Smart Waste Advisor: Real-Time Waste Classification and Interactive Feedback on Embedded Edge Devices

Yuan Jiang
Columbia University
M.S. in Electrical Engineering
yj2848@columbia.edu

Lingxi Zhang
Columbia University
M.S. in Electrical Engineering
lz2991@columbia.edu

Hongrui Huang
Columbia University
M.S. in Electrical Engineering
hh3084@columbia.edu

ABSTRACT

Improper waste sorting remains a persistent challenge in recycling systems, often due to a lack of accessible, real-time guidance. We present **Smart Waste Advisor**, an interactive embedded solution that performs real-time object classification and large language model (LLM)-powered feedback to assist users at the point of disposal. Our system runs a quantized MobileNetV1 model on a XIAO ESP32S3 microcontroller with an onboard camera, achieving inference for ten waste categories within 1 second. Classification results are sent via serial to a Python script, which queries the OpenAI GPT-3.5 API to generate concise disposal advice. This response is both displayed on a 1.28-inch circular screen (Seeed Studio Round Display for XIAO) and played via an I2S-connected speaker using lightweight text-to-speech (TTS). The complete end-to-end pipeline, from image capture to feedback delivery, achieves a typical latency of approximately 1.1 seconds. In real-world testing, the system achieved an overall recognition success rate of approximately 60% across 10 common waste categories. Our results highlight the potential of lightweight, conversational AI systems to promote sustainable disposal practices in everyday environments.

CCS CONCEPTS

• **Computing methodologies** → **Object classification**; • **Computer systems organization** → *Embedded hardware*; • **Applied computing** → *Environmental sciences.*

KEYWORDS

Edge AI, Waste Classification, Embedded Systems, GPT Integration, Real-time Feedback

ACM Reference Format:
Yuan Jiang, Lingxi Zhang, and Hongrui Huang. 2025. Smart Waste Advisor: Real-Time Waste Classification and Interactive Feedback on Embedded Edge Devices. In *Proceedings of Columbia University ELENE6908 Embedded AI Final Project (ELENE6908 Final Project).* ACM, New York, NY, USA, 4 pages.

1 INTRODUCTION

Improper waste sorting continues to undermine recycling efforts worldwide, with up to one-third of recyclable materials mis-sorted due to the lack of timely, accessible guidance at the point of disposal. Traditional tools—such as posters, bin labels, and smartphone apps—require user initiative and attention, creating friction in public or fast-paced environments. Recent advances in embedded AI and lightweight machine learning now enable real-time, on-device classification without network dependency. When coupled with large language models (LLMs), these systems can transform raw predictions into natural language advice, promoting user understanding and encouraging sustainable disposal behavior.

In this project, we introduce Smart Waste Advisor, an embedded system that combines real-time object classification [1], LLM-based disposal guidance, and multimodal feedback, including on-device speech and display. Our solution is built around the XIAO ESP32S3 [2] microcontroller and utilizes a quantized MobileNetV1 model to detect ten waste categories from camera input. Once identified, the item label is sent to OpenAI's [3] GPT-3.5 API via serial communication to generate a concise eco-tip, which is then rendered visually on a 1.28-inch round screen and spoken aloud using a lightweight text-to-speech (TTS) engine. By situating AI intelligence directly at the disposal point, our system reduces friction and offers a seamless, hands-free user experience.

To support reproducibility and encourage future extensions, we have open-sourced our implementation on GitHub, and published a live demonstration video on YouTube.

2 METHODOLOGY

The Smart Waste Advisor system is organized as a real-time pipeline comprising three key modules: image classification, language-guided feedback generation, and multimodal user interaction. These modules are distributed across an embedded microcontroller and a host computer, connected via serial communication.

2.1 On-device Waste Classification

At the core of the embedded device is a XIAO ESP32S3 microcontroller equipped with an onboard camera. We trained a custom MobileNetV1 image classification model using the Edge Impulse platform [4], utilizing a combined dataset of public waste classification images and self-collected samples. After training, the model was quantized to INT8 format and deployed directly to the device via Arduino library export. During inference, the ESP32S3 captures a 320×240 JPEG image, converts it to RGB, resizes it to the 96×96 input resolution, and performs classification in approximately 700 ms. The output includes a predicted label and its confidence score, which are printed to the serial interface in JSON format.

2.2 LLM-Driven Disposal Feedback

To translate raw predictions into user-friendly guidance, the host machine runs a Python script that listens to the ESP32S3's serial output. When a label and confidence score are received, the script constructs a prompt and queries the OpenAI GPT-3.5 Turbo API. The prompt instructs the model to identify the recyclability of the

ELENE6908 Final Project, Spring 2025, New York, NY, USA

item and suggest an appropriate disposal method in no more than 20 words. To reduce redundancy and API load, the system caches and reuses responses for repeated labels. The total round-trip time for GPT response generation is approximately 400 ms.

2.3 Multimodal Output: Display and TTS

Once the GPT-generated advice is received, it is relayed back to the ESP32S3 via serial. The embedded device then renders the response on a 1.28-inch circular TFT screen (Seeed Studio Round Display for XIAO), using line wrapping to fit up to 26 characters per line. Simultaneously, the system converts the message to speech using an I^2S-connected speaker and a lightweight text-to-speech engine (ESP8266SAM). The combined feedback—visual and auditory—is typically delivered within 1.1 seconds of initial image capture, enabling seamless, hands-free interaction.

3 MODEL TRAINING

We trained a custom image classifier using the *Garbage Classification (12 Classes)* dataset from Kaggle [5], selecting 10 relevant waste categories. After filtering and relabeling, our final dataset contained 14,275 images across common waste types like plastic, glass, metal, cardboard, and battery.

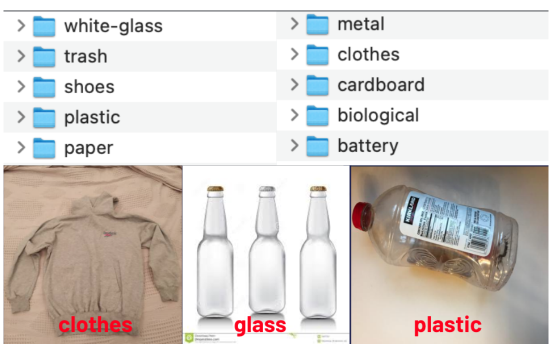

Figure 1: Sample images from dataset in 10 waste categories

The model was built on Edge Impulse using MobileNetV1, selected for its low-latency inference on microcontrollers. We applied quantization-aware training and real-time data augmentation to improve generalization. The final INT8-quantized model was exported as an Arduino library and deployed on the XIAO ESP32S3, achieving inference latency of 700 ms without cloud dependence.

4 SYSTEM IMPLEMENTATION

Smart Waste Advisor is composed of a compact yet complete embedded system that integrates hardware peripherals and modular software components for image classification, GPT querying, and multimodal feedback. Figure 2 shows the high-level architecture.

4.1 Hardware Architecture

Our hardware stack centers on the **XIAO ESP32S3**, a compact dual-core microcontroller with integrated Wi-Fi and onboard PSRAM, well-suited for lightweight AI workloads. It interfaces with:

- A **GC9A01-based 1.28-inch circular LCD display** (Seeed Studio Round Display for XIAO) for visual feedback.
- A **built-in camera module** for real-time image acquisition.
- An **I^2S-connected mini speaker** for audio feedback via text-to-speech.

4.2 Software Stack

We used the **Arduino IDE** to deploy firmware on the ESP32S3, which runs the Edge Impulse–generated inference code and handles peripheral control (camera, display, audio). On the host computer, a lightweight **Python script** listens to serial outputs and forwards classification labels to the **OpenAI GPT-3.5 API**, receiving a short eco-tip in return. This message is sent back to the microcontroller for display and voice playback.

4.3 Modular Pipeline

The system is structured around three modular stages:

(1) **Camera Inference:** The onboard camera captures snapshots that are fed into a MobileNetV1 model, returning a waste label and confidence score.
(2) **GPT Query:** The Python host reads the serial output and queries OpenAI's LLM to generate a disposal suggestion.
(3) **Synchronized Feedback:** The resulting tip is returned via serial, shown on the circular display, and vocalized using on-device TTS.

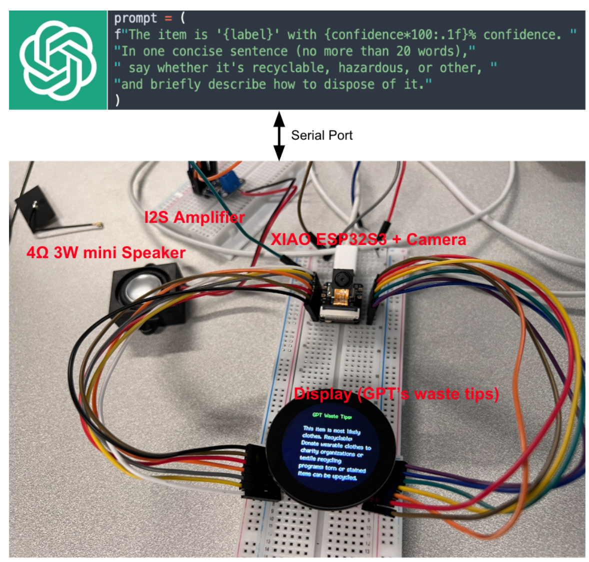

Figure 2: System Architecture of Smart Waste Advisor

5 EXPERIMENTS AND EVALUATION

We evaluated Smart Waste Advisor across three dimensions: classification accuracy, system latency, and real-world detection success.

5.1 Classification Accuracy

We trained and tested our quantized MobileNetV1 model on a curated dataset of 14,275 labeled images across 10 waste categories.

On the held-out test set, the model achieved a **top-1 classification accuracy of 61.8%**. Despite the constraints of edge deployment, this performance is sufficient to trigger meaningful and context-aware eco-feedback from the LLM.

5.2 Latency and Responsiveness

The system's end-to-end latency was measured across multiple runs in realistic usage conditions. Inference on the XIAO ESP32S3 microcontroller took approximately **700 ms** per image. Serial communication and GPT-3.5 API response added another **400 ms**, for a total response time of around **1.1 seconds** from item detection to feedback delivery. This latency was deemed acceptable in interactive demos and user tests.

5.3 Real-World Recognition Success

In field testing, Smart Waste Advisor successfully detected all 10 target waste categories at least once, demonstrating complete category coverage. The overall recognition success rate, including cases where the top-1 prediction matched ground truth, was approximately **60%**. Example output is shown in Figure 3, where real items were classified and matched with appropriate GPT-generated guidance.

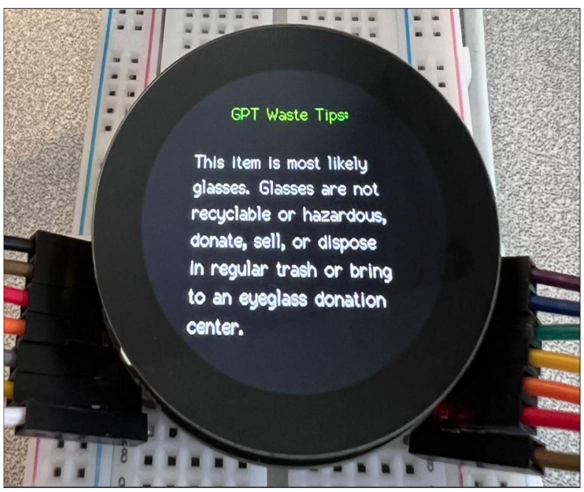

Figure 3: Circular display shows the GPT response

6 RESULTS AND DISCUSSION

We demonstrated Smart Waste Advisor in real-world settings to assess accuracy, responsiveness, and practical usability.

6.1 System Performance

The device successfully recognized all 10 target categories in live tests, displaying GPT-generated eco-tips on a 1.28-inch circular screen and speaking them aloud via a lightweight TTS engine. The full feedback loop—camera input, inference, GPT query, and response rendering—averaged **1.1 seconds**, with on-device classification taking **700 ms** and GPT response **400 ms**.

6.2 Limitations and Future Work

Despite its responsiveness, the system achieved only **61.8%** classification accuracy on a held-out dataset of 14,275 images. Improving this recognition rate remains the most critical goal. Current limitations include limited training diversity, class imbalance, and visually ambiguous items.

We observed three key factors influencing our training results and data quality. First, class imbalance: the clothes category has about 4,000 images—roughly one-third of the dataset—resulting in much higher classification confidence than smaller classes like metal or battery. Second, lighting: bright, even illumination consistently improves inference accuracy, while dim or uneven lighting introduces noise and misclassification. Third, object color: light-colored items are easier to classify. For instance, pale-yellow cardboard, dominant in the cardboard class, is recognized more reliably than darker or multicolored variants.

To address these issues, we aim to expand the dataset with more balanced and diverse real-world examples, and fine-tune the model architecture to improve robustness under variable lighting and color conditions.

Another challenge lies in audio quality. The ESP32S3's built-in TTS engine and DAC deliver intelligible but sometimes distorted speech in noisy environments. To address this, we plan to experiment with higher-fidelity TTS libraries, pre-recorded audio prompts, or external audio codecs.

Furthermore, the system's reliance on internet connectivity to query GPT introduces potential latency and availability issues. Future work will explore caching frequent responses, switching to on-device LLMs like TinyLlama, or adopting hybrid online-offline fallback designs. Support for more waste categories, multilingual eco-tips, and adaptive regional policies could also enhance generalizability and practical impact.

7 CONCLUSION

We presented **Smart Waste Advisor**, a real-time, voice-enabled waste classification system built on low-cost embedded hardware. Through a quantized MobileNetV1 model and GPT-generated guidance, our prototype offers interactive disposal tips with a full response time of 1.1 seconds. While early results show encouraging real-world usability, especially with multimodal feedback, future work is needed to improve classification accuracy, audio clarity, and network independence. This work underscores the potential of compact, AI-powered assistants to promote sustainable behavior in everyday environments.

ACKNOWLEDGMENTS

This project was supported by the ELENE6908 Embedded AI course, with guidance from Professor Xiaofan Jiang and assistance from teaching assistants Emily Bejerano and Kaiyuan Hou.

REFERENCES

[1] Columbia University. Elene6908 tutorial 1: Embedded ai. https://courseworks2.columbia.edu/courses/214858/files/folder/tutorials?preview=23207695, 2025.

[2] Seeed Studio. Seeed studio xiao esp32s3 hardware reference. https://github.com/Mjrovai/XIAO-ESP32S3-Sense, 2023.

[3] OpenAI. Gpt-3.5 turbo api. https://platform.openai.com/, 2023.

[4] Edge Impulse. Edge impulse documentation. https://docs.edgeimpulse.com/docs, 2023.

[5] Mostafa Abla. Garbage classification (12 classes). https://www.kaggle.com/datasets/mostafaabla/garbage-classification, 2023.