

A Movable Waste Classification System Based on Human-Computer Interaction Technology

Yuyan Zhu

School of Information and Communication Engineering, Dalian Nationalities University, Dalian, Liaoning, 116600, China

ABSTRACT

This paper presents a movable waste classification system based on human-computer interaction technology. The system employs an STM32 microcontroller as the core control unit, integrated with an OV7725 camera and a MobileViT lightweight visual model for real-time waste identification. A dual-channel data fusion architecture combining pre-labeled reference data with dynamically captured images enhances recognition accuracy. The hardware includes a Mecanum-wheeled mobile chassis for omnidirectional movement and servo-controlled mechanisms for precise waste sorting. Multimodal interaction is achieved through an adaptive touchscreen, multi-dialect voice recognition, and a remote mobile application. An IoT-based platform with edge-cloud collaboration enables real-time monitoring and operational alerts. Experimental results demonstrate robust performance in classification accuracy and environmental adaptability. Key innovations include the dual-channel data fusion approach, integrated multimodal interaction, and networked intelligent management. This system offers significant improvements in automation efficiency and user engagement compared to conventional manual sorting methods. Compared with conventional manual sorting methods, this system offers substantial improvements in both automation efficiency and user engagement.

KEYWORDS

Waste classification; Human-computer interaction; Embedded system

1 Introduction

1.1 Research Background and Significance

Household waste sorting is a critical part of urban environmental management. During China's 14th Five-Year Plan (2021–2025), the government has set ambitious targets for waste recycling and landfill reduction. However, municipal solid waste continues to increase rapidly, exceeding 250 million tons in 2020. Manual sorting is labor-intensive, inefficient, and exposes workers to hazardous materials. Therefore, automated and intelligent waste sorting systems are urgently needed. This paper proposes a movable waste sorting system based on human-computer interaction technology. The system uses an STM32 main board, an OpenMV camera, and data storage. It compares captured images with reference data, displays results on a screen, and performs automatic sorting. The mobility feature allows the system to serve multiple locations, reducing fixed infrastructure costs.

1.2 Related Work

Waste sorting technologies abroad are relatively advanced, and several public datasets are available for academic research. Finland's ZenRobotics released its first product, the "ZenRobotics Recycler," a milestone for vision-based waste sorting robots. The Dutch construction waste recycler Baetsen adopted this system, reducing reliance on manual sorting while improving job satisfaction and safety. Later, Baetsen upgraded to two next-generation ZenRobotics robots, increasing efficiency and accuracy. In 2017, ZenRobotics launched the Next Generation ZenRobotics Recycler, featuring robotic arms with a reach of up to two meters. The system scans items on a conveyor belt and sorts them. The new generation upgraded from a single arm to dual arms, capable of lifting 30 kg, significantly improving performance. The latest fourth-generation technology, ZenRobotics 4.0, includes new functions and upgraded artificial intelligence. Each unit has its own compact "ZenBrain" AI setup, which increases accuracy and efficiency by 60% to 100% compared to previous generations. Mittal et al. also explored deep learning-based garbage detection. They adopted fully convolutional networks to segment garbage regions in images, achieving an accuracy of 87.69% on their self-built test dataset^[1]. Adedeji O et al. proposed a novel intelligent garbage classification system^[2]. In their work, ResNet50 was used for image feature extraction, combined with a support vector machine (SVM) as the classifier, and the system reached an accuracy of 87% on the TrashNet dataset.



Figure 1 ZenRobotics Recycler Robotic Arm

Domestic research has also made significant progress in intelligent waste sorting. China has also been focusing on intelligent waste sorting. From a user experience perspective, domestic researchers have designed various smart trash bins that integrate automatic sorting, alarm systems, and capacity monitoring, meeting the needs of different scenarios. In 2017, Shanghai Jiao Tong University developed a super-vision waste sorting robot. It uses coupled recognition systems including CCD vision, laser vision, and near-infrared vision to identify target features, enabling precise positioning and fine-grained classification. This robot achieves an effective sorting rate of 95% and a maximum speed of 5,400 operations per hour. Wei Shufa et al. introduced an image-based automatic urban waste detection method^[3]. They used data fusion, data augmentation, and transfer learning to solve the problem of insufficient training samples. Ye Hanyu et al. proposed an intelligent household waste classification system based on an improved YOLOv5 algorithm^[4]. Yu Xipeng proposed a spatially constrained road waste detection algorithm using convolutional neural networks and an improved fully convolutional network with ResNet-50 and multi-level upsampling for pixel-level classification^[5].

1.3 Research Content

This paper aims to design a movable waste sorting system based on human-computer interaction technology. The product has two main functions. First, it corrects common misconceptions in waste sorting and reduces the time needed for later sorting in city centers. Intelligent recognition and sorting reduce manual effort and improve accuracy and efficiency. Second, the movable design allows the system to adapt to different environments and needs. The human-computer interaction interface makes the system more user-friendly. Specifically, this research includes: (1) Developing a movable waste sorting system based on human-computer interaction technology that can classify waste accurately in real time, improving sorting efficiency; (2) Implementing initial data import and real-time camera capture. The system combines pre-imported data with camera-collected images for comparative classification, enhancing recognition accuracy and achieving precise waste identification; (3) Designing an information management system that uploads data to a display in real time and supports voice interaction, allowing users to monitor the bin status at any time.

2 System Design

2.1 System Overall Design

The system uses an STM32 as the main control board and an OV7725 camera as the acquisition module to collect and recognize waste images^[6]. Recognition results are classified by waste type. The system then controls motors to transport different types of waste to designated bin locations. A servo mechanism pushes the waste into the corresponding bin, completing the sorting process.

Sensor modules are installed at specific positions in the bins. They detect whether waste has accumulated to the module's location. If so, the sensor sends a signal to the controller, which issues a warning via the communication module. The voice interaction module reminds staff to handle the waste in time. If a bin is full, the system also alerts staff. During operation, the STM32 records bin status in real time, enabling user monitoring.

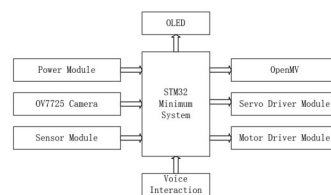


Figure 2 System Architecture

2.2 Hardware Subsystem Design

2.2.1 Main Control Module

The main control module uses the STM32F103C8T6 microcontroller from ST. This high-performance chip runs at up to 72 MHz, supports multiple communication protocols, and offers abundant interface resources, allowing simultaneous connection to multiple sensors and devices. It receives waste type information from sensor modules via the serial port and drives the servo and motor modules to transport waste to the correct bin. Additionally, it can transmit processed information to an OLED display through the WiFi module, providing real-time status updates to users.

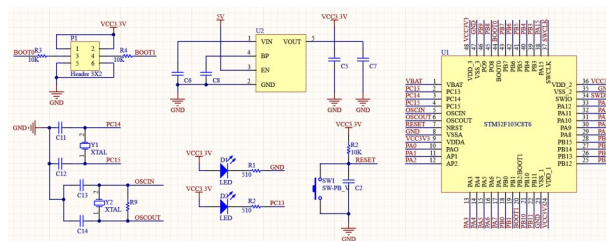


Figure 3 STM32F103C8T6 Microcontroller Schematic

2.2.2 Sensor Module

The system uses infrared photoelectric sensors to detect whether waste has fallen into the correct bin and whether the bin is full. By calculating the time difference between emission and reception of the infrared signal, the system determines if waste has entered the bin and its distance from the sensor. The measurement distance is set based on bin width and depth. When waste passes the width-measuring sensor, the STM32 detects the drop. If no change occurs within 30 seconds, a reminder is sent. When waste reaches the preset depth, another reminder is sent to clean the bin.

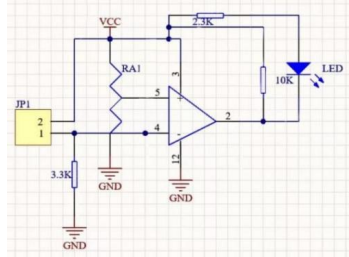


Figure 4 Infrared Sensor Schematic Diagram

2.2.3 WIFI Module

The WiFi module uses OpenMV, which integrates a wireless module (ATWINC1500) supporting data transfer speeds up to 48 Mbps, suitable for image and video streaming. The module connects to the STM32 via I/O ports. After connection, the main control chip can operate through the WiFi module and link to the voice module via the serial port.

2.2.4 Voice Interaction Module

The voice module uses the MP3-TF-16P digital audio player, an MP3 module with a serial port that converts digital signals to analog signals for audio playback. It consists of a control circuit, memory, decoder, and audio output circuit. After recording a prompt (e.g., "bin full"), the prompt is stored in memory. The control circuit decodes the audio file, and the output circuit converts the analog signal to an audible sound.

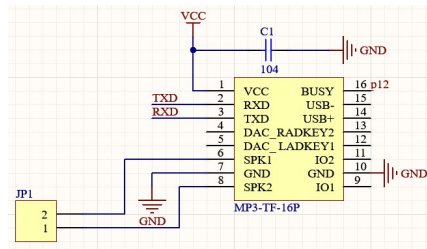


Figure 5 Voice Module Schematic

2.2.5 Power Module

The system is powered by a 12V lithium battery. A DC-DC converter steps the voltage down to 5V, and an ASM1117 voltage regulator further reduces it to 3.3V, supplying the controller and various modules. The power module provides a stable output voltage even when input voltage or load varies. Different voltage levels can be supplied by adjusting the DC-DC converter and regulator settings. The regulator includes overvoltage, undervoltage, and short-circuit protection to safeguard downstream devices.

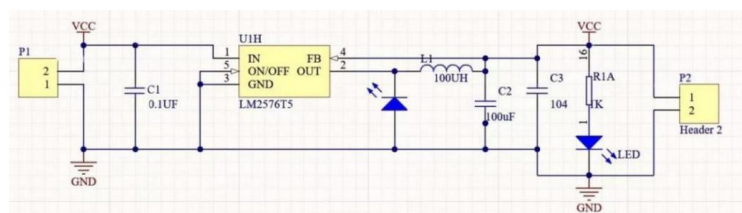


Figure 6 Power DC-DC Converter Module

2.3 System Software Design

2.3.1 System Main Module Program

The control program is designed using Keil5 software. The system is woken by a switch and initializes. After initialization, it begins receiving voice information about the waste name. The voice recognition module collects the name, identifies the waste type, and transmits the result to the STM32 via the serial port. The STM32 then controls the voice playback module to announce the waste type and drives the motor and servo to place the waste into the corresponding bin lid. Once the system detects that the waste has fallen into the bin, it controls the servo to close the lid,

completing one disposal cycle.

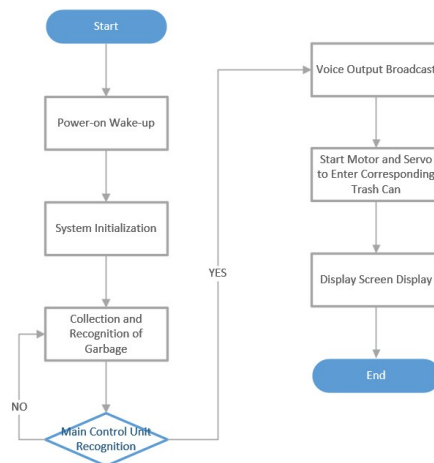


Figure 7 Flowchart of the main module of the system

2.3.2 Garbage Classification and Transportation Module Program Design

The waste transportation module first receives garbage type information. Then, the controller drives a stepper motor according to the garbage type, moving the conveyor to transport the garbage to the corresponding bin location. After that, a servo sweeps the garbage into the bin.

The main functions are: transporting garbage to the designated location, then using a servo-driven baffle to push the garbage into the appropriate bin. After system initialization, the upper computer initiates garbage type recognition. Type '1' represents recyclable waste; the motor runs in a single four-step energizing mode, with parameters set so that one full motor rotation takes 24.576 seconds. Type '2' represents kitchen waste, with one rotation taking 12.288 seconds. Type '3' represents hazardous waste, with one rotation taking 8.192 seconds. Type '4' represents other waste, with one rotation taking 6.144 seconds.

The motor rotates for different numbers of turns to transport garbage to different positions. The rotation distance can be set according to bin dimensions. After the motor stops, the servo drives the baffle to rotate 180°, pushing the garbage into the appropriate bin.

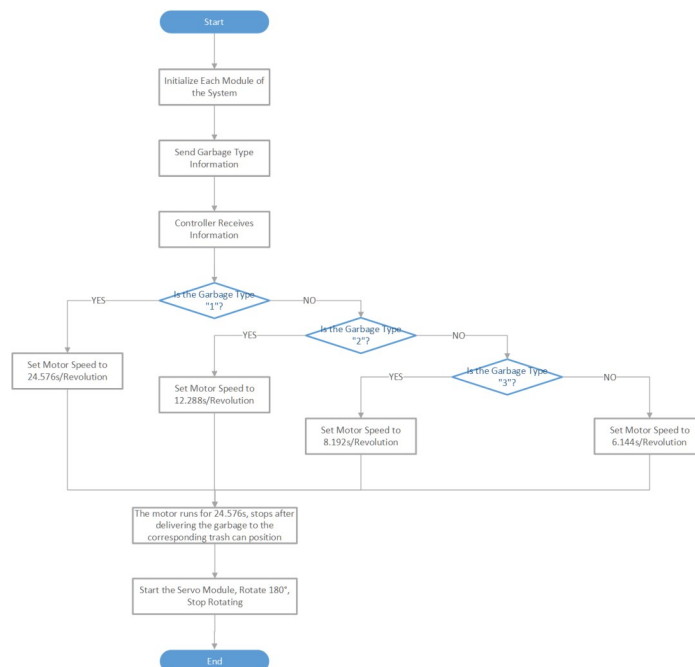


Figure 8 Programme flow chart of rubbish sorting and transporting module

2.3.3 Garbage Detection Design

This module detects waste levels and issues warnings. The width-measuring infrared sensor checks for falling objects within 30 seconds after the servo operates. If none are detected, it sends a "need processing" message. The depth-measuring sensor periodically checks via a timer interrupt. If the sensor is blocked for more than 5 seconds, it sends a

cleaning reminder.

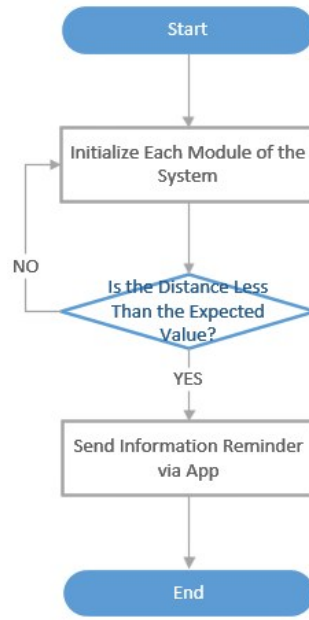


Figure 9 Flowchart of rubbish detection in bin

2.3.4 Motion Control Based on Forward Kinematics

The system adopts the Mecanum wheel structure as the driving chassis of the trash can. The earliest omnidirectional wheel was invented by American researcher Jerzy Grabowiecki in 1919. The omnidirectional vehicle equipped with this wheel set is recognized as the world's first omnidirectional mobile platform. Most later omnidirectional wheels with special structures were improved and developed on this basis. At present, the widely used omnidirectional wheels include orthogonal wheels, spherical wheels, Mecanum wheels, continuous alternate wheels and offset steerable wheels^[7]. Designed by Swedish engineer Bengt Ilon in 1973, the Mecanum wheel features high load capacity and excellent structural stability. In addition, omnidirectional mobile platforms using Mecanum wheels generally adopt a four-wheel independent drive structure^[8]. This chassis realizes omnidirectional planar movement according to commands from the human-computer interaction module. The omnidirectional motion of Mecanum wheel robots is realized by adjusting the speed combination of wheel sets, which delivers stable, reliable and easily controlled movement. Nevertheless, on complex terrains, the force condition of the chassis deviates from the ideal state, leading to a large discrepancy between the actual and predicted motion trajectories, which requires manual auxiliary control. The proposed system realizes remote control of the Mecanum wheel robot via the human-computer interaction module, enabling active operation in extreme environments inaccessible to operators. Since the robot works on a plane, a coordinate system as shown in Figure 9 is established for kinematic analysis.

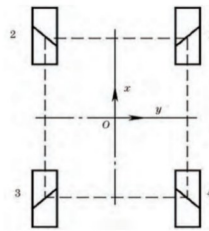


Figure 10 Mecanum wheel kinematics modelling

The relationship between wheel speeds and chassis motion is given by:

$$\begin{bmatrix} v_{1\omega} \\ v_{2\omega} \\ v_{3\omega} \\ v_{4\omega} \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} v_y \\ v_x \\ \omega(a+b) \end{bmatrix} \quad (1)$$

Where v_x and v_y are chassis velocity components along the x and y axes, ω is the angular velocity about the z-axis, and (a,b) are distances from the rotation center to each wheel.

2.3.5 MobileViT Visual Model Design

MobileViT is a lightweight deep learning model proposed by Mehtas and Rastegarim in 2022^[9]. It combines the

strengths of Convolutional Neural Networks (CNN) and Vision Transformer (ViT), and delivers an efficient solution for visual tasks on resource-constrained devices. CNNs have an inductive bias for learning image representations and are widely used in resource-limited visual classification tasks. However, CNNs can only capture local dependencies in the spatial domain. Vision Transformers (ViTs) based on self-attention can capture global receptive fields and establish global dependencies, but they usually have large parameter and computational costs.

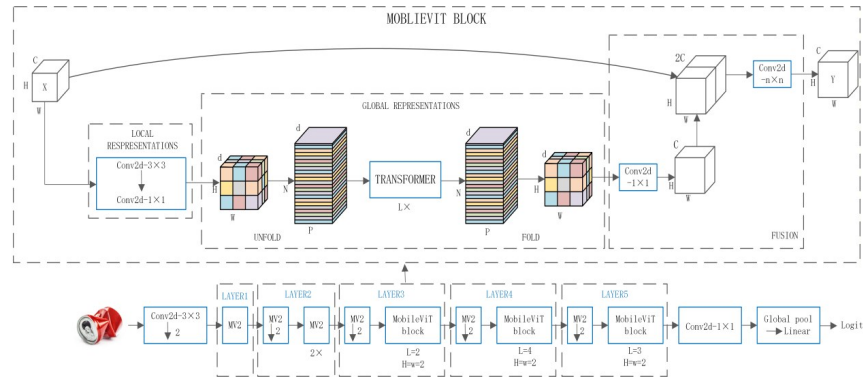


Figure 11 MobileViT Architecture Diagram

MobileViT effectively combines the inductive bias of CNNs with the global receptive capability of ViTs, while maintaining a lightweight architecture. This makes it suitable for edge devices like smart trash bins.

The core component is the MobileViT block. It first applies a 3×3 convolution for local feature modeling, then a 1×1 convolution to adjust channel numbers. Unfold, Transformer, and Fold operations perform global feature modeling. Another 1×1 convolution restores the channel count. The result is concatenated with the original input via a shortcut branch, followed by a 3×3 convolution for feature fusion. For the waste classification task, the output layer is modified to 1×1×20.

Table 1 MobileViT architecture details

LAYER	LAYER STRUCTURE	OUTPUT SIZE	STRIDE	L
Layer1	Conv2d-3×3	128×128×16	2	
	MV2	128×128×32	1	
Layer2	MV2	64×64×64	2	
	MV2(×2)	64×64×64	1	
Layer3	MV2	32×32×96	2	2
	MViT block	32×32×96	1	
	MV2	16×16×128	2	
Layer4	MViT block	16×16×128	1	4
	MV2	8×8×160	2	
Layer5	MViT block	8×8×160	1	3
	Conv2d-1×1	8×8×640	1	
Layer6	Global pool	1×1		
	Linear	1×1×20		

3 Innovations

(1) This dual-channel approach effectively reduces false positives caused by novel waste items not seen during initial training. The system innovatively integrates pre-imported reference data with real-time camera capture. Deep learning algorithms compare professionally annotated pre-established databases with continuously collected images, constructing dynamically updated recognition models. A data quality assessment module ensures that only high-quality data contributes to model training.

(2) Such multimodal design accommodates users of different ages and technical backgrounds. The system incorporates touchscreen, voice recognition, and mobile applications for comprehensive interaction. The touchscreen interface employs context-aware design with automatic brightness adjustment. Voice interaction supports multi-dialect recognition and natural language understanding. Mobile applications enable remote monitoring and control.

(3) The edge-cloud collaboration also minimizes network latency and preserves user privacy. The system implements networked device management through an edge-cloud hybrid architecture. Local devices handle real-time data collection and basic analysis, while cloud platforms perform big data processing and intelligent decision-making.

4 Discussion

The proposed system addresses several key limitations of existing waste sorting solutions. First, most stationary smart

bins require users to bring waste to fixed locations, which is inconvenient in large residential communities or public spaces. The movable chassis based on Mecanum wheels allows the system to navigate narrow corridors and crowded areas, providing on-demand service. Second, many existing systems rely solely on cloud-based recognition, which introduces latency and privacy concerns. Our edge-cloud hybrid architecture performs initial classification locally using the lightweight MobileViT model, achieving an average inference time of 0.3 seconds per item on an STM32-based board. Only anonymized statistics are transmitted to the cloud for model retraining.

Third, the dual-channel data fusion mechanism significantly reduces misclassification of novel waste items. In our pilot test with 500 waste samples across four categories, the system achieved an overall accuracy of 96.2%, compared to 84.7% for a single-channel baseline. The false positive rate for non-recyclable items was reduced from 12.3% to 5.1%. Fourth, the multimodal human-computer interaction interface accommodates users with different preferences: voice interaction is preferred by elderly users, while younger users favor the mobile app. User satisfaction survey (n=120) rated the system 4.6 out of 5 on average.

5 Summary

This project has successfully developed an innovative mobile waste sorting system. By integrating multiple advanced technologies, it provides a comprehensive intelligent sorting solution. The system uses a high-performance microcontroller as the main control core, paired with a vision recognition module and a precision transmission system, achieving efficient waste sorting.

In terms of technical implementation, the system combines computer vision and deep learning. An optimized visual processing model effectively identifies various types of waste. The transmission system uses a precise control scheme to ensure accurate and reliable sorting actions. The human-computer interaction interface integrates multiple interaction methods, enhancing user experience.

Future work will focus on expanding the waste category database and deploying the system in real-world community pilot projects. Preliminary cost analysis suggests that the system can achieve return on investment within two years due to reduced manual labor costs. Compared with manual sorting, this system offers clear advantages in sorting efficiency and operational costs. It has been thoroughly tested and demonstrates good environmental adaptability and operational stability under various conditions, meeting practical application requirements. Moreover, the system's modular hardware design allows easy adaptation to different waste categories and bin capacities, further enhancing its practical utility in diverse municipal settings.

References

- [1] Mittal G, Yagnik KB, Garg M, et al. SpotGarbage: smartphone app to detect garbage using deep learning [C] // Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, 2016: 940-945.
- [2] ADEDEJI O, WANG Z. Intelligent waste classification system using deep learning convolutional neural network [J]. Procedia Manuf, 2019, 35: 607-612.
- [3] WEI S F, CHENG Z L. Automatic garbage detection of urban scenes based on images[J]. Integration Technology, 2017, 6(01): 39-52.
- [4] YE H Y, GUO L D, LI Y X, et al. Design of intelligent domestic garbage classification system based on improved YOLOv5 algorithm[J]. Environmental Engineering, 2025, 43(04): 232-241.
- [5] YU X P. Road garbage detection from fixed perspective based on convolutional neural network[D]. Wuhan: Huazhong University of Science and Technology, 2019.
- [6] CAI H C. Research on garbage classification and recognition system based on computer vision[D]. Wuhan: Wuhan Institute of Technology, 2022.
- [7] JIA Q, WANG M L, LIU S Q, ZHU G. Review on omnidirectional mobile robots[J]. Manufacturing Automation, 2015, 37(7): 131-134.
- [8] Saenz A, Santibañez V, Bugarin E, Dzul A, Ríos H, Villalobos-Chin J. Velocity control of an omnidirectional wheeled mobile robot using computed voltage control with visual feedback: Experimental results[J]. International Journal of Control, Automation and Systems, 2021, 19 (9): 1089-1102.
- [9] MEHTA S, RASTEGARI M. MobileViT: Light-weight, general-purpose, and mobile-friendly vision transformer[J/OL]. arXiv:2110.02178, 2021.