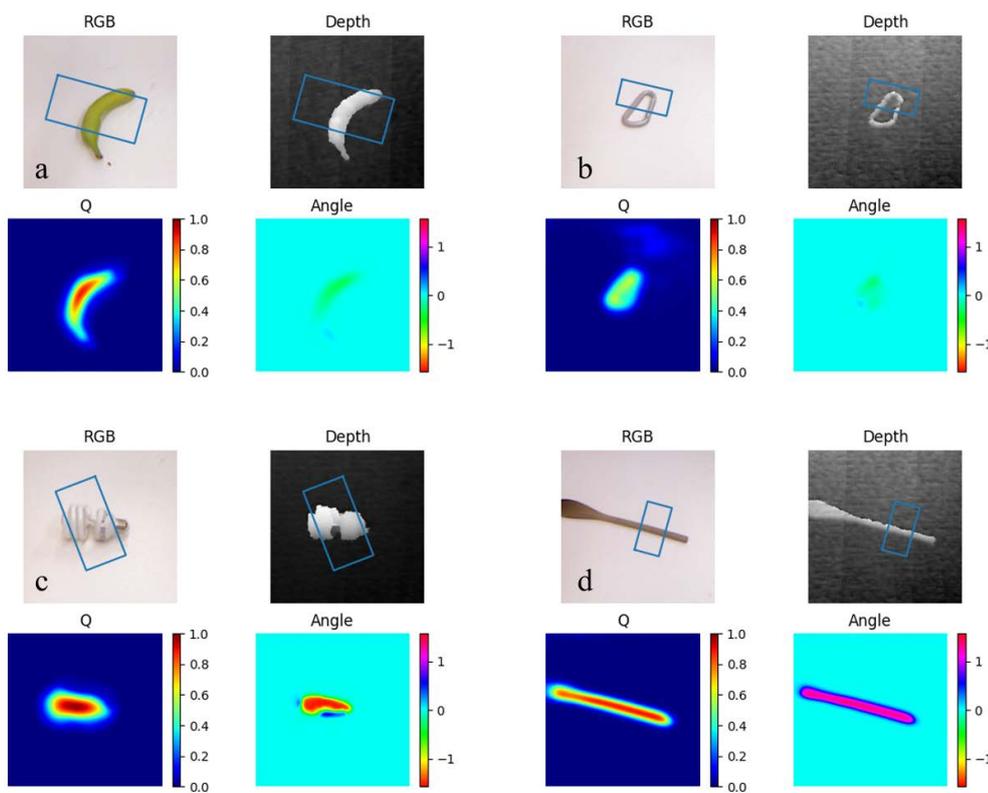
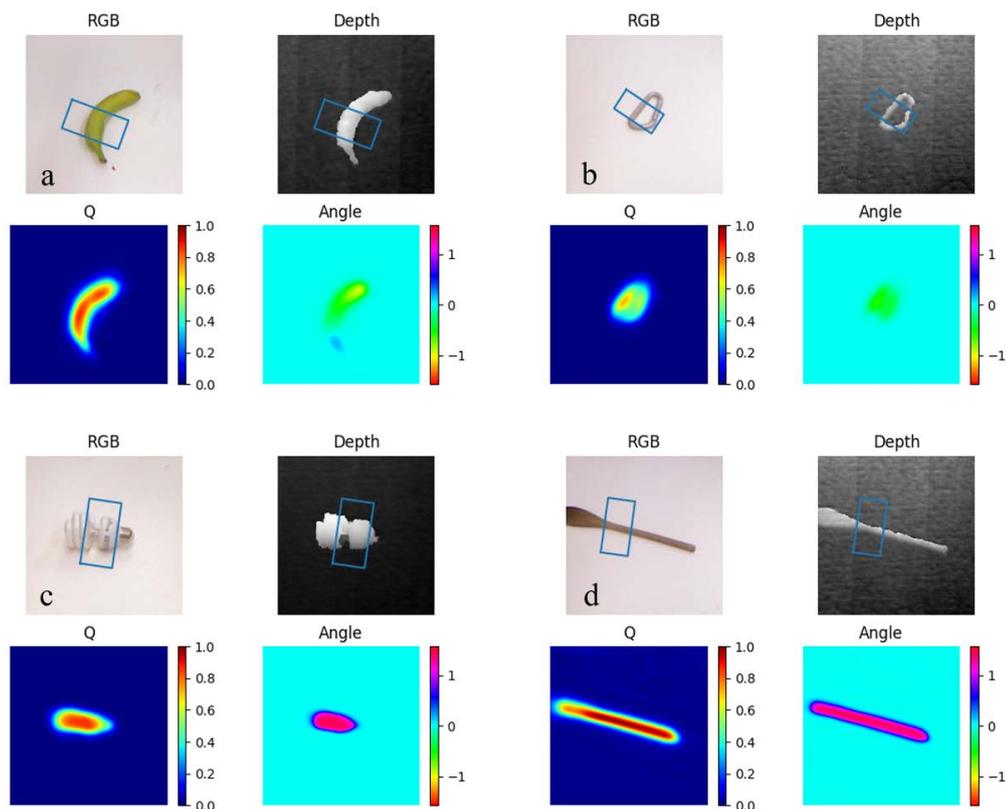


Figure 6. Visualization of grasping prediction before improvement

图 6. 改进前的抓取预测可视化效果

从预测出的结果可以看出, 图 6(a)物体预测框过大, 预测框抓取范围模糊, 图 6(b)物体存在抓取位置偏移的问题, 图 6(c)物体位置预测正确但是抓取框的角度存在偏差, 图 6(d)物体预测位置非最佳。由可视化图可以看出 GR-CNN 算法得到的预测结果存在偏差, 预测出的抓取结果并不是抓取最优结果, 这种预测差异会导致夹爪在实际抓取时存在偏移, 可能会出现抓取脱落的风险, 最终将导致夹爪抓取物体失败。通过加入注意力机制模块对该算法进行改进, 以降低原算法存在的预测框过大和位置角度偏差问题。改进后的抓取可视化效果如图 7 所示:



**Figure 7.** Visualization of grasping prediction after improvement  
**图 7.** 改进后的抓取预测可视化效果

改进后的预测抓取框在抓取位置和偏转角度都更加接近真实理论抓取，预测效果达到了理论预期标准。表明该抓取优化算法针对抓取检测问题可以取得很好的预测效果，改进后的算法可以有效的提高抓取预测准确率。

## 4. 实验与分析

### 4.1. 仿真平台

当前机器人成本居高不下，在真实机器人下开发和测试算法是一个复杂而昂贵的过程，通过仿真可以降低工程设计成本，加快开发周期。本实验在 Ubuntu 系统 18.04 Melodic 版本 ROS1 环境下进行测试，在 Gazebo 中使用 UR5 机械臂仿真抓取以验证改进后 GR-CNN 算法的有效性。

ROS 是一个开源的机器人平台，具有丰富的编程语言特点。各个节点间通过 Topic、Service、Action 等方式进行通信，适用于机械臂抓取这种多节点任务。ROS 有内置的开发工具 Gazebo 和 Rviz。Gazebo 是一款 3D 动态模拟器，可以提供一个虚拟的测试和验证环境。不仅可以仿真机器人的运动，还可以仿真传感器数据，可以准确、高效地模拟机器人在工作环境中的状态。具有成本低、灵活性强、安全性高的优势；Rviz 是三维可视化工具，用于观察坐标变换(TF)、图像信息、机器人状态、物体状态、传感器数据，能将数据可视化显示；

机械臂模型选用 UR5 机械臂，UR5 具有高精度高灵敏度的特性，控制效率高，并且提供了 ROS 控制接口。UR5 完整的生态系统使其能够轻松地与其他技术集成，从而实现目标检测和复杂的抓取任务。机械臂夹爪选用 Robotiq 2F-85 二指夹爪，深度相机选用 RealSense D435i，相机的安装方式采用眼在手上

的方式，图 8 为在 Gazebo 中的实验场景，在 Gazebo 中搭建了两个桌子，分别用于固定 UR5 机械臂和放置待抓取物体，工作台桌面放置了四类物体分别为塑料瓶、扳手、香蕉、易拉罐，机械臂右端放置了一个蓝色收集方框，用于放置抓取完的物体。

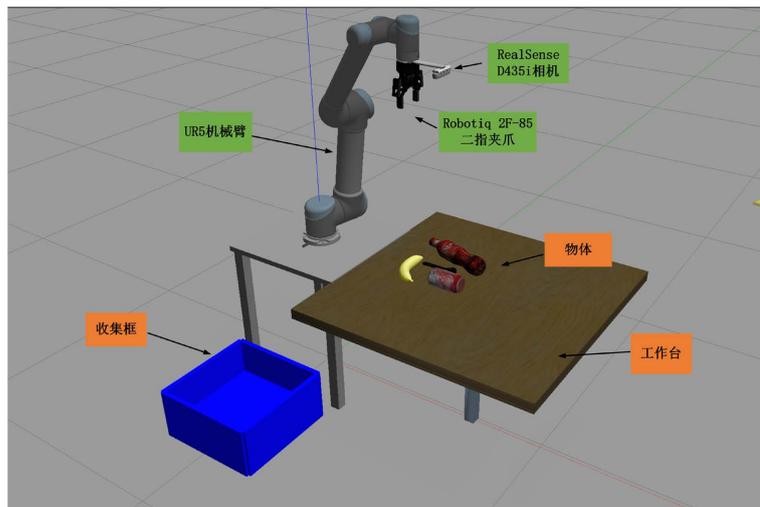


Figure 8. Gazebo simulation experiment scene  
图 8. Gazebo 仿真实验场景

## 4.2. 仿真实现

在 ROS 系统中，本文通过 URDF 以 XML 的格式描述 UR5 机械臂、相机、夹爪、桌子等结构，通过 ROS 中 C++ 内置解释器将 URDF 转换成可视化模型，并在 Gazebo 中搭建仿真环境，通过 MoveIt! 实现 UR5 机器人的运动控制。本仿真实验中主要使用到 Python、C++ 语言。

仿真系统由仿真环境模块、视觉模块、规划模块以及抓取执行模块组成[16]。仿真抓取实现过程分为 6 个步骤：1) 在 Gazebo 中搭建环境。2) 通过 RGBD 相机获取待抓取对象深度像素信息。3) GR-CNN 网络预测物体的抓取姿态。4) 坐标转换。5) 机械臂路径规划。6) 抓取执行。在 Gazebo 中由 RealSenseD435i 相机采集物体的深度图像信息，将相机得到的图像信息以话题的方式发布出来，机械臂订阅相关话题，得到抓取框中心点的像素坐标，通过坐标转换，得到抓取目标点在机器人坐标系下的位置，然后进行规划抓取。

执行抓取任务还需要建立 Gazebo 和 MoveIt! 之间的通信，主要利用 Follow Joint Trajectory、Joint Trajectory Controller、Joint State Controller 这三个 ROS 模块进行数据交互[17]。机械臂得到控制指令后由 MoveIt! 进行关节规划，将规划后的轨迹信息通过 Action 形式发布出来，Gazebo 接收数据并将轨迹数据传给关节轨迹控制器，经过插补运算后由机械臂位置伺服驱动器控制机械臂运动。

本文使用 UR5 机械臂对桌面上放置的四类物体进行抓取，抓取过程如图 9 所示。图 9(a)~(d)分别为塑料瓶、扳手、香蕉、易拉罐的抓取过程。图中展示了机械臂抓取的初始准备状态、抓取执行状态、抓取放置状态。机械臂从初始状态运动接近目标后打开夹爪，抓取目标，然后闭合夹爪，控制机械臂将抓取到的物体放置到指定的收集框中。

## 4.3. 实验结果分析

本文在 Gazebo 环境中搭建 UR5 机械臂、RealsenseD435i 深度相机、Robotiq 2F-85 二指夹爪模型，其中深度相机用于检测物体图像信息，通过图像信息为夹爪确定抓取姿势。使用四类生活中常见的物体

作为抓取对象，这四类物体是随机排列在工作平面上的，只有当物体被抓取并放置到蓝色收集框中才算是抓取成功的。

对这四类物体进行 30 次抓取实验，表 1 为 UR5 机械臂对四类物体抓取实验后得到的统计结果，抓取成功率平均在 86% 左右。改进后的算法可以在杂乱随机的场景下抓取物体，当然，在抓取过程中也存在抓取失败的情况，有时虽然机械臂抓取到物体了，但在机械臂运动时出现打滑脱落的情况，原因是由于预测的深度图像存在偏差。

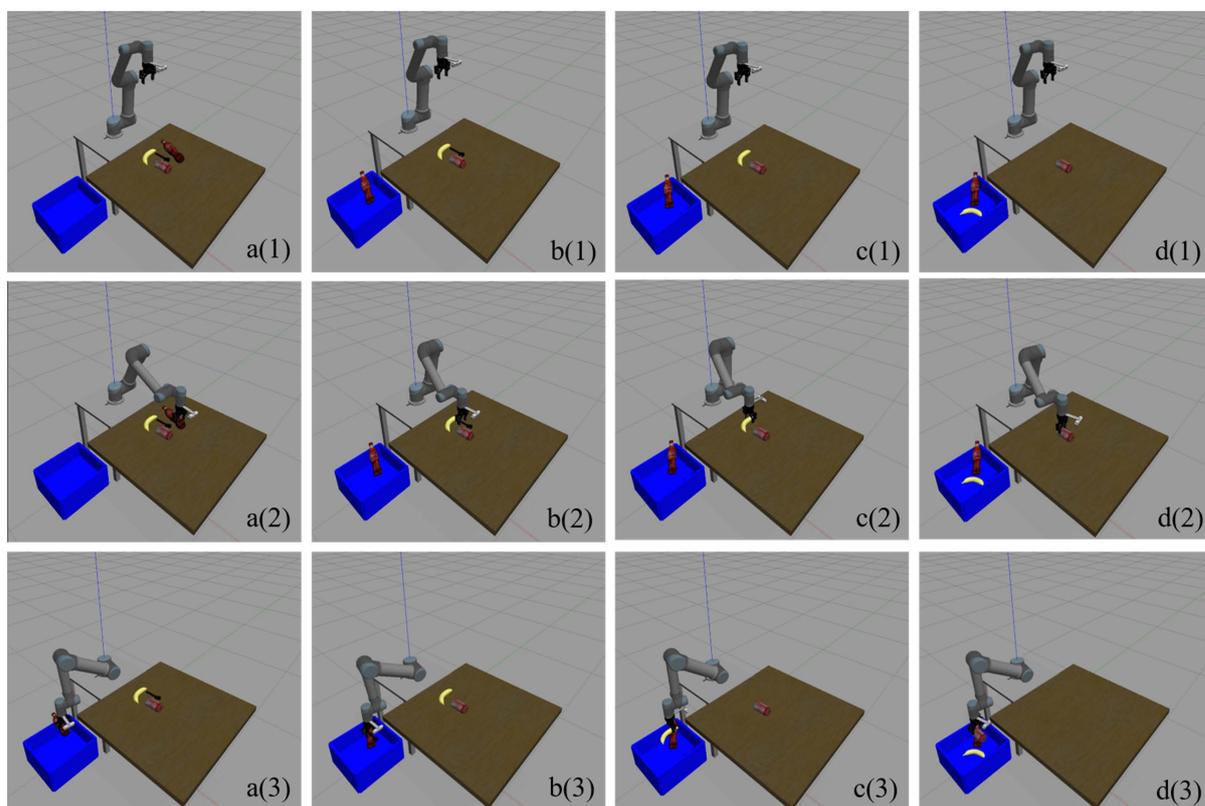


Figure 9. Gazebo simulates the grasp process

图 9. Gazebo 下仿真抓取过程

Table 1. Grasp experimental statistics

表 1. 抓取实验统计结果

种类	总抓取次数	抓取成功次数	成功率(%)
塑料瓶	30	26	86.7
扳手	30	25	83.3
香蕉	30	26	86.7
易拉罐	30	27	90.0

## 5. 结束语

为了解决 GR-CNN 抓取预测框偏移问题，本文对 GR-CNN 算法进行改进，在不损失原算法精度的前提下，通过加入 CBAM 注意力模块对算法进行优化。使得改进后的网络主要集中于学习可抓取的信息，

解决了预测框偏移问题。最后基于 ROS 搭建了 MoveIt!-Gazebo 联合仿真实验平台, 在 Gazebo 中使用 UR5 机械臂对四类常见的物体进行仿真抓取操作, 抓取成功率平均在 86% 左右, 实验结果表明该算法适应性更强, 对规则物体与形状结构复杂的物体都可以有效抓取。

## 基金项目

洛阳市科技重大专项(2101018A)。

## 参考文献

- [1] Caldera, S., Rassau, A. and Chai, D. (2018) Review of Deep Learning Methods in Robotic Grasp Detection. *Multi-modal Technologies and Interaction*, **2**, 57. <https://doi.org/10.3390/mti2030057>
- [2] Ye, L., Xia, X., Chai, B., et al. (2021) Application of Deep Learning in Workpiece Defect Detection. *Procedia Computer Science*, **183**, 267-273. <https://doi.org/10.1016/j.procs.2021.02.058>
- [3] 刘艳, 冯志杰. SSD 深度学习算法在机械手抓取系统中的应用[J]. 组合机床与自动化加工技术, 2020, 556(6): 67-70.
- [4] Redmon, J. and Angelova, A. (2015) Real-Time Grasp Detection Using Convolutional Neural Networks. 2015 *IEEE International Conference on Robotics and Automation (ICRA)*, Seattle, 26-30 May 2015, 1316-1322. <https://doi.org/10.1109/ICRA.2015.7139361>
- [5] 刘亚欣, 王斯瑶, 姚玉峰, 等. 机器人抓取检测技术的研究现状[J]. 控制与决策, 2020, 35(12): 2817-2828.
- [6] 陈春朝, 孙东红. 基于 YOLOv5 的角度优化抓取检测算法研究[J/OL]. 计算机工程与应用, 1-9. <http://kns.cnki.net/kcms/detail/11.2127.tp.20230118.1624.007.html>, 2023-06-07.
- [7] Asif, U., Tang, J. and Harrer, S. (2018) EnsembleNet: Improving Grasp Detection Using an Ensemble of Convolutional Neural Networks. *BMVC*, **10**, 1-12.
- [8] Lenz, I., Lee, H. and Saxena, A. (2015) Deep Learning for Detecting Robotic Grasps. *The International Journal of Robotics Research*, **34**, 705-724. <https://doi.org/10.1177/0278364914549607>
- [9] Gariépy, A., Ruel, J.C., Chaib-Draa, B., et al. (2019) Gq-stn: Optimizing One-Shot Grasp Detection Based on Robustness Classifier. 2019 *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Macau, 3-8 November 2019, 3996-4003. <https://doi.org/10.1109/IROS40897.2019.8967785>
- [10] Morrison, D., Corke, P. and Leitner, J. (2020) Learning Robust, Real-Time, Reactive Robotic Grasping. *The International Journal of Robotics Research*, **39**, 183-201. <https://doi.org/10.1177/0278364919859066>
- [11] Yu, H., Lai, Q., et al. (2019) A Cascaded Deep Learning Framework for Real-Time and Robust Grasp Planning. In *Proceedings of the 2019 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Dali, 6-8 December 2019, 1380-1386. <https://doi.org/10.1109/ROBIO49542.2019.8961531>
- [12] Wang, S., Jiang, X., Zhao, J., et al. (2019) Efficient Fully Convolution Neural Network for Generating Pixel Wise Robotic Grasps with High Resolution Images. In *Proceedings of the 2019 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Dali, 6-8 December 2019, 474-480. <https://doi.org/10.1109/ROBIO49542.2019.8961711>
- [13] Kumra, S., Joshi, S. and Sahin, F. (2020) Antipodal Robotic Grasping Using Generative Residual Convolutional Neural Network. 2020 *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Las Vegas, 24 October 2020-24 January 2021, 9626-9633. <https://doi.org/10.1109/IROS45743.2020.9340777>
- [14] Fu, H., Song, G. and Wang, Y. (2021) Improved YOLOv4 Marine Target Detection Combined with CBAM. *Symmetry*, **13**, 623. <https://doi.org/10.3390/sym13040623>
- [15] 关立文, 孙鑫磊, 杨佩. 基于关键点估计的抓取检测算法[J]. 计算机工程与应用, 2022, 58(4): 267-274.
- [16] 孙鑫磊. 基于 RGB-D 图像的抓取检测深度学习算法研究[D]: [硕士学位论文]. 成都: 电子科技大学, 2021.
- [17] Jung, H., Kim, M., Chen, Y., et al. (2020) Implementation of a Unified Simulation for Robot Arm Control with Object Detection Based on ROS and Gazebo. 2020 *17th International Conference on Ubiquitous Robots (UR)*, Kyoto, 22-26 June 2020, 368-372. <https://doi.org/10.1109/UR49135.2020.9144984>