World Journal of Innovation and Modern Technology, Vol. 8, Issue 5, (May) 2025
ISSN 2682-5910
DOI: 10.53469/wjimt.2025.08(05).14

Design and Implementation of Intelligent Speech Recognition Classification Trash Can Based on STM32 Single Chip Computer

Zhimo Cai, Zhihuang Jiang

Guangdong Peizheng University, Guangzhou 510830, China

Abstract: ***This paper designs and develops a new type of smart trash bin, and elaborates on the composition structure, design concept, and implementation technology of the system modules. The smart trash bin can achieve intelligent classification of garbage, automatic lid opening and guiding for disposal, detection of fullness level, and automatic odor removal through means such as voice recognition and ultrasonic ranging. These intelligent functions improve garbage recycling efficiency, prevent bacterial transmission, ensure hygiene and safety, save labor costs, and protect the social environment.***

Keywords: Intelligence; Speech recognition; Garbage sorting; Brim over.

1. INTRODUCTION

China is taking a big step towards urbanization development, with the urban population growing day by day. With the development of the economy and the gradual improvement of people's living standards, the total amount of urban waste production is also rapidly increasing. In the context of garbage classification, efficient disposal of garbage is urgently needed. At present, the treatment of urban waste mainly adopts the methods of open-air stacking and sanitary landfill. However, the scarcity of land resources and resource waste limit the effectiveness of traditional waste treatment methods. If garbage can be classified correctly and effectively recycled and reused, it can alleviate or alleviate the shortage of land resources and environmental pollution to a certain extent, while also playing a role in conserving resources and reducing waste. Currently, some researchers have attempted to develop and design intelligent garbage bin systems using a series of IoT technologies to address issues such as efficient classification and recycling of garbage.

After a thorough understanding of user needs and market research analysis, this article designs and develops an intelligent speech recognition garbage bin based on STM32 microcontroller. It is equipped with hardware such as human infrared sensors, ultrasonic ranging sensors, odor sensors, and ozone generation devices to achieve intelligent recognition and classification of garbage, automatically open the lid to guide garbage disposal, automatically detect the overflow degree of the garbage bin, and remove odors. This system adopts modular design, with low development cost and high practicality. It uses technological means to overturn traditional garbage bins and provides intelligent and humanized services to make people's lives more convenient, comfortable, and hygienic. In healthcare infrastructure, Xu [1] proposed a graph convolutional network (GCN)-based approach to optimize the structural and functional design of medical facilities, enhancing sustainability. Similarly, Liu et al. [2] developed a capsule neural network framework combined with a modified spring search algorithm to improve the control efficiency of spider-like medical robots. Complementing these innovations, Feng et al. [3] explored green building technologies for old building renovations, while Yang [4] integrated large language models (LLMs) and knowledge graphs for medical text mining, improving knowledge extraction efficiency. Beyond healthcare, AI-driven solutions are reshaping urban development. Chen [5] introduced geospatial neural networks to enhance smart city planning through location intelligence, whereas Wang [6] optimized last-mile delivery logistics using AI algorithms. Legal and economic considerations are also being addressed, as Wang [7] examined the balance between enterprise naming rights and prior legal protections, and Li, Evans, and Zhang [8] developed interactive data exploration tools for smart city analytics from a user-centered perspective. Financial applications are expanding rapidly, with Yang, Zhang, and Wang [9] proposing an AI-based economic cycle prediction model using big data, while Yang et al. [10] integrated LLMs for real-time cross-asset risk monitoring in equity, fixed income, and currency markets. Further refining financial strategies, Yang et al. [11] introduced dynamic hedging techniques in derivatives markets using LLM-driven sentiment and news analytics. Industrial and security applications also benefit from AI advancements. Zhao et al. [12] optimized steel production scheduling via deep learning, achieving significant efficiency gains. In cybersecurity, Li et al. [13] improved image steganalysis

through active learning and hyperparameter optimization, enhancing detection accuracy. Medical AI continues to evolve, with Yuan [14] leveraging contrastive multimodal learning to synergize text and image data for chest X-ray analysis. Finally, Ge, Zhang, and Meng [15] demonstrated AI's role in urban planning and green building technologies, presenting global case studies that highlight sustainable development strategies.

2. FUNCTIONAL DESIGN OF INTELLIGENT SPEECH RECOGNITION FOR GARBAGE BIN CLASSIFICATION

The main functions of the intelligent voice recognition garbage bin include voice recognition of garbage types and automatic opening of the lid to guide garbage classification and disposal, overflow detection and odor treatment, automatic battery life, etc. The intelligent speech recognition classification garbage bin system is based on the application of STM32 microcontroller in mechanical motion control and adopts modular design. Its main system modules consist of the following two parts.

2.1 Voice recognition of garbage types, automatic lid opening guidance for garbage classification and disposal

When selecting the main chip, considering the diverse usage environments of the intelligent voice classification trash can, the need for convenient chip replacement, and low cost factors, the STM32F103 "enhanced" series with powerful functions, low power consumption, and high cost-effectiveness is chosen as the main chip of the system [1]. The normal clock frequency is generally 36MHz, while the STM32F103 clock frequency can be increased to 72MHz, which is the most cost-effective among similar microcontroller products, and has the advantages of high accuracy, good device stability, and strong reliability. The STM32F103C8T6 microcontroller is mainly composed of a core microcontroller ARM Cortex-M3, which is mainly used for controlling external circuits and executing programs. Using the STM32F103C8T6 chip as the control core for the garbage bin system can greatly improve the performance of speech recognition garbage classification.

The speech recognition module uses the LD3320 chip because it has a low market price and is a high-precision chip that can recognize speech with high performance. The LD3320 chip is a non-specific human speech recognition chip, which is a speech recognition technology based on "key word list" recognition technology and ASR (Auto Speech Recognition) technology. After processing the voice signal through the chip, the signal is transmitted to the microcontroller, which then processes it and sends the signal to the external circuit, providing accurate voice recognition for the intelligent voice sorting garbage bin.

Automatically activate speech recognition technology and use infrared to receive and send commands. The diode of TCRT5000 sensor continuously emits infrared until the emitted infrared is not reflected back to the sensor or the signal is not clear enough when it is reflected back. At this time, the photosensitive triode will be in a closed state, and the output terminal of the module will be at a high level, indicating that the diode will remain off; When the human body approaches the detectable range, infrared rays are quickly reflected and transmitted back to the sensor, and the signal induction is strong enough. The photosensitive transistor saturates, and the output terminal of the module will switch to a low level, indicating that the diode will light up, thus automatically turning on the speech recognition function.

The five wire four phase stepper motor drive module is used to control the opening and closing of the garbage can lid. The five wire four phase stepper motor is affordable and has stable performance. Its main control chip controller is STM32, and its main function is to grasp the required speed profile. The driver amplifies the pulse signal to drive the winding current of the stepper motor, and sets its rotation angle to 90 ° to meet the needs when opening the garbage can lid.

2.2 Overflow degree detection and odor treatment

The overflow detection mainly uses HC-SR04 ultrasonic ranging sensor to monitor the overflow degree of the garbage bin. By detecting the distance from the garbage to the top of the garbage cover through the ultrasonic sensor, the overflow degree of the garbage bin can be accurately calculated to solve the problem of garbage overflow.

During the use of garbage bins, odor is a common phenomenon. MQ-135 odor sensor and ozone generator are used to remove the odor. When the odor is generated, the odor sensor senses the odor and turns on the ozone generator to treat it.

3. IMPLEMENTATION OF INTELLIGENT SPEECH RECOGNITION FOR GARBAGE BIN CLASSIFICATION

3.1 Infrared sensor detects whether someone is approaching the device, and automatically activates the device function when approaching

The control component switch is mainly determined by dual signals of infrared radiation and natural light. In infrared sensing switch components, the optical principle of Fresnel lens is adopted. This special lens can accurately sense the "high sensitivity zone" and "blind zone" in alternating scenes, which can expand the receiving range of the detector and improve the detection sensitivity. The sensing device is mainly a human body pyroelectric infrared sensor. The highly sensitive infrared detection element can accurately sense the signals emitted by the human body, and can detect whether a person is approaching without contact. It can also convert and amplify the output voltage signal to achieve the automatic driving function of the control circuit controlled by human body thermal sensing. The sensor has two detection elements, and the wiring method between the detection elements is reverse polarity series connection, effectively avoiding the problem of excessive heat leading to abnormal operation.

3.2 Using STM32 as the main control board, the recognition module receives garbage classification information

The speech recognition garbage classification module adopts the STM32 microcontroller with the model number F103C8T6 as the microcontroller unit (MCU). The STM32F103C8T6 main controller is based on the ARM Cortex-M3 core microcontroller, which mainly controls external circuits and executes command programs. Using STM32F103C8T6 as the main controller of the garbage classification system can integrate the functions of timers, counters, analog-to-digital converters and other modules, greatly improving the performance of speech recognition garbage classification.

The principle of speech recognition algorithm is achieved through methods such as endpoint detection, feature extraction (MFCC parameter extraction), and recognition matching (dynamic time warping algorithm DTW). This system mainly consists of four aspects, namely speech signal preprocessing, endpoint detection, feature extraction, template matching, etc.

3.3 LD3320 intelligent speech recognition broadcasting system realizes accurate classification broadcasting of corresponding garbage

Input garbage information into the LD3320 chip through a microphone, perform frequency spectrum analysis and extract features inside the microcontroller chip to achieve speech recognition function; At the same time, the LD3320 chip can also perform speech recognition on keywords through the keyword manual. Simply input the keywords into the LD3320 chip and they can be successfully recognized in subsequent processes. In MCU programming, by setting the registers of the LD3320 chip, easily recognizable key words such as "hello" can be dynamically passed into the chip, and the LD3320 chip can recognize pre-set key words. All key words can be words, idioms, or sentences. This speech recognition system can dynamically change the keywords in the list during runtime, so that only one system is needed to support multiple recognition methods, and users do not need to do related recording recognition training. It has the following characteristics:

(1) Feature extraction:

By using signal processing techniques, the original speech waveform signal is converted into a sequence of feature vectors as input for the speech recognition system. Speech signals are short-term, stable, and random processes that require frame segmentation operations. Commonly used 16KHZ/8KHZ sampling, typically using 25ms frame length and 10ms frame shifted speech for frame segmentation, followed by feature extraction. Through feature extraction, the time-domain waveform of the speech signal is first converted into a frequency-domain signal, and then processed through signal processing methods such as filtering and discrete cosine transform to obtain the final

speech features. In speech recognition, commonly used features include short-term spectral features such as MFCC, PLP, and fbank.

(2) Acoustic model:

After obtaining the speech features, it is necessary to characterize the speech. The acoustic model is a probabilistic model that has the property of characterizing speech segments. The key technology in speech recognition systems is acoustic modeling, which plays an important role in the performance of the system. In the development of speech recognition over the years, hidden Markov models have become the most common acoustic models. The Hidden Markov Model mainly consists of: (A) Modeling units, each basic speech modeling unit (word, factor, three factor) corresponds to HMM; (B) Topology structure: from left to right, states 1-5, each HMM state corresponds to a basic speech segment; (C) In HMM, there are two sequences: the state sequence and the observation sequence.

(3) Language model:

The pronunciation dictionary is the link between language models and acoustic models. It includes the mapping of words to acoustic modeling units (phonemes). Pronunciation dictionaries are usually compiled by pronunciation experts through summarization. The language model is a probabilistic model that can be combined with a decoder to guide search algorithms and distinguish similar pronunciations. The most commonly used is the N-gram language model.

(4) Decoder:

The ability to simplify decoding algorithms can be divided into two main types: dynamic decoding algorithms and static decoding algorithms. Static decoders have faster decoding speeds than dynamic decoders. The decoding algorithm can be divided into one pass decoding and multiple pass decoding. One pass decoding is to use a set of models to unlock the speech signal, and the result is obtained after one pass decoding. The so-called multi pass decoding often uses a small model to first decode the word lattice containing multiple paths, and then uses a larger model to search on the word lattice. Generally speaking, multiple decoding iterations can achieve better recognition results. The most common technique in static decoders is the Weighted Finite State Machine (WFST).

3.4 Use HC-SR04 sensor ultrasonic distance measurement to monitor the overflow degree of garbage cans

Equipped with automatic overflow detection function, a remaining capacity display is set outside the garbage bin. When the garbage bin is filled, the lid will no longer automatically open and relevant personnel will be contacted immediately to ensure timely cleaning of the garbage; Simultaneously equipped with odor detection function, when the odor reaches the predetermined value, the ozone generation module is automatically activated to eliminate the odor. When the trash can overflows. The HC-SR04 sensor emits ultrasonic waves and immediately returns when encountering obstacles. The actual distance measured at this time is the distance between HC-SR04 and the garbage, thus determining the occurrence of overflow. When overflowing, inform the staff through automatic phone calls.

3.5 Use MQ-135 odor sensor and ozone generator to remove odors

When the odor sensor detects an odor and reaches the threshold set by the sensor, it will issue a command to the Arduino. Upon receiving the command, the Arduino will start supplying power and turn on the ozone generation module to eliminate the odor. When the odor level returns below the threshold, it will automatically power off and turn off the ozone generation module.

4. CONCLUSION

This design uses STM32 microcontroller as the main control chip, combined with hardware such as human pyroelectric infrared sensor, HC-SR04 ultrasonic ranging sensor, odor sensor, and ozone generation device, to achieve intelligent recognition and classification of garbage, induction of human body automatic lid opening guidance for garbage disposal, automatic detection of garbage bin overflow degree and removal of odor, and other intelligent functions. This system has low development cost, high practicality, effectively improves the efficiency of garbage collection, and has good application prospects. This system can improve the health and safety guarantee

for people's lives, provide a material basis for environmental protection for society, and enable the garbage disposal industry to better adapt to the intelligent development of society.

REFERENCES

[1] Xu, Haoran. "Sustainability Enhancement in Healthcare Facility Design: Structural and Functional Optimization Based on GCN." (2025).
[2] Liu, Z., Jian, X., Sadiq, T., Shaikh, Z. A., Alfarraj, O., Alblehai, F., & Tolba, A. (2025). Efficient control of spider-like medical robots with capsule neural networks and modified spring search algorithm. Scientific Reports, 15(1), 13828.
[3] Feng, Zhang, et al. "Research on old building renovation strategies by using green building technologies." 2024 6th International Conference on Civil Architecture and Urban Engineering (ICCAUE 2024). Atlantis Press, 2025.
[4] Yang, Jinzhu. "Integrated application of llm model and knowledge graph in medical text mining and knowledge extraction." Soc. Med. Health Manag 5 (2024): 56-62.
[5] Chen, J. (2025). Geospatial Neural Networks: Enhancing Smart City through Location Intelligence.
[6] Wang, J. (2025). Smart City Logistics: Leveraging AI for Last-Mile Delivery Efficiency.
[7] Wang, H. (2024). The Restriction and Balance of Prior Rights on the Right of Enterprise Name.
[8] X. Li, L. Evans, and X. Zhang, “Interactive data exploration for smart city analytics: A user-centered perspective,” 01 2025.
[9] Yang, W., Zhang, B., & Wang, J. (2025). Research on AI Economic Cycle Prediction Method Based on Big Data.
[10] Yang, J., Tang, Y., Li, Y., Zhang, L., & Zhang, H. (2025). Cross-Asset Risk Management: Integrating LLMs for Real-Time Monitoring of Equity, Fixed Income, and Currency Markets. arXiv preprint arXiv:2504.04292.
[11] Yang, Jie, et al. "Dynamic Hedging Strategies in Derivatives Markets with LLM-Driven Sentiment and News Analytics." arXiv preprint arXiv:2504.04295 (2025).
[12] Zhao, H., Chen, Y., Dang, B., & Jian, X. (2024). Research on Steel Production Scheduling Optimization Based on Deep Learning.
[13] Bohang, Li, et al. "Image steganalysis using active learning and hyperparameter optimization." Scientific Reports 15.1 (2025): 7340.
[14] Yuan, J. (2025). Contrastive Multimodal Learning for Chest X-ray Analysis: Leveraging Text and Image Synergy. Authorea Preprints.
[15] Ge, Minyue, Zhang Feng, and Qian Meng. "Urban planning and green building technologies based on artificial intelligence: Principles, applications, and global case study analysis." Applied Science and Engineering Journal for Advanced Research 3.5 (2024): 18-27.