

# 基于神经网络识别的垃圾分类平衡机器人

高 昂<sup>1</sup>, 贺文嘉<sup>2</sup>, 董文硕<sup>2</sup>

<sup>1</sup>辽宁科技大学计算机与软件工程学院, 辽宁 鞍山

<sup>2</sup>辽宁科技大学机械工程与自动化学院, 辽宁 鞍山

收稿日期: 2023年9月25日; 录用日期: 2023年11月15日; 发布日期: 2023年11月24日

## 摘 要

本机器人具有广泛应用前景, 解决日益严重的环境污染和资源浪费问题, 能在复杂环境中解决人工操作难以实现的自动化, 精确化, 智能化问题。该平衡机器人具有自主定位、自主移动、自主避障、通过垃圾、垃圾抓取与搬运、路径规划等功能。利用外界算力平台, 通过神经网络构建垃圾模型, 计算特征点, 打包传入嵌入式AI计算设备(TX2), 通过TX2解算相应的模型, 更精确的视觉识别物品模型。同时设计相对应程序使得机器人每个任务都安全可行。并且机器人可在狭窄环境下实现换向。加入视觉识别相机实现垃圾识别、以及放置到指定位置。其次, 该机器人加入机械夹爪, 可对大体积垃圾进行抓取放置。该机器人的底盘使用STM32作为主控芯片, 可以迅速反映, 快速获取电机, 舵机等反馈数据。且应用PID算法进行控制, 根据目标轨迹与当前误差, 动态地调整控制参数, 机器人四周都装有距离传感器可自动避障, 且下方装配了定位轮实时监测机器人位置, 若发生问题可以及时做出反应。

## 关键词

神经网络, 平衡机器人, PID, TX2, STM32, 视觉识别

# Garbage Sorting Balancing Bot Based on Neural Network Recognition

Ang Gao<sup>1</sup>, Wenjia He<sup>2</sup>, Wenshuo Dong<sup>2</sup>

<sup>1</sup>School of Computer and Software Engineering, University of Science and Technology, Anshan Liaoning

<sup>2</sup>School of Mechanical Engineering and Automation, University of Science and Technology, Anshan Liaoning

Received: Sep. 25<sup>th</sup>, 2023; accepted: Nov. 15<sup>th</sup>, 2023; published: Nov. 24<sup>th</sup>, 2023

## Abstract

The robot has a wide range of application prospects, to solve the increasingly serious environmental pollution and resource waste problems, can solve in the complex environment manual op-

eration difficult to achieve automation, precision, intelligent problems. The balance robot has the functions of autonomous positioning, autonomous movement, autonomous obstacle avoidance, garbage, garbage grab and handling, path planning and so on. Using the external computing power platform, the garbage model is constructed through the neural network, the feature points are calculated, and the garbage model is packaged into the embedded AI computing device (TX2). The corresponding model is solved through TX2, and the object model is more accurate visual recognition. At the same time, the corresponding program is designed to make each task of the robot safe and feasible. And the robot can change direction in a narrow environment. A visual recognition camera is added to recognize garbage and place it in a designated position. Secondly, the robot adds a mechanical claw, which can grasp and place large volumes of garbage. The chassis of the robot uses STM32 as the main control chip, which can quickly reflect and obtain feedback data such as motor and steering gear. Moreover, PID algorithm is applied for control, and the control parameters are dynamically adjusted according to the target trajectory and current error, Distance sensors are installed around the robot to automatically avoid obstacles, and the positioning wheel is equipped below to monitor the robot's position in real time, so that it can respond in time if problems occur.

## Keywords

Neural Network, Balancing Robot, PID, TX2, STM32, Visual Identification

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

随着人民环保意识的提高,可回收垃圾分类问题得到广泛关注,垃圾分类机器人能够有效的提高垃圾处理效率。如今垃圾分类机器人正处于发展初期阶段,在垃圾分类和垃圾处理方面存在一些问题和挑战。其中包括:机器人识别垃圾的准确性、受限于实际工作环境的复杂性、且可回收垃圾种类繁多,难以通过视觉方法进行材质分类。本文针对以上问题,提出了一种基于神经网络[1]的取放一体平衡机器人[2]。(1)通过设计机械夹爪,通过深度学习在位姿估计等领域的优越性能,为机械夹爪提供准确的位姿信息,提高了垃圾分类任务准确性;(2)设置定位轮,实现机器人全自动定位,通过定位轮使用两个坐标系二分法精准位置解算使得机器人能稳定到达目的地,并通过摄像头避障,实现时刻勘探周围环境,使其在复杂的环境下仍能安全作业;(3)提出基于卷积神经网络[3]的垃圾图像分类算法(Garbage Classification Net, GCNet),在网络结构中融合了注意力机制模块与特征融合模块,提高了模型在垃圾分类任务上的准确性与鲁棒性,解决其难以分类的问题。

## 2. 系统设计

### 2.1. 系统框架

本机器人的运动控制系统采用 STM32F429IGT6 作为主控。STM32F429IGT6 具有高性能的 ARM Cortex-M4 处理器,可以高效地执行各种指令,处理复杂的算法和任务。该处理器还具有安全性如存储器保护单元和执行保护,有助于保护机器人系统免受恶意攻击。此外,它还提供丰富的外设功能包括,通信接口、如 UART、SPI、I2C、CAN)、模拟/数字转换器(ADC/DAC)、定时器、PWM 控制。可以依此实现机器人系统的多样化复杂功能。此外,STM32F429IGT6 还具有低功耗特性,同时成本优势突出,易于

获取开发工具，这使得它们适用于成本敏感的应用。机器人的视觉识别系统采用 Jetson TX2。Jetson TX2 是一款高性能的嵌入式计算模块，特别设计用于机器学习、计算机视觉、自动驾驶和人工智能等应用。可提供超过 1TFLOPS 的高性能计算能力，并支持多种编程模型，包括神经网络、图像处理和流媒体编码等，还支持多种传感器、控制器和执行器的连接，并配备了丰富的开发工具和样例，方便开发人员对机器人应用进行复杂开发。Jetson TX2 拥有庞大的开发社区和生态系统，可获得大量的文档、示例代码和支持资源，帮助我们快速上手该垃圾分类机器人的设计。由此可知，它的高性能、低功耗和丰富的功能使其在本机器人设计研发中起到了重大作用。

## 2.2. 具体步骤

首先机器人开电启动，摄像头采集环境图像，随后将其传递给图像处理模块(TX2)，该模块对图像进行处理后将处理后的数据传递给运动控制模块。在数据处理过程中，运动控制模块运用 PID [4]控制算法实现实时控制机器人的运动状态[5]。最终，在运动控制模块的规划和控制下，机器人将到达目标位置，并利用机械夹爪拾取目标垃圾并将其存储。

控制流程如图 1 所示。

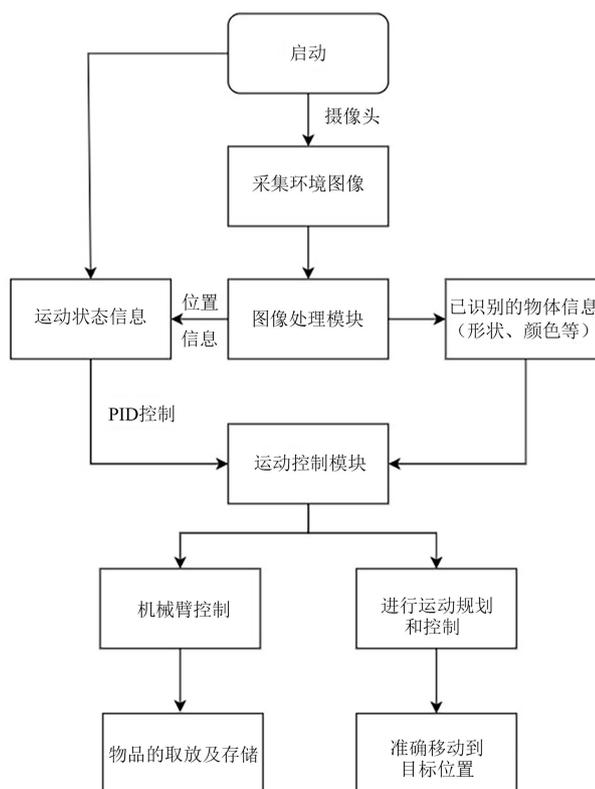


Figure 1. Control flowchart

图 1. 控制流程图

## 2.3. 系统模块

### 2.3.1. 感知模块

以神经网络进行图像的识别和处理，环境感知以及目标检测，使机器人更精确，更复杂地理解环境，处理各种垃圾。采用了 Jetson TX2。

### 2.3.2. 运动控制模块

使用 STM32F427IGT6 作为主控，通过 CAN 通信方式获得各部分电机数据，以 PID 算法实现控制，机器人可以根据目标轨迹和当前位置的误差，动态地调整控制参数，以达到预期的控制轨迹，多重闭环可来控制电机，YAW 轴电机与 PITCH 轴电机。

### 2.3.3. 定位模块

加入定位轮，实现机器人全自动定位，通过定位轮使用两个坐标系二分法精准位置解算使得机器人能稳定到达目的地，并通过摄像头避障，实现时刻勘探周围环境，保证线路的安全。

## 3. 硬件设计

### 硬件选型

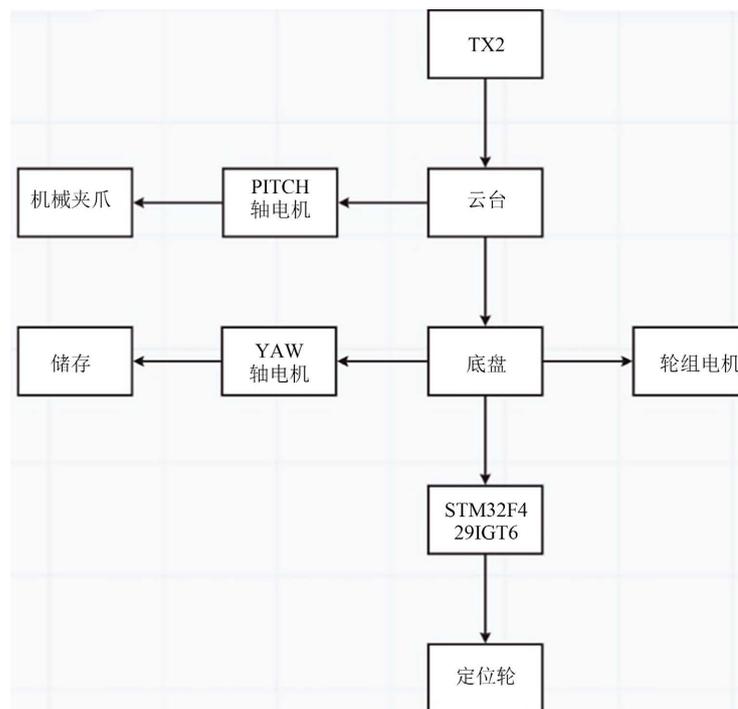


Figure 2. Hardware connection diagram

图 2. 硬件连接图

1. Jetson TX2 用于处理视觉系统采集到的物体图片，并输出检测结果：Jetson TX2 是用于机器人的高性能计算平台，特别适用于视觉处理和深度学习任务。它可以接收来自摄像头等感知设备的图像数据，使用内置的神经网络和图像处理功能，进行垃圾检测、识别和分析。然后，它可以输出检测到的垃圾的位置、类别等信息，为机器人的决策和控制提供关键数据。

2. STM32F429IGT6 用于底盘主控及云台底盘信息传递：STM32F429IT6 是一款嵌入式微控制器，用于底盘主控和云台底盘信息传递。它负责控制底盘上的电机，包括轮组电机、YAW 轴电机和 PITCH 轴电机。此外，它还通过 CAN 通信协议与云台上的其他电机和控制系统交互，确保底盘和云台之间的协同工作。

3. 陀螺仪选用单片机内置的 BMI088：BMI088 是一款高性能的陀螺仪传感器，通常用于测量机器人的姿态和运动。在这个系统中，它似乎与单片机内置的陀螺仪一起使用，以提供精确的姿态信息，以支

持机器人的稳定性和导航。

4. 大疆 M3508 电机：YAW 轴选用了大疆的 M3508 电机，具有有感 FOC (Field-Oriented Control) 控制，可提供出色的扭矩和稳定性。它们用于驱动轮组和 YAW 轴，确保机器人在不同速度和负载下保持平稳的动力响应。

5. 大疆 M2006 电机：根据需求，PITCH 轴选择了大疆的 M2006 电机，这些电机具有高精度的控制、强大的输出功率和小巧的体积，用于为机械夹爪提供动力支持。

6. 云台和底盘：云台上不仅部署 Jetson TX2 以外，而且还有控制机械臂的 PITCH 轴电机、YAW 轴电机和轮组电机。底盘安装了 STM32 主控以及定位轮，便于精准定位机器人位置。

硬件连接图如图 2。

## 4. 软件设计

### 4.1. 软件架构

在这个项目中，我们选择了 NVIDIA 的 Jetson TX2 作为嵌入式计算平台，并在其上配置了 Ubuntu 16.04 操作系统以及 OpenCV 3.4, CUDA 等关键软件，以构建高性能的计算环境，训练神经网络，机器人利用 TX2 平台进行垃圾图像的采集和处理，并将处理后的图像输入神经网络模型中进行分类。这个计算平台的主要任务是处理从摄像头获取的图像数据，并将处理结果通过 UART 串口通信传输给 STM32 单片机。

利用摄像头和 Jetson TX2，我们能够捕获空间图像信息，并使用像素级评估技术进行深度图像分析。同时，我们还能够构建目标物体的三维模型，并提取其关键特征点。这使得我们能够在 TX2 上进行复杂的图像解算，以准确测量机器人和垃圾之间的距离。

此外，项目还在 Jetson TX2 平台上建立了一个强大的机器视觉开发环境，利用 OpenCV 等技术，与神经网络学习框架集成，以实现神经网络模型的训练与部署。且实现了本地端图形方式的信息传递功能。这意味着我们能够更加直观地与机器人的感知和决策系统交互，从而更好地控制机器人执行任务。

### 4.2. 平衡控制算法

为了生成程序骨架并得到垂直方向上的平衡控制，可以使用 STM32CubeMX 工具并添加相关库和驱动程序，需要添加与平衡控制相关的库和驱动程序。这包括陀螺仪传感器的库，CAN 总线通信的库，以及 PID 控制算法的库。并读取陀螺仪每个时刻的测量值，来实现控制的平衡。在平衡控制中，PID 控制算法起着关键作用。PID 控制器由比例(P)、积分(I)和微分(D)三个部分组成。P 系数决定了系统的响应速度和稳定性，I 系数用于消除静态误差，D 系数有助于减小震荡。选择合适的 PID 系数至关重要。这些系数的选择可以通过实验和仿真来进行优化，以实现更为稳定的平衡控制。通常，需要不断调整 PID 系数，直到达到满意的平衡性能。在程序中，我们将根据 PID 控制器的输出来控制电机的速度或位置等参数，以实现平衡控制。为了防止电机失控，通常要为 PID 加限幅。最后需要对系统进行全面的测试，以测试系统的性能和可靠性。

### 4.3. 软件总体实现

在这个嵌入式系统中，我们采用了一种基于 HAL 库(Hardware Abstraction Layer)和 freeRTOS (Real-Time Operating System)的任务分配架构。通过这个架构，我们将不同的功能模块分配给不同的任务，以实现系统的多任务执行。

Chassis Task (底盘任务): Chassis Task 任务专门负责获取和处理底盘控制相关的信息。这可能包括底

盘的速度、位置和方向等数据。底盘任务的主要责任是确保底盘的稳定运行和控制。

**Gimbal Task (云台任务):** Gimbal Task 任务专注于获取和处理与云台控制相关的信息。这包括云台的方向、俯仰和偏航等数据。云台任务负责确保云台的准确定位和控制。

**Imu Task (陀螺仪任务):** Imu Task 任务的主要任务是解算板载陀螺仪的数据。陀螺仪通常用于测量系统的角速度和角度变化。这些数据对于导航和姿态控制非常重要。

**Annex Task (附件任务):** Annex Task 任务负责控制一些附加设备或附件，例如舵机等。这些设备可能用于执行各种任务，如摄像头控制、传感器操作等。

**System Inform Task (系统信息任务):** System Inform Task 任务充当系统的数据协调者和交互中心。它负责更新系统数据、处理交互请求以及加载必要的信息。这个任务确保系统各部分之间的协调和信息传递。

软件实现图如图 3。

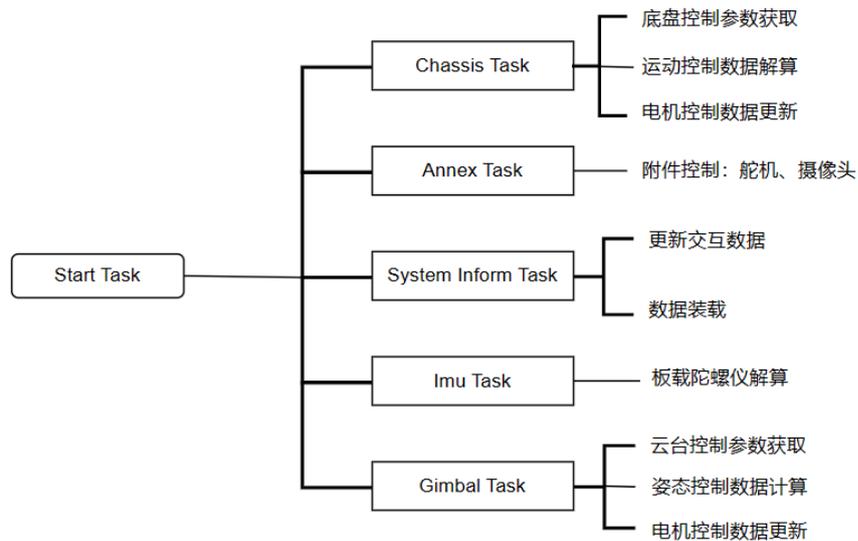


Figure 3. Software implementation diagram  
图 3. 软件实现图

## 5. 神经网络算法设计

### 5.1. 模型结构

本文建立的 GCNet 模型包括特征提取器和分类器两个部分[6]，总体结构如图 1 所描述。在该图中，特征提取器采用了 ResNet101 作为主干部分，共包括 5 个瓶颈层，并在不同的瓶颈层之后加入了关注力机制模块。同时，它还对不同模块提取的特征进行了特征融合，如图 4 中的虚线所示，以从输入中提取图像的特征信息。

$$F_1 = M_e(x)$$

其中， $M_e$  表示特征提取器。 $F_1, y_i$  分类器由两层全连接层和一个 Softmax 分类器组成，对提取到的特征信息进行分类，以得到图像在每个类别下的最终得分：

$$y_i = M_c(F_1), i \in \{1, 2, \dots, n\}$$

其中  $M_c$  表示分类器。

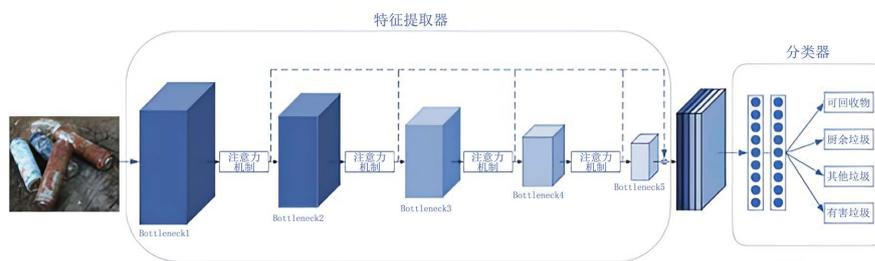


Figure 4. GCNet network structure diagram  
图 4. GCNet 网络结构图

### 5.2. 注意力机制

注意力机制是受人类视觉研究启发而来的，因为人类会根据实际需要选择视网膜内的特定区域进行集中关注，这样可以将有限的处理资源分配到最重要的部分。由于相同类别的垃圾可能在特征表示上存在较大的差异，这可能会妨碍正确分类图像，因此需要准确关注图像中的显著区域。受到这一思想的启发，通过构建注意力机制模块，我们能够使网络模型集中注意力于有利于分类的特征区域，以实现更出色的特征提取功能。该模块的具体结构如图 5 所展示。

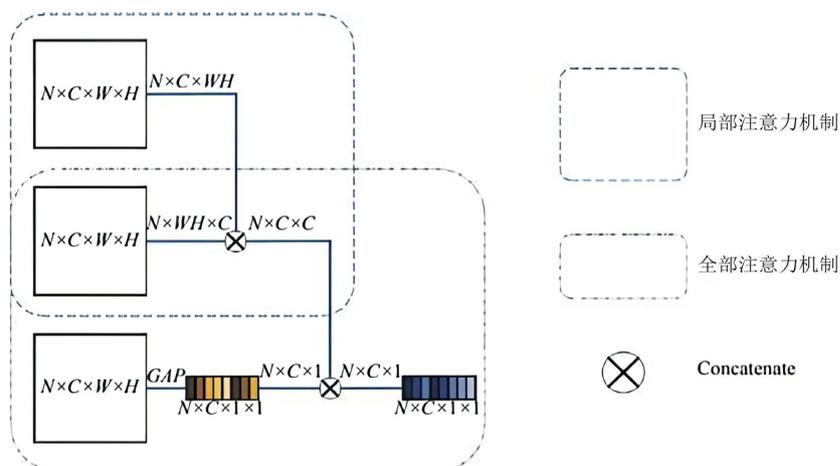


Figure 5. Schematic diagram of attention mechanism  
图 5. 注意力机制示意图

将每个 bottleneck 提取的特征用  $F_i$  表示，利用格拉姆矩阵构建局部注意力机制模块，将  $F_i$  与其转置矩阵  $F_i^T$  相乘，以此得到局部特征  $F_{local}^i$ ：

$$F_{local}^i = F_i^T \cdot F_i$$

这一步骤有助于减小特征图中每个元素与自身的相关性数据，将原本数值较大的元素突显出来，同时抑制了对判断没有帮助的特征。这样可以使每个局部区域的重要性更加显著，有助于更明确地确定类别，同时抑制了对分类判断产生负面影响的特征。

对  $F_i$  进行全局平均池化操作(GAP)，可以保留特征  $F_i$  获取到的空间、语义信息，并获得全局特征  $F_{global}^i$ 。

$$F_{global}^i = GPA(F_i)$$

在  $F_{global}^i$  中表现了特征  $F_i$  中最显著的特征，这也符合所定义的全局注意力机制。

最后，将局部特征  $F_{local}^i$  与全局特征  $F_{global}^i$  相乘，从而获得包含局部和全局信息的总体特征  $F_{fusion}^i$ 。

$$F_{fusion}^i = F_{local}^i \cdot F_{global}^i$$

### 5.3. 特征融合机制

在垃圾分类任务中，同一类别的垃圾通常在外观上存在较大差异，例如可回收物类别包括纸板箱和玻璃杯，尽管它们在外观上有显著不同，但它们属于相同的类别，这会增加分类器的挑战。此外，随着网络层数的增加，单一图像特征可能会失去某些区域的信息，导致模型的分类性能下降。为了解决这个问题，通常采用多尺度的特征融合模块，通过不同卷积核的池化操作来构建，但这会增加模型的计算负担，同时也增加了模型训练的难度。

由于我们的特征提取器中包含了不同的池化操作，因此只需提取不同瓶颈层下的特征即可获得不同尺寸的特征。基于这一基础，本文提出了改进的特征融合机制，以弥补不同尺度特征之间的差异，减少信息损失，提高模型的分类性能，而不需要引入额外的计算负担如图 1 中虚线所示，将每个 bottleneck 经过注意力机制提取后的特征融合：

$$F_1 = w_c * Concat(F_{fusion}^1, \dots, F_{fusion}^5) + b_c$$

其中  $w_c$  为卷积核， $b_c$  为偏置， $Concat$  为融合操作，该操作旨在利用不同尺度特征信息，有效避免了信息丢失，提高了系统鲁棒性。

## 6. 测试与分析

### 6.1. 神经网络算法实验及结果分析

本实验采用大量标注类别的数据集，包括厨余垃圾、可回收物、其他垃圾、和有害垃圾等四个大类。在实验中，采用 4:1 的比例划分数据集，80% 作为训练数据，20% 作为测试数据。此外，为了增强模型的泛化能力和鲁棒性，还对训练样本进行了数据增强处理，包括随机反转，局部翻折、随机裁剪等。

### 6.2. 测试结果

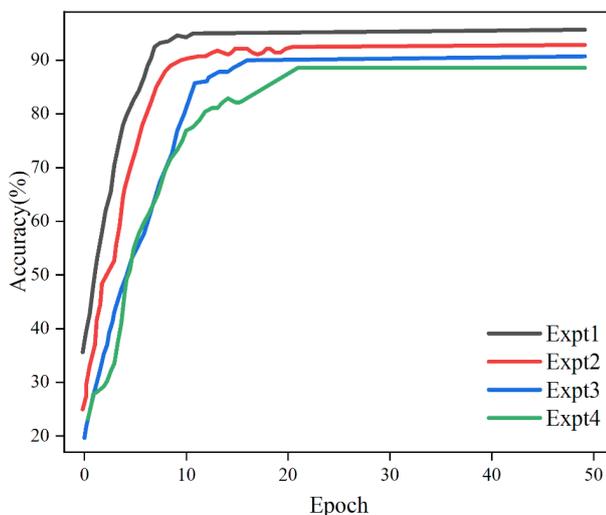
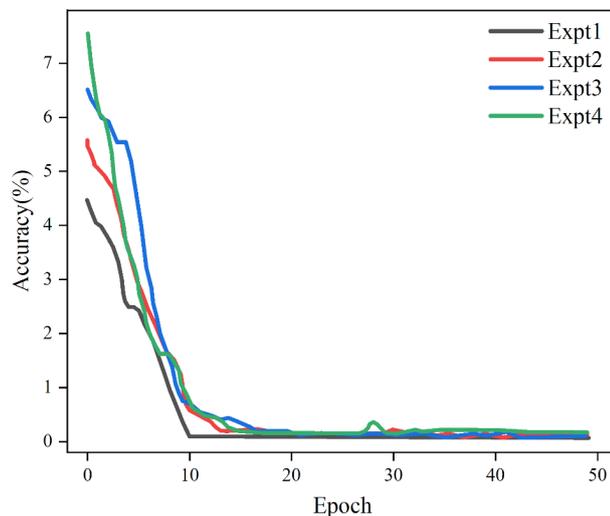


Figure 6. Accuracy iteration curve  
图 6. 准确率迭代曲线



**Figure 7.** Loss function iteration curve  
**图 7.** 损失函数迭代曲线

在训练过程中，对 GCNet 模型通过消融实验以分别验证注意力机制和特征融合机制的功能，实验 1 为含有注意力机制和特征融合机制的模型，实验 2 为只含有注意力机制的模型，实验 3 为只含有特征融合机制的模型，实验 4 为不含有注意力机制和特征融合机制的模型。各个实验的训练过程迭代曲线如图 6，图 7 所示。

从图 6 中可以观察到，实验 1 (Expt1) 的准确率提升和损失值下降速度明显快于其他三个实验，这表明训练过程更快地收敛。实验 2 (Expt2) 和实验 3 (Expt3) 的训练速度几乎相近，而实验 4 (Expt4) 的训练速度最慢。为了进一步验证上述训练的准确性，进行了相应的测试集测试。实验结果如表 1 所示。实验 1 包含了注意力机制模块和特征融合模块，取得了 97.06% 的最佳准确率，同时在各个子类别上都取得了最高的准确率。这说明所构建的模型在泛化能力方面表现出色。在类内差异较大的“其他垃圾”这一类别中，注意力机制和特征融合机制均能够显著提高模型的准确性。

**Table 1.** Accuracy of GCNet model ablation experiment (unit: %)

**表 1.** GCNet 模型消融实验准确率(单位: %)

序号	可回收垃圾	有害垃圾	厨余垃圾	其他垃圾	平均准确率
1	95.81	96.45	98.11	97.85	97.06
2	94.23	90.42	93.21	90.39	92.09
3	93.75	90.61	92.46	87.24	91.02
4	89.75	90.06	87.42	86.22	88.36

本文模型与其它模型对比如表 2 所示，由此可见本文构建的 GCNet 模型平均准确率最高，在各个类别中也取得了最高准确率，说明本文构建的注意力机制和特征融合机制充分提取了有利于垃圾分类的特征，使其结果准确。

**Table 2.** Comparison of accuracy of different model experiments

**表 2.** 不同模型实验准确率对比

模型	可回收垃圾	有害垃圾	厨余垃圾	其他垃圾	平均准确率
GCNet	95.81	96.45	98.11	97.85	97.06

**Continued**

TrashNet	92.78	93.24	93.06	91.56	92.66
CaffeNet	90.53	89.19	93.81	87.43	90.24

**6.3. 测试结果分析**

本网络通过注意力机制和特征融合机制能有效提取图像特征、降低类别差异性带来的影响，并在相关数据集上取得了 97.06% 的平均准确率，相比较于现有的分类算法提高了约 4% 的准确率，满足了要求，具有良好前景。

**7. 结束语**

本文介绍了一种基于神经网络的垃圾分类平衡机器人，该机器人基于卷积神经网络视觉识别技术实现垃圾分类，现有的图形分类算法在垃圾处理领域的应用较少，且存在准确率不足、泛化性能差、处理效率低的缺点。这是一个前沿的研究方向。通过训练神经网络来识别四类垃圾，并与现有图像分类算法作对比，针对现有方法的不足，通过特征融合机制和注意力机制，具有很大的创新性和可研发性。相比于 TrashNet 和 CaffeNet 算法提高了 4% 的准确率。

运用 PID 控制算法并使用，能够实时获取各种传感器数据，对周边环境实时得到信息反馈。该机器人采用了两轮平衡车的设计，实现了优秀的自平衡和控制功能。这种设计不仅可以应用于垃圾分类机器人，还可以应用于其他类型的机器人控制，如移动机器人、机械臂等。这种设计可以帮助机器人在不平坦的地面上保持平衡，提高了机器人的运动性能。总之，该平衡机器人的设计和实现为机器人技术的发展提供了新的思路和方向，推动智能制造发展，运用许多现代科技，包括神经网络、PID 控制和计算机视觉等等，推动了智能制造技术的发展。提高垃圾的处理效率，使得废弃物更容易被回收和处理，且减轻人工负担，帮助人们更方便的垃圾分类。

**基金项目**

本论文由辽宁科技大学大学生创新创业训练计划项目经费支持(编号: S202310146061)。

**参考文献**

- [1] 阮激扬. 基于 YOLO 的目标检测算法设计与实现[D]: [硕士学位论文]. 北京: 北京邮电大学, 2019.
- [2] 段焜. 基于 STM32 的智能车控制算法设计[J]. 电子设计工程, 2023, 31(18): 140-144.
- [3] 朱骏宇. 基于卷积神经网络的图像识别的技术分析[J]. 长江信息通信, 2023, 36(8): 66-68.
- [4] 戎新萍, 徐海璐, 季春天. 基于解耦的工业机器人模糊 PID 抑振研究[J]. 自动化与仪表, 2023, 38(8): 40-45. <https://doi.org/10.19557/j.cnki.1001-9944.2023.08.009>
- [5] 苏继平. 多姿态两轮移动机器人控制方法研究[D]: [硕士学位论文]. 天津: 天津职业技术师范大学, 2023. <https://doi.org/10.27711/d.cnki.gtjgc.2023.000024>
- [6] 刘子路. 基于卷积神经网络的生活垃圾分类和检测[D]: [硕士学位论文]. 焦作: 河南理工大学, 2022.