ECE445

SENIOR DESIGN LABORATORY

# DESIGN DOCUMENT

## Vision-Guided Robotic Arm for Waste Sorting

Team #19

Zihan Zhou [zihanz16]
Dailin Wu [dailin2]
Jinyang Chen [jc129]
Tinghao Pan [tinghao4]

Advisor: Meng Zhang

March 31, 2026

# Contents

# 1. Introduction

## 1.1 Problem Statement

In modern waste management and resource recovery, accurate material sorting is a critical prerequisite for effective recycling. However, in small-scale environments—such as research laboratories, educational facilities, or localized collection points—this process remains heavily dependent on manual labor. Manual sorting is inherently repetitive, inefficient, and susceptible to human error, particularly when operators are required to distinguish between diverse materials like plastics, metals, and paper under continuous or fast-paced conditions.

While large-scale industrial sorting facilities utilize high-speed automated systems, these solutions are often prohibitively expensive, physically massive, and overly complex for localized or low-volume applications. There is a distinct technical gap for a compact, cost-effective, and highly integrated automated system capable of performing precise waste classification and physical separation within a desktop-scale footprint. Current small-scale research setups often lack the seamless integration of real-time computer vision with industrial-grade collaborative robotics.

This project aims to develop a **Vision-Guided Robotic Arm for Waste Sorting** to address these challenges through an intelligent "perception-to-action" pipeline. The system is designed to automate the identification and handling of three primary categories: plastic bottles, aluminum cans, and paper-based cardboard. To bridge the gap between manual labor and industrial automation, the system integrates the following key technologies:

- **Advanced Perception:** Utilizing a camera to capture high-resolution color and depth data, allowing for the precise spatial localization of objects on a moving conveyor.
- **Edge Intelligence:** Leveraging an platform to run deep learning inference for real-time object detection and material classification.
- **Precision Manipulation:** Employing a robotic arm equipped with a vacuum suction system to reliably pick and place diverse objects into designated bins according to their material type.

By consolidating these components into a unified, desktop-scale platform, this project provides an efficient and repeatable solution for automated waste management. The proposed system aims to minimize human intervention, improve sorting consistency, and offer a scalable prototype for future low-cost, smart recycling infrastructure.

### 1.2 Solution Overview & Visual Aid

The proposed solution is an automated sorting station designed to classify and segregate recyclable materials within a **static horizontal workspace**. Unlike conveyor-based systems, this design utilizes an "area-scanning" approach, which is ideal for compact, desktop-scale research and localized waste processing.

The operational workflow begins with the **camera**, which is mounted on a fixed overhead frame to provide a comprehensive top-down view of the entire **Sorting Area**. The camera captures high-resolution color images and spatial depth maps simultaneously. This data is processed by the **Processor**, where a deep learning model performs real-time inference to identify the material type (Plastic, Metal, or Paper) and calculate the precise 3D centroid coordinates of each item scattered in the workspace.

Upon successful detection and localization, the **robotic arm** executes a calculated motion trajectory to the target's position. Using a **vacuum suction end-effector**, the arm picks up the object and transfers it to one of the three designated **Collection Bins** positioned at the periphery of the workspace. The seamless integration of 3D depth sensing and high-precision collaborative robotics ensures that objects of varying shapes, sizes, and orientations can be handled reliably without manual intervention.

### 1.3 High-Level Requirements List

To consider the Vision-Guided Robotic Arm for Waste Sorting system successful, the following three high-level requirements must be met:

**High Classification Accuracy:** The integrated vision system, powered by the NVIDIA Jetson Orin Nano and the Gemini 335 RGB-D camera, must achieve a minimum of 90% accuracy in identifying and classifying three distinct material categories (Plastic, Metal, and Paper) under controlled indoor lighting conditions.

**Precise 3D Localization:** The system must successfully map 2D image coordinates to 3D spatial coordinates with an error margin of less than $\pm 1$ cm. This accuracy is essential for the DOBOT MG400 robotic arm to reliably align its vacuum suction end-effector with the center of mass of objects scattered in the workspace.

**Safety and Reliability:** The system must demonstrate fail-safe operational integrity. This includes a 100% success rate in "safe-state" transitions—where the DOBOT MG400 immediately halts all motion upon detecting an emergency stop trigger or software anomaly.

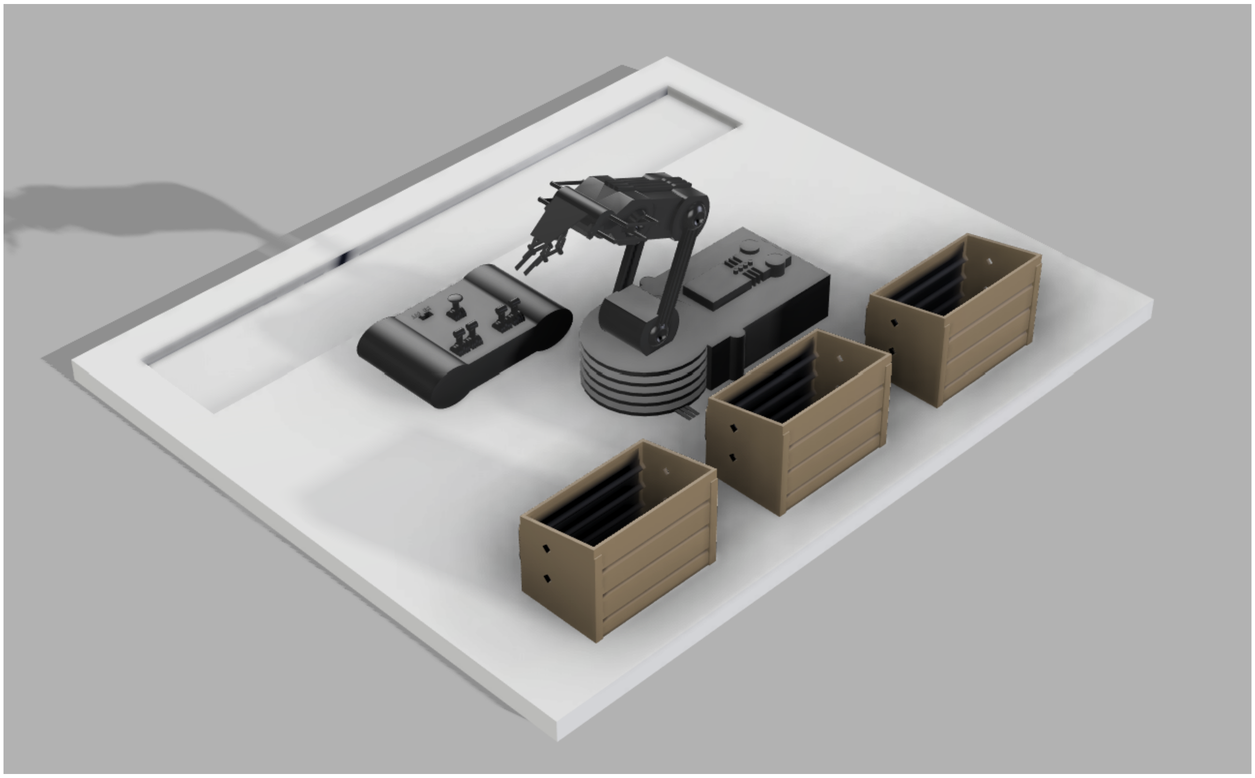

Figure 1: Visual aid of the whole system.

Furthermore, the system must maintain operational stability without "hand-eye" calibration drift for a continuous runtime of at least half hour.

# 2. Design

## 2.1 Overall System Block Diagram

Figure 2 illustrates the overall architecture of the vision-guided robotic arm waste-sorting system. The system is composed of six major functional parts: the object input region, the vision perception module, the classification and localization module, the task planning and control module, the robotic arm with end-effector, and the final sorting bins. These modules are integrated into a complete perception-to-action pipeline that enables automatic identification, grasping, and placement of recyclable objects.

The workflow begins when recyclable objects are placed on the input platform or conveyed into the working area. An RGB-D camera mounted above the workspace continuously captures the scene and sends image and depth information to the vision perception module.

Based on these inputs, the system detects the objects in the scene, determines whether each item belongs to the plastic, metal, or paper category, and estimates its position for grasp planning.

After the target information is obtained, the task planning and control module selects the object to be processed and generates the corresponding pick-and-place command. This command is then sent to the robotic arm controller, which drives the robotic arm and end-effector to move toward the target object, grasp it, and transfer it to the correct sorting bin. Once the object is released, the arm returns to its standby position and waits for the next sorting cycle.

In addition to the main perception and actuation chain, the overall system also includes supporting subsystems such as power distribution, communication interfaces, and emergency-stop protection. These subsystems ensure that the camera, embedded computing platform, robotic arm, and auxiliary actuators operate in a coordinated and safe manner. Therefore, the block diagram not only represents the functional decomposition of the system, but also shows how sensing, decision-making, and physical execution are linked together in an integrated waste-sorting platform.

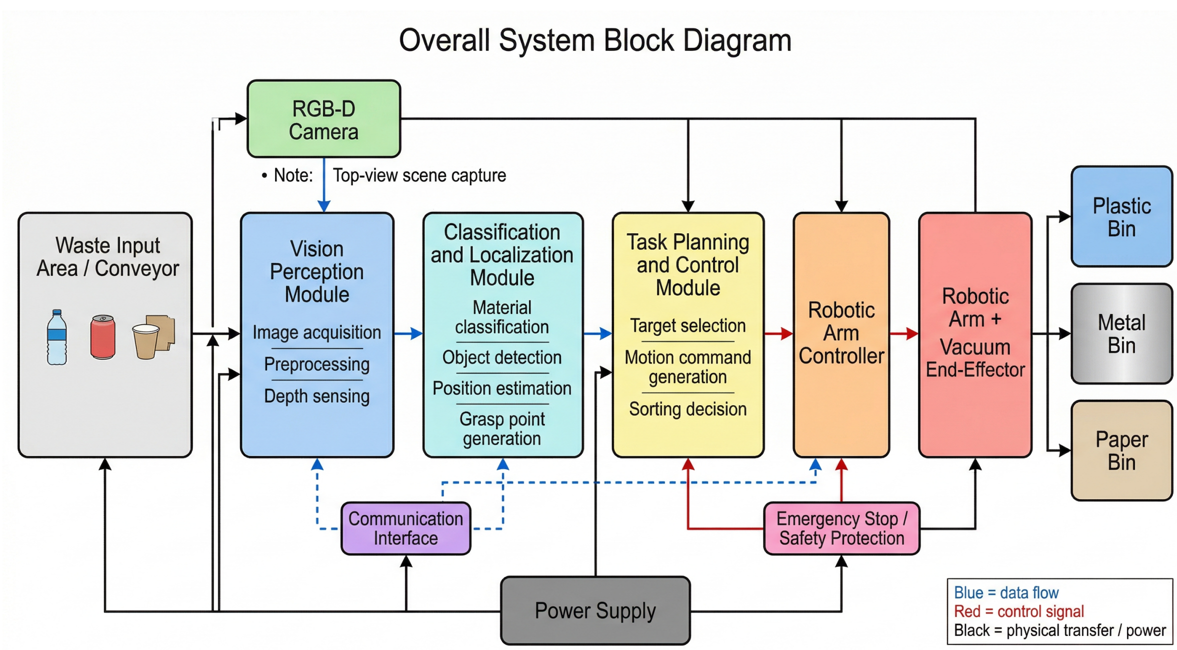


Figure 2: Overall block diagram of the vision-guided robotic arm waste-sorting system.

## 2.2 Vision Perception System

Figure 3 shows the workflow of the vision perception system, which serves as the sensing front end of the waste-sorting platform. Its primary function is to monitor the working area, identify recyclable objects, and provide the spatial information required for robotic grasping. In this project, an RGB-D camera is mounted above the sorting region to continuously capture both color images and depth information. The color channel supports object recognition and material classification, while the depth channel improves object localization and provides useful geometric information for grasp planning.

As illustrated in Figure 3, the perception module processes the input through a structured pipeline that includes image acquisition, preprocessing, object detection, material classification, and position estimation. The system then extracts the target information needed by the downstream controller, such as object category and grasp-related location data, and sends these results to the task planning module. In this way, the vision perception system forms the critical link between environmental sensing and physical sorting, directly affecting the accuracy and reliability of the overall robotic platform.

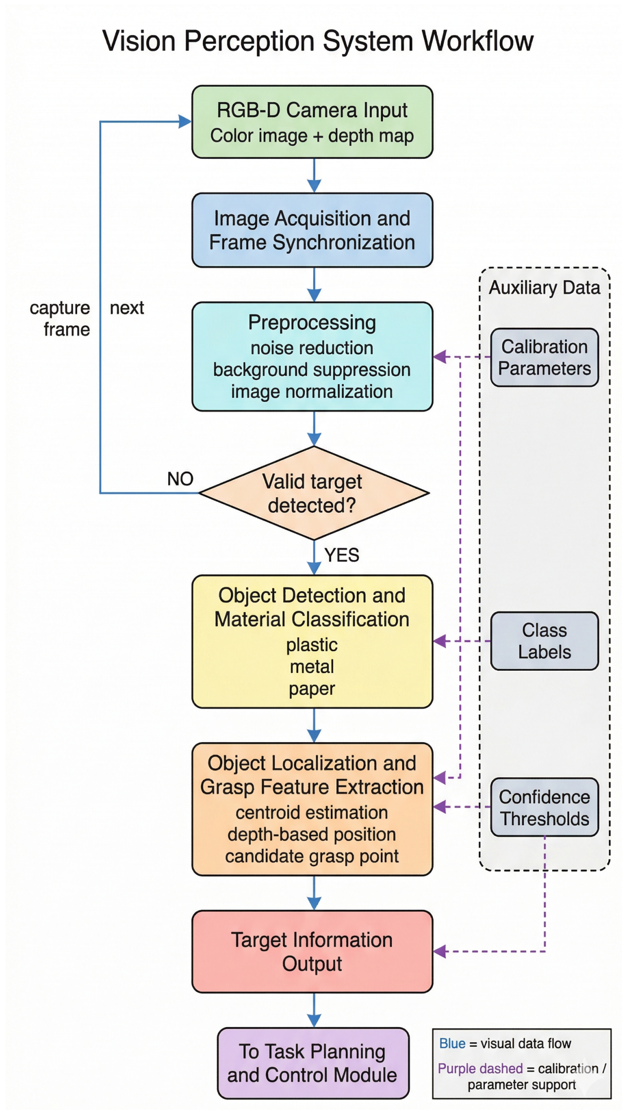


Figure 3: Workflow of the vision perception system for recyclable object identification and localization.

### 2.2.1 Sensing Configuration and Data Acquisition

The camera continuously captures two types of information: RGB images and depth data. The RGB images are mainly used to identify the visible appearance of objects, while the depth information provides basic geometric support for estimating object position in the workspace. In this way, the sensing system does not rely on color information alone, which improves the robustness of subsequent classification and grasp planning.

To ensure stable operation, the camera must remain rigidly mounted with respect to the robot base during the entire sorting process. The sensing region should cover the effective workspace where objects are presented for picking, so that each target can be fully observed before the manipulation stage begins. Therefore, this part of the design focuses on clear scene acquisition, stable sensor placement, and reliable visual input for the downstream processing module.

### 2.2.2 Visual Processing and Target Output

After the scene data are captured, the vision module processes the input to identify recyclable objects and determine their categories. In this project, the output classes are limited to plastic, metal, and paper, which keeps the classification task simple and consistent with the intended hardware demonstration. Once an object is detected, the system estimates its image-center position and combines it with depth information to obtain a basic spatial description of the target. A simple coordinate conversion can then be applied to relate the camera measurement to the robot workspace:

$$\mathbf{p}_r = T_{c\to r}\,\mathbf{p}_c \tag{1}$$

where $\mathbf{p}_c$ is the target position in the camera coordinate frame, $\mathbf{p}_r$ is the corresponding position in the robot coordinate frame, and $T_{c\to r}$ is the calibration transform between the camera and the robotic arm.

When multiple objects appear in the workspace, the system selects one target at a time for manipulation rather than attempting simultaneous sorting. A practical and simple rule is to prioritize the object with valid classification confidence and a reachable position near the center of the working area. The visual subsystem then sends structured target information to the control module, including the object category, estimated position, and confidence score. In this way, the output of the perception stage is converted from raw image data into a compact command-oriented representation that can be directly used for pick-and-place planning.

### 2.3 Robotic Arm and End-Effector System

The robotic arm and end-effector system is the main actuation part of the waste-sorting platform. After the target information is provided by the vision subsystem, this part of the system is responsible for physically executing the sorting task, including approaching the object, grasping it, lifting it, moving it to the designated bin, and releasing it. In this project, a compact tabletop robotic arm is combined with a simple vacuum-based end-effector so that the overall manipulation process remains practical and easy to implement.

Compared with a more complex industrial manipulation setup, this design focuses on lightweight recyclable objects and short-range pick-and-place operations within a fixed workspace. Therefore, the mechanical execution system emphasizes simplicity, repeatability, and compatibility with the perception-guided sorting pipeline. The following subsections describe the robotic arm configuration and workspace arrangement, followed by the design of the end-effector and the basic pick-and-place logic.

#### 2.3.1 Robotic Arm Configuration and Workspace

As a candidate manipulation platform, the system is planned to use a compact tabletop robotic arm for short-range pick-and-place operations in the sorting workspace. As shown in Figure 4, this type of manipulator provides a clearly defined working range and is suitable for transferring light recyclable objects from the input region to the corresponding collection bins. Compared with a larger industrial robot, a tabletop robotic arm is easier to integrate into a laboratory prototype and is sufficient for demonstrating the basic perception-guided sorting process.


Figure 4: Candidate tabletop robotic arm and its representative workspace.

### 2.3.2 End-Effector Design and Pick-and-Place Logic

For the end-effector, the current design plans to use a simple vacuum-based suction mechanism to handle light recyclable objects such as plastic containers, aluminum cans, and paper items. This solution is easier to implement than a more complex multi-finger gripper and is sufficient for a basic sorting prototype. In principle, the suction force can be approximated by

$$F = \Delta P \cdot A \tag{2}$$

where $F$ is the theoretical holding force, $\Delta P$ is the pressure difference, and $A$ is the effective suction area. As long as the available holding force is greater than the object weight with a reasonable safety margin, the object can be lifted and transferred reliably.

During operation, the robotic arm first moves the suction end-effector above the detected target, then descends to the grasping position and activates the vacuum source to pick up the object. After lifting the object, the arm transports it to the designated sorting bin and releases it by turning off the suction. This pick-and-place sequence keeps the manipulation logic simple and matches the overall goal of building an easy-to-understand and practical waste-sorting demonstration system.

### 2.4 Control System

Figure 5 illustrates the overall control architecture of the waste-sorting platform. The control system acts as the coordination layer between the vision subsystem, the robotic arm, and the auxiliary actuators. After receiving the detected object category and position from the perception module, it generates the corresponding commands for arm motion and suction control, thereby linking visual sensing to physical sorting in a simple and reliable execution flow.

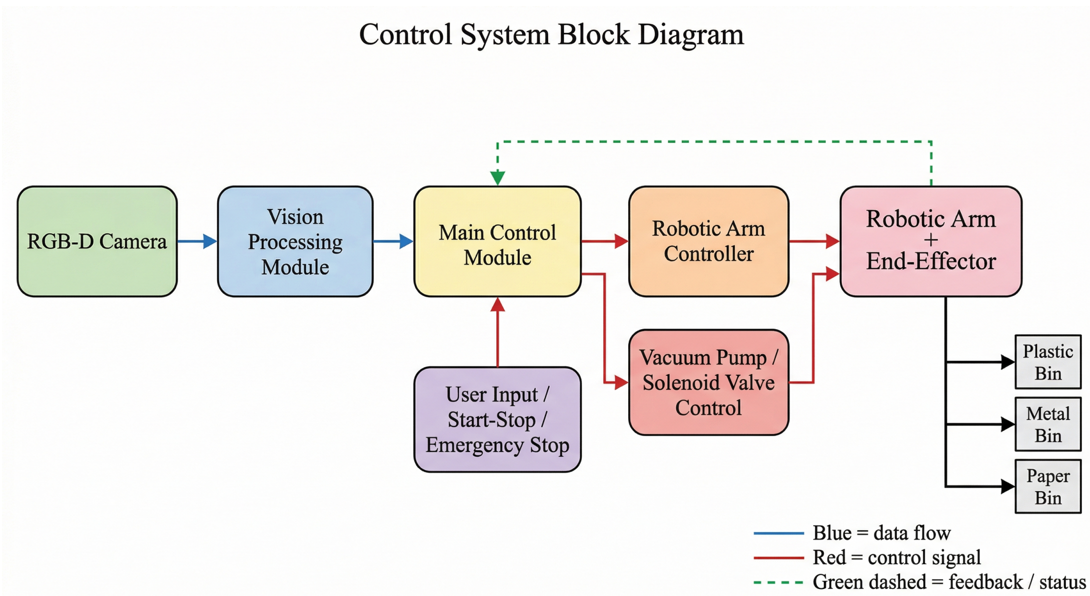


Figure 5: Control system block diagram of the vision-guided robotic arm waste-sorting platform.

#### 2.4.1 Control Logic and Communication

In the current design, the control function is planned to be centered on the Jetson computing platform, which receives image and depth data from the RGB-D camera, runs the visual inference process, and sends sorting commands to the robotic arm and auxiliary devices. To keep the prototype simple, low-speed peripheral components such as the photoelectric sensor, vacuum valve, and start-stop interface can be connected through a USB relay or RS485 I/O module. In this way, the control system does not rely on a complicated multi-controller architecture, but instead uses one main computing unit to coordinate perception, motion triggering, and suction actuation in a clear and practical manner.

The execution logic can be expressed in a simplified form as

$$u_{\text{pick}} = \begin{cases} 1, & c \geq c_{\text{th}} \text{ and } \mathbf{p} \in \mathcal{W}, \\ 0, & \text{otherwise}, \end{cases} \tag{3}$$

where $c$ is the classification confidence, $c_{\text{th}}$ is the confidence threshold, $\mathbf{p}$ is the estimated target position, and $\mathcal{W}$ denotes the reachable workspace of the robotic arm. In other words, the system performs a pick-and-place action only when the detected object is recognized with sufficient confidence and lies inside the valid manipulation region.

### 2.5 Power and Safety Considerations

The waste-sorting platform includes several powered modules, such as the RGB-D camera, the Jetson computing platform, the robotic arm, the vacuum actuation unit, and the conveyor or other auxiliary interfaces. Since this project is intended as a laboratory prototype, the power architecture is kept simple and practical rather than highly customized.

Instead of designing a complicated power network, the current plan is to use standard and modular power supplies for different parts of the system. This choice reduces circuit complexity, improves integration feasibility, and makes the prototype easier to assemble, test, and maintain.

Table 1: Main powered modules in the proposed system.

| **Module** | **Main Function** |
|---|---|
| RGB-D camera | Scene capture and depth sensing |
| Jetson platform | Vision processing and control coordination |
| Robotic arm | Pick-and-place motion execution |
| Vacuum unit | Object suction and release |
| Conveyor / interfaces | Object transport and auxiliary triggering |

Safety is also an important part of the design because the system contains moving mechanical parts and active suction components. During operation, the robotic arm and camera should remain rigidly mounted so that the workspace geometry and calibration remain stable.

The system should only execute a pick command when the detected object lies inside the valid workspace and the visual result is sufficiently reliable. In addition, an emergency-stop or manual shutdown mechanism should be included so that the system can be stopped immediately if abnormal behavior occurs.

Table 2: Basic safety considerations of the prototype system.

| **Safety Item** | **Design Consideration** |
|---|---|
| Rigid mounting | Prevent sensor shift and motion error |
| Workspace limitation | Only operate inside the valid manipulation area |
| Reliable detection | Avoid grasping based on low-confidence results |
| Emergency stop | Allow immediate shutdown in abnormal situations |
| Lightweight objects | Reduce risk during suction failure or drop |

### 2.6 Tolerance Analysis and Design Risks

The proposed prototype is affected mainly by perception error, calibration error, robot positioning error, and suction failure. For a successful sorting action, the system must first detect the correct object, then reach a valid grasping position, and finally place the object into the correct bin. Therefore, the main risk is not a single error source, but the accumulation of several small deviations along the full pipeline.

A simple position-error model can be written as

$$e_{\text{total}} \approx e_{\text{vision}} + e_{\text{cal}} + e_{\text{robot}} \tag{4}$$

where $e_{\text{vision}}$ is the visual localization error, $e_{\text{cal}}$ is the camera-to-robot calibration error, and $e_{\text{robot}}$ is the robotic arm positioning error. To achieve a reliable pick, the total error should remain smaller than the allowable grasp tolerance,

$$e_{\text{total}} < e_{\text{grasp}} \tag{5}$$

This is why the prototype uses a fixed workspace, rigid sensor mounting, and relatively simple object categories.

For the vacuum end-effector, the basic holding condition can be expressed as

$$F = \Delta P \cdot A, \qquad F > mg \tag{6}$$

where $F$ is the suction force, $\Delta P$ is the pressure difference, $A$ is the effective suction area, $m$ is the object mass, and $g$ is gravitational acceleration. In addition, the system should only execute a sorting action when the visual confidence is above a preset threshold,

$$c \geq c_{\text{th}} \tag{7}$$

so that uncertain detections do not trigger incorrect manipulation.

At the current stage, these are mainly preliminary design considerations for the prototype. The main risks may include incorrect classification, grasp failure, or object placement deviation, but their actual impact still depends on later system integration and testing results. As the prototype is built and debugged, the tolerance assumptions and control thresholds can be adjusted further according to real experimental performance.

## 3. Cost

The estimated hardware cost of the Vision-Guided Robotic Arm for Waste Sorting is summarized in Table 3. The current design focuses on a controlled tabletop sorting platform for three recyclable categories—plastic, metal, and paper. Based on the our current budget sheet, the total estimated cost is approximately **2379.1 RMB**. This is below the corrected budget ceiling of **5000 RMB**, leaving enough room for debugging materials, replacement parts, and minor adjustments during integration.

Table 3: Estimated project cost

| Part | Cost (RMB) |
|---|---|
| DOBOT MG400 robotic arm | 1298.0 |
| Gemini 335 RGB-D camera | 239.9 |
| Jetson Orin Nano Super developer kit | 207.0 |
| 512 GB NVMe SSD | 22.9 |
| Mini PVC conveyor belt | 129.9 |
| End-effector pneumatic components | 39.9 |
| Suction cup set and tubing | 17.9 |
| Photoelectric sensor and bracket | 8.8 |
| LED lighting modules | 26.0 |
| Aluminum frame and mounting base | 150.0 |
| 24V switching power supply | 12.9 |
| Emergency stop switch box | 12.0 |
| USB relay / RS485 I/O module | 15.9 |
| Sorting bins | 18.0 |
| Cables and miscellaneous accessories | 50.0 |
| Markers / 3D printed fixtures / mounting parts | 30.0 |
| Debugging reserve | 100.0 |
| **Total Cost** | **2379.1** |

The robotic arm is the dominant cost in the system because it determines the workspace, repeatability, and overall sorting capability. The camera and embedded computing platform are the next most important subsystems, since the project depends on vision-guided classification and robotic control. The remaining components mainly support transportation, actuation, sensing, mounting, and electrical safety. Overall, the current cost is appropriate for a proof-of-concept system that demonstrates a full perception-to-action pipeline, including visual recognition, robotic pickup, and category-based sorting.

## 4. Schedule

Since only two months remain for development, the project schedule must emphasize rapid implementation, subsystem integration, and requirement verification. The revised project scope focuses on a controlled tabletop waste-sorting prototype involving three predefined recyclable categories: plastic, metal, and paper. This narrower scope makes the remaining timeline realistic and achievable within the available time.

| **Week** | **Zihan Zhou** | **Dailin Wu** | **Jinyang Chen** | **Tinghao Pan** |
|---|---|---|---|---|
| Week 1 | Finalize system architecture and overall control logic | Confirm hardware list and purchasing plan | Finalize target categories and vision task definition | Finalize mechanical layout and workspace structure |
| Week 2 | Set up controller, Jetson environment, and communication interfaces | Purchase and assemble power, wiring, and I/O modules | Set up camera and collect initial object images | Assemble robotic arm platform, bins, and support structure |
| Week 3 | Build basic control pipeline between vision and arm commands | Integrate sensors, emergency stop, and power distribution | Develop object recognition / classification pipeline | Install and adjust arm mounting, end effector, and conveyor structure |
| Week 4 | Test robotic arm communication and pick-and-place control | Debug hardware connections and actuator interfaces | Improve recognition accuracy for plastic, metal, and paper objects | Optimize grasping workspace and fixture positions |
| Week 5 | Integrate classification results with motion execution | Assist with embedded control and system-level debugging | Provide object position / category outputs for grasping | Tune mechanical operation, object placement path, and structural stability |
| Week 6 | Build full sorting workflow and category-based placement logic | Verify reliable operation of power and interface modules | Test end-to-end vision performance under repeated trials | Improve grasp success rate and placement consistency |
| Week 7 | Perform repeated system debugging and optimize control timing | Check electrical safety, cable routing, and emergency stop behavior | Refine model robustness and reduce false detections | Refine mechanical reliability and repeated sorting performance |

| | | | | |
|---|---|---|---|---|
| Week 8 | Complete final integration, requirement verification, and presentation preparation | Prepare hardware demonstration and test records | Summarize experimental results and support report writing | Prepare final mechanical demo setup and assist with presentation |

In the first half of the remaining timeline, the priority is to ensure that the core hardware and software modules work independently, including camera-based recognition, robotic arm control, and the physical sorting workspace. In the second half, the priority shifts to full integration, repeated testing, and performance verification. The main schedule risk is system integration, because even when recognition and robotic motion work separately, converting classification outputs into stable and repeatable pickup actions usually requires multiple rounds of debugging. For this reason, the final two weeks are reserved for testing, refinement, and final presentation preparation.

## 5. Ethics and Safety

### 5.1 Ethics

From the ethics perspective, there are three main points we want to emphasize.

First, we need to clearly define the scope of the project. Our system is not intended to be a universal solution for all types of trash. Instead, it is a prototype designed for a controlled environment, focusing only on three recyclable categories: plastic, metal, and paper. We believe it is important to make this boundary clear in both our report and our presentation, so that the project is described accurately and not overstated.

Second, we need to consider the issue of false detection and misuse. As our instructor pointed out earlier, the concept of "trash" in real-world environments is too broad and complex. If the task scope is too large, the system may generate many false positives and behave unreliably. For this reason, we intentionally narrowed the project to a small number of target categories in a controlled workspace. We believe this is a more responsible approach, because it allows us to build a system that is testable, understandable, and realistic within the time available.

Third, we also need to consider privacy and data usage. Since the system uses a camera, we plan to limit image collection to the tabletop workspace only, rather than recording unnecessary information about people or the surrounding environment. Any images or videos collected during testing will be used only for project development and performance evaluation, and not for unrelated purposes.

Overall, we would like to present this project as a small-scale, controlled, and educational prototype for intelligent recyclable sorting, rather than as a complete real-world deployment solution.

### 5.2 Safety

For safety, we mainly consider mechanical safety, electrical safety, and limitations on test objects.

In terms of mechanical safety, the robotic arm and end-effector introduce potential collision and pinch hazards during operation. Our current plan is to clearly define the operating area during testing and demonstrations, and to make sure that team members keep their hands away from the workspace while the system is running. We also plan to include an emergency stop switch so that the system can be shut down immediately if any abnormal motion occurs.

In terms of electrical safety, we will pay close attention to power supply connections, cable routing, and interface reliability. Since the system includes a camera, a Jetson platform, a robotic arm, a power supply, and several auxiliary modules, poor wiring could affect both stability and safety. Therefore, before each test, we plan to check the power connections, communication links, and mechanical mounting conditions to make sure the system can operate safely.

We will also restrict the types of objects used in testing. At the current stage, the system is intended to handle only lightweight and relatively safe recyclable items, such as plastic bottles, aluminum cans, paper cups, and cardboard pieces. Sharp objects, fragile objects, liquid-filled items, contaminated waste, and hazardous materials will not be included in the test set. This is consistent with the project scope and also helps reduce the risk of grasp failure, dropping, or unsafe operation.

Finally, in the control logic, we want the system to behave conservatively. In other words, if the vision result is not reliable enough, or if the object position is uncertain, the system should skip the grasp instead of attempting a risky action. For a course project prototype, we believe that safety and stability are more important than making the system attempt every possible pickup.

## References