15 Trash Sorter Final demo

18-844 project final

Kejia Hu, Bowei Li, Zijia Liu

Carnegie Mellon University

Abstract

objectives and motivation:

a smart device that helps people put trash in the right bin

target users:

1. General public using public recycle bins
2. Organizations running sustainability or recycling programs
3. People who have difficulty manually sorting trash

Carnegie Mellon University

Introduction

a smart device that helps people put trash in the right bin

trash

Nicla Vision

LCD

Why use Embedded ML?

Bandwidth: eliminate heavy upstream data transmission, conserve bandwidth

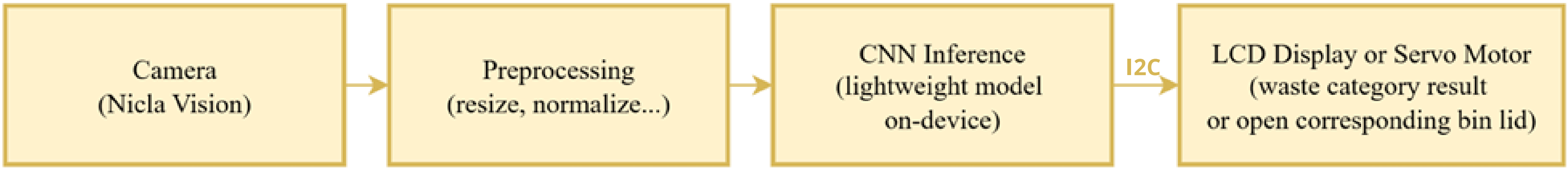
Latency: inference locally at MCU, achieve low latency

Economics: reduce recurring costs by eliminating dependence on cloud servers

Reliability: local processing ensures continuous operation even during network outages

Privacy: prevents personal image data from leaving the device

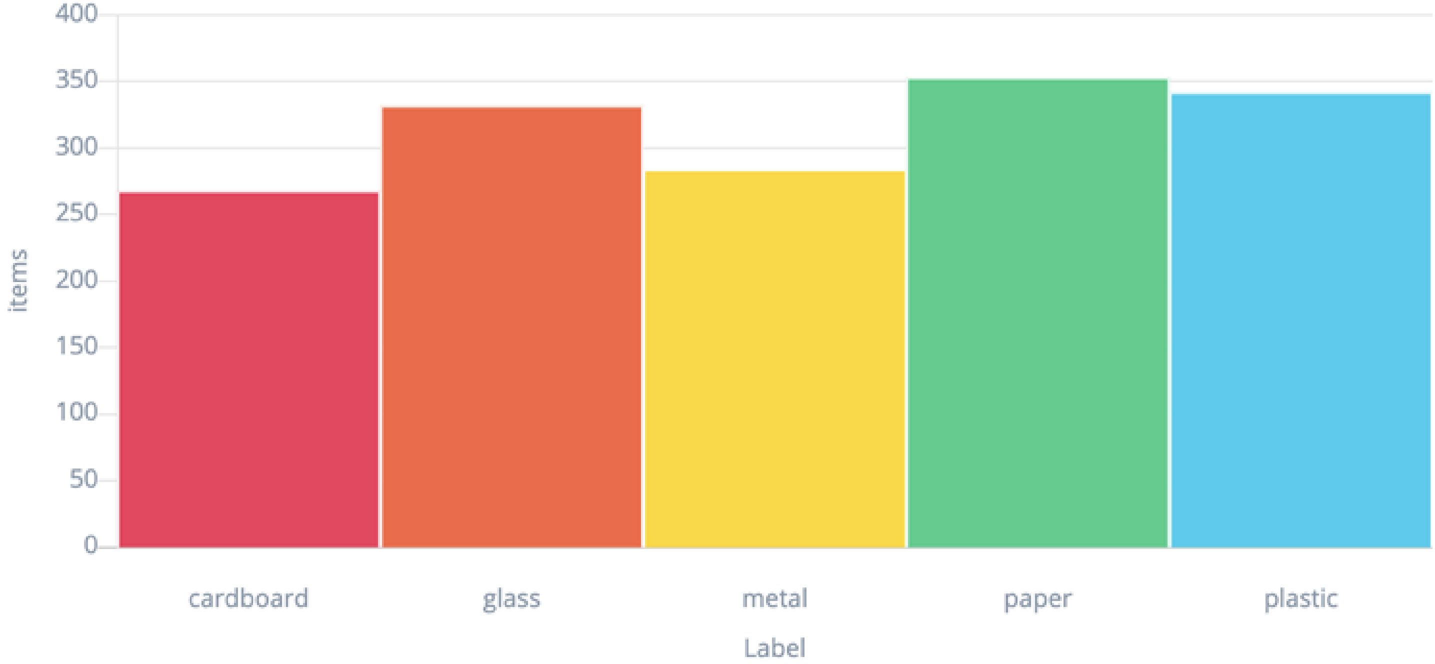
Overall block diagram



Methods - data collection

1. **Kaggle Baseline**: Started with the public Kaggle Garbage Classification dataset as the initial training source.
2. **Manual Cleaning**: Manually removed blurry, mislabeled, and ambiguous images to fix quality issues found in the original dataset.
3. **Real-World Capture**: Collected new images using the actual Nicla Vision camera to ensure the data matched the specific hardware and lighting of the deployment environment.
4. **Standardization**: Resized all images to **96x96 RGB** pixels and split them into training and validation sets for Edge Impulse.

Dataset distribution



Methods - ML models

1. **Model Architecture** Utilized **MobileNet V2** (0.35 alpha), selected for its superior balance of speed and accuracy on embedded devices compared to EfficientNet.
2. **Optimization Strategy** Applied **INT8 Quantization** to compress the model for the STM32H7 MCU, achieving an ultra-low inference latency of **~97ms**.
3. **Training Pipeline** Input images are resized to **96x96 RGB** and processed using **Transfer Learning** based on pre-trained ImageNet weights.
4. **Final Performance** The deployed quantized model achieved a **74.17%** test accuracy, successfully meeting the project's real-time classification goals.

Methods - embedded systems implementations

The **Nicla Vision** captures an RGB image of the waste item using its onboard camera every one second and performs on-device inference using a lightweight convolutional neural network deployed through the Arduino IDE. For each frame, the model outputs a predicted waste class along with a confidence score.

After inference, the Nicla Vision communicates the result to an external **LCD** module over the I²C bus. The predicted class label and its associated confidence value are displayed in real time.

Results

Validation accuracy

Last training performance (validation set)

ACCURACY
73.7%

LOSS
0.67

Confusion matrix (validation set)

	CARDBOARD	GLASS	METAL	PAPER	PLASTIC
CARDBOARD	98.2%	0%	0%	0%	1.8%
GLASS	12.5%	67.2%	9.4%	4.7%	6.3%
METAL	3.3%	16.4%	68.9%	3.3%	8.2%
PAPER	6.8%	3.4%	6.8%	78.0%	5.1%
PLASTIC	6.6%	14.5%	9.2%	7.9%	61.8%
F1 SCORE	0.85	0.66	0.70	0.79	0.69

Testing accuracy

ACCURACY
79.54%

Confusion matrix

	CARDBOAI	GLASS	METAL	PAPER	PLASTIC	UNCERTAII
CARDBOARI	86.4%	6.1%	1.5%	1.5%	4.5%	0%
GLASS	2.4%	77.1%	13.3%	3.6%	3.6%	0%
METAL	0%	14.5%	82.6%	0%	2.9%	0%
PAPER	5.7%	6.8%	5.7%	78.4%	3.4%	0%
PLASTIC	2.4%	7.1%	12.9%	2.4%	75.3%	0%
F1 SCORE	0.86	0.74	0.74	0.85	0.80	

Carnegie Mellon University

Video

Challenges - Power Consumption

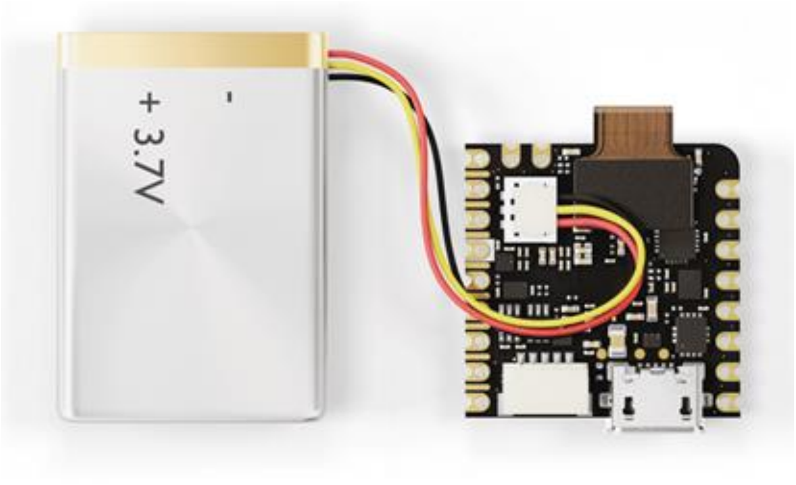


	IMAGE	TRANSFER LEA...	TOTAL
LATENCY	1 ms.	46 ms.	**47 ms.**
RAM	4.0K	133.2K	**133.2K**
FLASH	-	293.7K	-
ACCURACY			**74.17%**

Device Operating Mode

- **Camera + Inference Model:** 250–350 mW
- Based on a **3.7V Li-Po battery**, this corresponds to: → **70–95 mA** (full-speed inference)

Display Module

- LCD consumption: **20–25 mA**

Total System Current

→ ≈ **90–120 mA** during continuous inference

With a typical 5000 mAh Li-Po battery, our system (camera + inference + LCD) can run for **approximately 2-3 days** of continuous operation.

Carnegie Mellon University

Challenges - Power Consumption

- Power Saving with Event-Triggered Inference
- Instead of running inference continuously, the device can **only activate the camera and model when a button press is detected**.
- For the rest of the time, it **stays in a low-power mode (< 50 mW, ~10–15 mA)**.
- This reduces the average current from about **90–120 mA to roughly 15–25 mA**,
- leading to a **5–10× increase in battery life**.
- Example:
  - Continuous inference → **5000 mAh ≈ 2 - 3 days**
  - Event-triggered inference → the same 5000 mAh can last for **several weeks**.

Or simply install a power outlet nearby and plug the device in directly.

Challenges - Class limitation

One limitation of our system is that it can only classify items belonging to a small set of recyclable categories: **cardboard, glass, metal, paper, and plastic**. Real-world waste includes many additional types—such as food waste, textiles, electronics, compostable materials, or hazardous waste—that our model is not trained to recognize. As a result, the device may fail to provide useful guidance when users present items outside the supported categories.

Thank you !