Image Classification of Waste Items

Shahla Kazimzada, Sahib Rzayev, Yashar Ibrahimxalilov

Introduction

This project aims to develop machine learning models for classifying waste items into 9 major material types, such as cardboard, food organics, glass, metal, and plastic. The dataset was collected within an authentic landfill environment, presenting real-world challenges such as occlusions, lighting variations, and diverse waste appearances. The primary goal is to build robust image classification models capable of supporting automated waste segregation systems.

Two Convolutional Neural Network (CNN) architectures, ResNet18 and VGG16, were implemented and trained for the task. These models were chosen for their contrasting design philosophies: ResNet18 employs residual connections to enable deeper architectures by addressing vanishing gradients, while VGG16 focuses on simplicity with a stack of convolutional layers and fixed kernel sizes.

The cross-entropy loss function was used to optimize the models, as it is well-suited for multi-class classification problems. Two optimizers, Stochastic Gradient Descent (SGD) and Adam, were experimented with to analyze their effects on training speed and model performance. The experiments provide insights into the working principles of CNN components, the differences between architectures, and the role of optimizers and loss functions in achieving effective waste classification.

Results of Waste Classification Models

The performance of ResNet18 and VGG16 is summarized in the table below, including metrics for test accuracy, F1-Score, and classwise accuracy.

Table 1: Performance Comparison of ResNet18 and VGG16

Model **Classwise Accuracy**	**Optimizer**	**Loss**	**Accuracy**	**F1 Score**
ResNet18 [0.5753, 0.8909, 0.7903, 0.7217, 0.5161, 0.6486, 0.7445, 0.5000, 0.6949]	SGD	CE	0.6793	0.6767
VGG16 [0.2333, 0.4482, 0.6250, 0.4252, 0.0571, 0.7727, 0.4000, 0.3200, 0.8148]	Adam	CE	0.4342	0.4118

TensorBoard Visualizations

The training and evaluation process for the models was monitored using TensorBoard. The following plots provide a detailed view of accuracy during training, validation, and testing.

Accuracy During Training

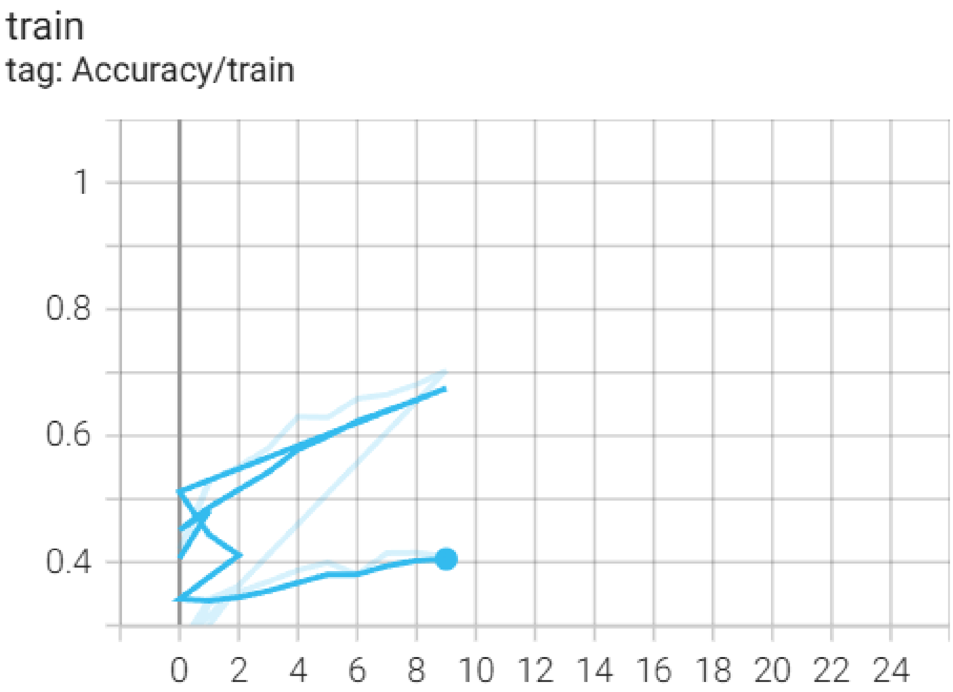

Figure 1: Training accuracy across epochs.

Accuracy During Validation

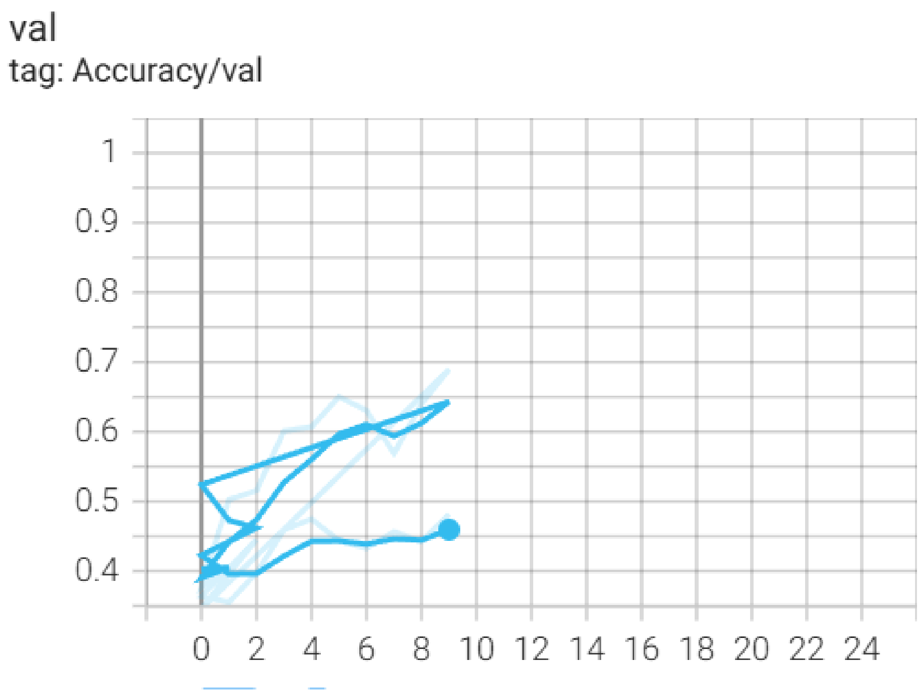

Figure 2: Validation accuracy across epochs.

Test Accuracy

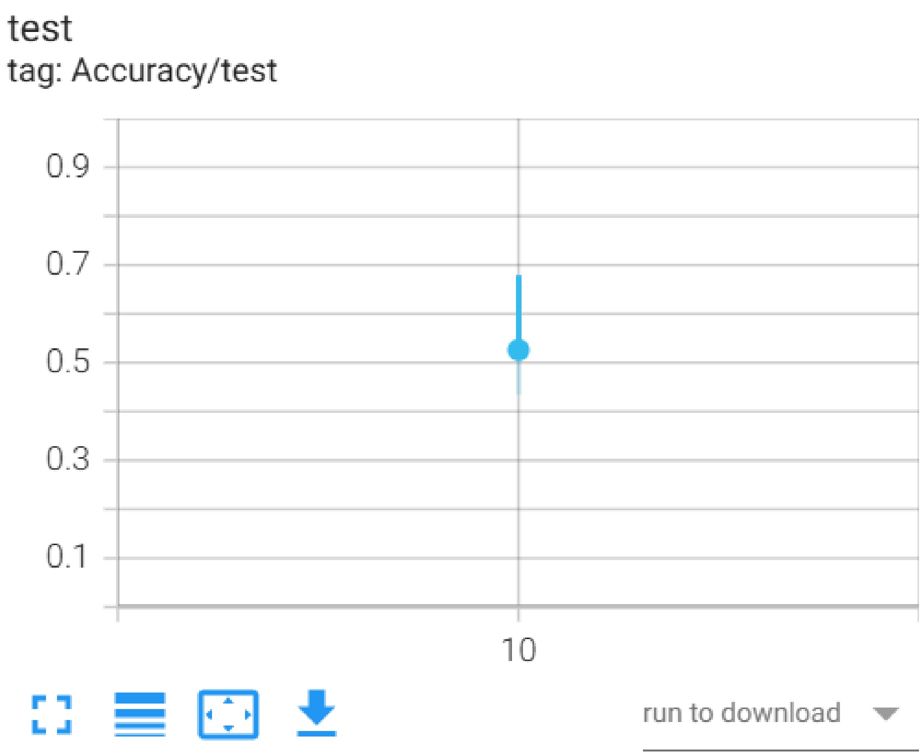

Figure 3: Test accuracy at the final epoch.

Model Architectures and Principles

ResNet18 and VGG16 are CNN architectures designed for image classification but differ significantly in their principles and design:

ResNet18

ResNet18 (Residual Network) uses residual connections, or "skip connections," which allow gradients to flow through the network more easily. This addresses the vanishing gradient problem and enables training of deeper networks. The architecture includes 18 layers, with blocks that compute the residual mapping of inputs, allowing the network to learn identity mappings.

VGG16

VGG16 is a simpler and more traditional CNN architecture that consists of 16 layers. It uses fixed-size convolutional filters (3x3) and focuses on stacking multiple convolutional layers followed by max-pooling layers. This design results in a deeper network that captures hierarchical features but lacks the advanced mechanisms to combat gradient-related issues seen in deeper networks.

Loss Function and Optimizers

Loss Function: Cross-Entropy (CE)

The cross-entropy loss function measures the difference between predicted probabilities and the true class labels. It is widely used for classification tasks as it penalizes incorrect predictions and rewards accurate ones.

Optimizers: SGD and Adam

- **SGD (Stochastic Gradient Descent):** A traditional optimizer that updates weights using the gradient of the loss function. It often requires manual tuning of the learning rate and benefits from momentum to accelerate convergence. - **Adam (Adaptive Moment Estimation):** Combines the advantages of RMSProp and momentum. It adapts the learning rate for each parameter based on estimates of first and second moments of gradients, often converging faster than SGD.

Conclusion

The ResNet18 model trained with SGD outperformed the VGG16 model trained with Adam in both accuracy (67.93%) and F1-Score (0.6767). However, the classwise accuracy results show that some classes, such as "Food Organics" and "Vegetation," achieved high accuracies in both models, while "Miscellaneous Trash" remained challenging to classify. These results highlight the importance of architectural design and optimization techniques in achieving robust waste classification.