Understanding Convolutional Neural Networks (CNN): A Technical Overview

Machine learning has transformed how computers interpret images, and one of the most important tools in this transformation is the CNN. CNNs are specifically designed to detect patterns in visual data and combine them to recognize complex objects. In this case study, you will use a CNN to classify images of waste into three categories: **Recyclable, Compostable, and Trash**. This document provides a technical foundation to help you understand how the method works and why it is well-suited to this problem.

1. What makes CNNs different from regular Neural Networks?

Traditional neural networks treat an image as a long list of pixel values, ignoring the spatial structure of the image. This leads to too many parameters, poor performance on visual patterns, and models that do not generalize well. CNNs solve this by taking advantage of the spatial locality of images: nearby pixels are usually related. Instead of flattening images, CNNs process them in 2D.

2. The Core Building Block: Convolution

A *convolution* is a mathematical operation where a small matrix called a **filter** slides over the image and performs element-wise multiplication.
Each filter detects a different pattern:

- A vertical edge
- A horizontal edge
- A curve
- A texture

When the filter moves across the image, it produces a **feature map—**a transformed version of the image that highlights where that pattern appears.

Why this matters for waste classification:

Different waste types exhibit distinct patterns:

- Plastics have smooth, reflective surfaces
- Compostable items often have irregular organic textures
- Trash can appear highly varied

CNNs learn filters that pick up on these subtle differences.

3. Building up Complexity with Layers

A CNN is composed of multiple types of layers, each with a specific role:

a. Convolutional Layers

Extract visual features (edges, textures, shapes)

b. Activation Layers (ReLU)

Introduce non-linearity so the model can learn complex patterns.

c. Pooling Layers

Reduce the size of feature maps to:

- Make computation faster
- Reduce overfitting
- Focus on the “most important” features

d. Fully Connected Layers

Toward the end of the network, feature maps are flattened and passed into standard neural network layers to make the final prediction.

4. The CNN Workflow in this Case Study

When you train a CNN on the RealWaster dataset:

1. Input images enter the convolutional layers.
2. The network extracts low-level features (edges, corners)
3. Deeper layers extract higher-level features (shapes, textures)
4. Fully connected layers combine everything into a prediction:
    a. Recyclable / Compostable / Trash
5. Your model learns these features through **backpropagation**, adjusting filter values to reduce classification error.

The RealWaster dataset originally contains nine classes, but many of them are severely unbalanced:

- Some categories have hundreds of images
- Others have fewer than 100

CNNS trained on imbalanced data tend to:

- Favor majority classes
- Misclassify minority classes
- Performs poorly on underrepresented categories

This is why the original nine-class approach performed poorly (53% accuracy)

Why regrouping helps:

By combining categories into **three broader groups, we:**

- Increase the number of images per class
- Reduce imbalance
- Improve the model’s ability to generalize
- Reflect real-world waste sorting more accurately

This regrouping is one of the most important technical decisions in the case study.

6. Understanding Evaluation Metrics

CNN performance is evaluated using several metrics:

Accuracy

- Percentage of all predictions that were correct

Precision

- Of all items predicted to be in a class, how many truly were?

Recall

- Of all real items in a class, how many were identified correctly?

F1 Score

- The harmonic mean of precision and recall.
- Useful when classes are imbalanced.

Confusion Matrix:

- A 3x3 table showing:
    - Where the model succeeds
    - Where it confuses categories
    - Which classes are hardest to distinguish

In your results, you should interpret patterns such as:

- The model accurately identifies recyclables
- Compostables sometimes overlap with trash
- Trash often shows the greatest variability

7. Why CNNs Are a Good Fit for This Task

CNNs are ideal for waste classification because they:

- Automatically learn visual patterns
- Scale well to thousands of images
- Are robust to lighting, perspective, and background differences
- Can learn complex distinctions between material types

Even a small, simple CNN achieves strong performance on grouped categories, as demonstrated by the 83% accuracy achieved in the original project.

8. What You Will Implement

In this case study, you will:

- Load the RealWaster dataset
- Apply the 3-class regrouping
- Build a small CNN
- Train the model
- Evaluate performance
- Interpret misclassifications and uncertainty

This technical overview should give you the conceptual foundation to understand the architecture you will implement in the starter code.