Turning Wastes into Gold: *D3- Embedded* Capstone Goals and Expectations in Brief

Robert Ke, Jason Wang, Richard Gu, Jerry Wang, Sam Sun

Catalog:

1

Literature review: D3 and the world: The power of recycle and reuse electronics

2

Descriptive analysis of example dataset: what will we be working as a U of R student team

3

Goals and expectations: What we, students and D3, will present in this semester as a grand finale of our DS degree.

1

Literature review: D3 and the world: The power of recycle and reuse electronics

Literature review: D3 and the world: The power of recycle and reuse electronics

D3 Embedded is a company that ...

1. **Designs** hardware and software for their customers' need.
2. **Highlights** efficiency: lowering cost and maximum applicability
3. **Dedicates** collaborations: collaborating with the world leading tech-giants to provide advanced and innovative services and products
4. **Devotes** to the sustainable growth: enhancing the waste management and making effort to tackle excessive electronic waste

Our job as U of R students is to...

Problem: EWASTE+ processes over a million pounds of electronics waste a year

Challenges: Efficiency, correctness, room for improvement of material usage, and lack of predictiveness of electronic part price.

Our Goals: Use CV, predictive models, algorithms and tools to identify electronic parts, classify into market categories, and predict prices

Image Source: D3-Embedded

Devastating Status-Quo: Ewaste is costing the Earth

“A record **62 million** tonnes of e-waste was produced in 2022, Up **82%** from 2010; On track to rise another **32%**, to **82 million** tonnes, in 2030; Billions of dollars worth of strategically-valuable resources squandered, dumped; Just **1%** of rare earth element demand is met by e-waste recycling…

…

Meanwhile, less than one quarter (**22.3%**) of the year’s e-waste mass was documented as having been properly , leaving **US$ 62 billion** worth of recoverable natural resources unaccounted…” (UNITAR 2024)

Ewaste is costing our lives: A case study in China

Soil and water contamination; Air pollution - increasing disease risks (WHO 2024)

Case Study as an admonishment: Guiyu, Guangdong, China

“Every household dismantles, every home emits smoke; acid flows into rivers, black clouds cover the sky.” – (Ministry of Ecology and Environment of People’s Republic of China 2019)

60,000 e-waste workers in Guiyu who processed the more than 100 truckloads that were transported to the **52-square-kilometre** area every day. "**Electronic graveyard of the world.** " (Johnson 2006)

Children under the age of 6 are especially vulnerable to lead poisoning, which can severely affect their mental and physical development or even be fatal. Lead can result in irreversible brain damage to their still-developing brains. (Times 2009)

Image Source: CNN 2013

The world is dedicated to recycle Ewaste:

As of **February 2025, 25,** U.S. states and the District of Columbia have enacte **s.** (Environmental Protection Agency 2024)

Bamako Convention: This treaty, primarily involving African nations, prohibits the import of hazardous waste, including e-waste, into Africa, aiming to safeguard the continent from becoming a dumping ground for toxic materials. (IUPAC 2024)

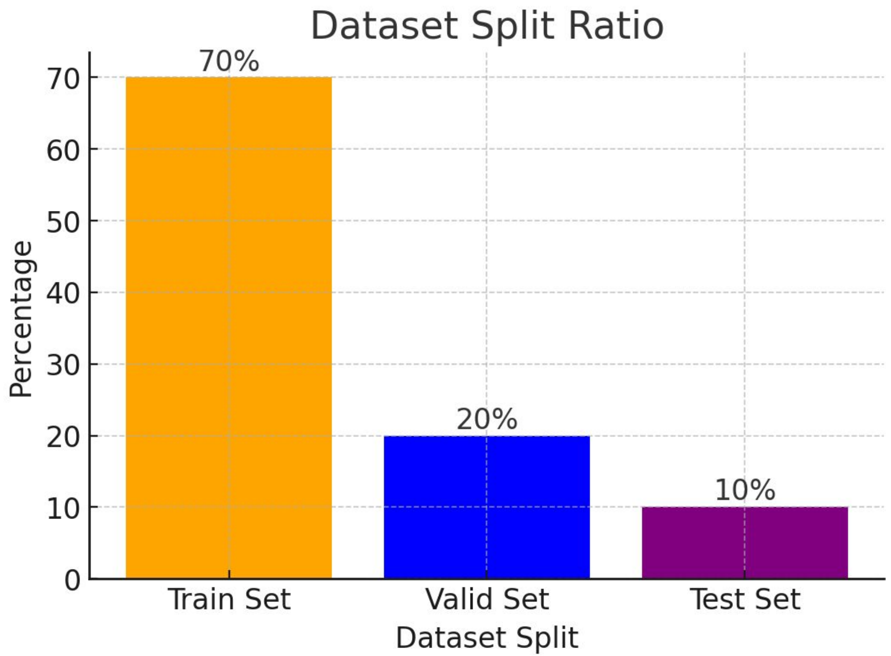
The EU has implemented the Waste **Electrical and Electronic Equipment** (WEEE) Directive, mandating member states to establish systems for collecting, treating, and recycling e-waste. (EU 2024)

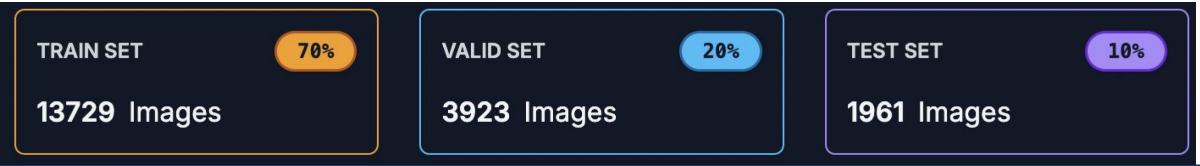
2

Descriptive analysis of example dataset: what will we be working as a U of R student team

Roboflow Data Descriptive Analysis

The dataset has 19613 (13729 in train set, 3923 in valid set, 1961 in test set)annotated images and 77 classes. The dataset has mixed bounding-box and polygon annotations.





Dataset Example

Dataset Classification

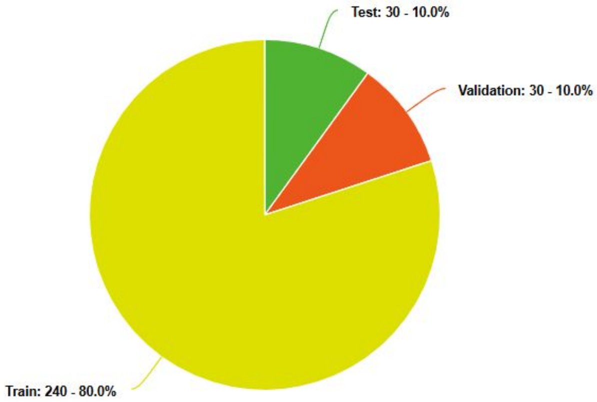
(94 air conditioner, 80 bar-phone, 139 battery, 391 blood pressure monitor, 527 boiler, 108 calculator, 1 camera, 51 ceiling fan, 30 christmas light, 135 cloth iron, 120 coffee machine, 49 Compact-Fluorescent-Lamps, 507 Computer-Keyboard, 644 Computer-Mouse, 77 Cooled-Dispenser, 1020 Cooling-Display, 38 CRT-Monitor, 87 CRT-TV, 1 Dehumidifier, 82 Desktop-PC, 181 Digital-Oscilloscope, 143 Dishwasher, 1047 Drone, 27 Electric-Bicycle, 1000Electric-Guitar, 38 Electrocardiograph-Machine, 285 Electronic-Keyboard, 40 Exhaust-Fan, 674 Flashlight, 461 Flat-Panel-Monitor, 139 Flat-Panel-TV, 90 Floor-Fan, 27 Freezer, 335 Glucose-Meter, 18 HDD, 21 headphone, 853 Laptop, ..

Kaggle dataset introduction

Overview

The Kaggle Dataset is a collection of images representing electronic waste items categorized into 10 distinct classes. There are 300 images in each class (dividing into 240 training; 30 testing; 30 validating), amounting a 300*10 = 3000 images. The dataset is designed for tasks such as image classification, object detection, and other computer vision applications.

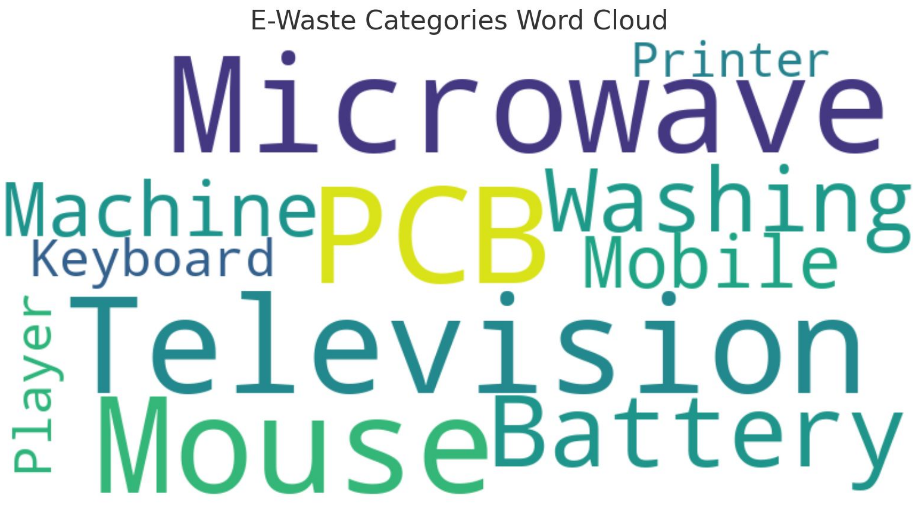
1. **PCB (Printed Circuit Board)**
2. **Player**
3. **Battery**
4. **Microwave**
5. **Mobile**
6. **Mouse**
7. **Printer**
8. **Television**
9. **Washing Machine**
10. **Keyboard**



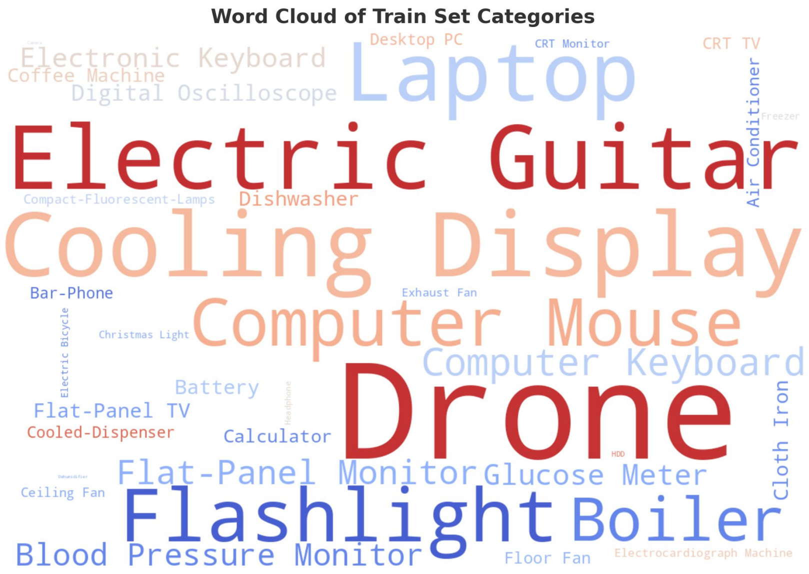
Data Sources

The images in this dataset were collected from diverse sources, including open datasets, image repositories, and proprietary sources. Efforts were made to ensure a representative and diverse collection of electronic waste items.

Wordcloud Visualization



Kaggle Dataset



Roboflow Dataset

3

Goals and expectations: What we, students and D3, will present in this semester as a grand finale of our DS degree.

Project Stages

- **Target Identification & Data Acquisition**
- **Data Preprocessing & Quality Assurance**
- **Model Benchmarking & Initial Training**
- **Model Fine-Tuning & Optimization**
- **Real-World Testing & Validation**
- **Final Solution Deployment & Report Completion**

Data Collection & Preparation

Data Sources:

- Real-world images from e-waste streams
- Public datasets (e.g., Kaggle, Roboflow)

Data Annotation:

- Using RoboFlow to manually label categories (e.g., PCB, laptop, power supply)
- Scene text annotation for brand/model detection

Synthetic Data Generation:

- Augmentation to increase dataset diversity (lighting, orientation) (Using RoboFlow)
- Creation of artificial images when real data is scarce. (**Unity**, **Unreal**, or **Blender**)

Core Computer Vision Techniques

Object Detection and Classification

- Frameworks: Faster R-CNN, YOLO, SSD
- Identifies and classifies multiple e-waste items in a single frame

Scene Text Detection and Recognition (STDR)

- Extracts product labels or part numbers
- Enables rapid identification of device make/models

Fine-Grained Classification

- Distinguishes subtle differences among similar device types (e.g., model variants)

Model Architectures & Transfer Learning

- **Pretrained CNNs (e.g., ResNet, EfficientNet)**
  - Fine-tune on e-waste images for classification
- **Vision Transformers (ViT)**
  - Can capture complex relationships among parts
- **Transfer Learning Advantages**:
  - Reduces training time
  - Works well with limited datasets

Hardware Acceleration & Deployment (Real-time sorting)

GPU/Edge Inference

- NVIDIA Jetson (TensorRT optimization) for on-site sorting
- Low-power accelerators for embedded applications

Performance Considerations:

- Real-time detection speed for high-throughput sorting
- Model compression (quantization, pruning)

Additional Techniques & Extensions

Predictive Analysis

- Market and commodity price forecasting for resale vs. scrap decisions

6D Pose Estimation (if needed)

- Helps understand object orientation for automated handling or robotics. (if using robotics for sorting)

Weekly Plan

Weeks 1-2: Target Identification & Dataset Collection

• Define classification accuracy metrics for PC/laptop sorting.

• Gather diverse images of devices (functional, semi-functional, scrap).

• Label key attributes (model type, condition, serial numbers) for dataset structuring.

Weeks 3-4: Annotation & Dataset Optimization with Roboflow

• Implement Scene Text Detection & Recognition (STDR) for enhanced labeling.

• Optimize dataset quality by refining annotations and reducing noise.

• Conduct exploratory data analysis (EDA) to assess dataset balance and distribution.

Weeks 5-6: Model Benchmarking & Initial Training

• Evaluate YOLO family, HOG, R-CNN, RetinaNet, EfficientDet for: Accuracy, speed, and cost-efficiency.

• Performance on labeled dataset.

• Select top-performing models for further tuning.

Weekly Plan

Weeks 7-8: Hyperparameter Tuning & Model Optimization
- **Optimize key parameters (batch size, learning rate, augmentation).**
- **Train the best model with improved precision and throughput.**
- **Compare post-tuning performance and finalize the best model.**

Weeks 9-10: Real-Time Pipeline Development
- **Design an automated sorting pipeline for real-time classification.**

Weeks 11-12: System Testing & Optimization
- **Conduct pilot testing to validate real-world performance.**
- **Identify processing bottlenecks and refine the pipeline.**
- **Improve real-time decision-making for increased accuracy.**

Weeks 13-14: Performance Metrics & Validation
- **Measure improvements in speed, accuracy, and scrap reduction.**
- **Compare AI-assisted classification with baseline manual sorting.**