A Deep Analysis of Modeling, Architecture, and Dataset Within Waste Classification

Tamir Yirga
Khoury College of Computer Sciences
Northeastern University
Seattle, WA, USA
yirga.t@northeastern.edu

Nihal Sandadi
Khoury College of Computer Sciences
Northeastern University
Seattle, WA, USA
sandadi.ni@northeastern.edu

Lana Sleek
Khoury College of Computer Sciences
Northeastern University
Seattle, WA, USA
sleek.l@northeastern.edu

***Abstract*—Automated waste classification is essential for improving recycling efficiency and sustainable waste management. This study presents a comprehensive evaluation of modern CNN architectures for waste classification across multiple datasets. We compared ResNet-50, ConvNeXt-Tiny, EfficientNet-B3, and MobileNetV3 variants on three datasets: TrashNet (2,527 images, 6 classes), Garbage Classification v2 (10 classes), and a Kaggle waste dataset. Transfer learning proved critical, with pretrained ResNet-50 outperforming random initialization by 19% in F1-score. ConvNeXt-Tiny achieved the highest accuracy across datasets, reaching 96.6% on TrashNet and 96.73% on Garbage Classification v2, while MobileNetV3-Small demonstrated competitive performance (95.61%) with significantly lower computational requirements. Cross-dataset evaluation revealed that dataset characteristics significantly impact performance, with identical ResNet-50 models achieving 70.3% versus 80.4% accuracy. All architectures struggled with ambiguous categories, particularly "trash" (mixed waste). Our findings highlight the importance of transfer learning, dataset diversity, and architecture selection for real-world waste sorting systems.**

***Index Terms*—waste classification, CNN, transfer learning, ResNet, ConvNeXt, EfficientNet, MobileNet**

I. Introduction

The global waste crisis presents a significant environmental challenge, with improper waste disposal and inadequate recycling contributing to pollution, resource depletion, and climate change. Effective waste management relies heavily on accurate waste classification, which traditionally requires manual sorting which is labor-intensive, time-consuming, and an often hazardous process. Automating this task through computer vision and deep learning offers a promising solution to improve recycling efficiency, reduce contamination in waste streams, and support sustainable waste management practices.

Deep learning, particularly convolutional neural networks (CNNs), has revolutionized image classification by learning hierarchical feature representations directly from raw image data. In waste management contexts, this translates to automatically identifying different types of waste materials such as plastic, paper, metal, glass, and organic waste. However, selecting optimal architectures for deployment requires understanding trade-offs between accuracy, computational efficiency, and the characteristics of available training data.

Transfer learning has emerged as a critical technique, where models pre-trained on large-scale datasets like ImageNet are fine-tuned for specific tasks [5]. This approach is particularly valuable for waste classification, where labeled datasets may be limited. By leveraging knowledge learned from millions of general images, transfer learning enables models to achieve high accuracy even with relatively smaller domain-specific datasets.

This study presents a comprehensive evaluation of modern CNN architectures across multiple waste classification scenarios to provide practical deployment guidance. We investigate four prominent architectures: ResNet-50 [1], known for its residual connections enabling deep network training; ConvNeXt-Tiny [2], a modern architecture incorporating transformer-inspired design principles; EfficientNet-B3 [4], employing compound scaling for parameter efficiency; and MobileNetV3 [3], a lightweight architecture designed for mobile and edge devices. Each represents different trade-offs between model complexity, computational efficiency, and classification accuracy.

We conduct three complementary experiments to understand architecture performance, transfer learning impact, and dataset characteristics' influence on model generalization. First, we systematically compare all four architectures on the TrashNet dataset (2,527 images, 6 classes: cardboard, glass, metal, paper, plastic, trash), including a randomly-initialized baseline to quantify transfer learning benefits. Second, we evaluate three architectures (ResNet-50, ConvNeXt-Tiny, MobileNetV3-Small) on the Garbage Classification v2 dataset, comprising 10 waste categories: battery, biological waste, cardboard, clothes, glass, metal, paper, plastic, shoes, and trash. Third, we perform cross-dataset comparison by training identical ResNet-50 models on both TrashNet and a larger Kaggle waste dataset to isolate dataset characteristics' impact on performance. Together, these experiments provide comprehensive insights into model selection, training strategies, and data requirements for practical automated waste sorting systems.

II. Related Work

Waste classification using deep learning has been an active area of research in computer vision, with numerous studies exploring various architectures and techniques for automated recycling systems.

Deep Learning Architectures for Waste Classification: Several studies have demonstrated the effectiveness of CNNs for waste classification tasks. Bircanoğlu et al. [6] proposed RecycleNet, a modified DenseNet121 architecture that reduced parameters from 7 million to 3 million while achieving 95% test accuracy with transfer learning on the TrashNet dataset. Their work demonstrated that architectural modifications to dense blocks can improve both efficiency and performance for waste classification. Similarly, Ruiz et al. [7] compared multiple CNN architectures including VGG, Inception, and ResNet on TrashNet, with their combined Inception-ResNet model achieving 88.6% accuracy. These foundational works established benchmarks for waste classification on controlled datasets.

Transfer Learning in Environmental Applications: The application of transfer learning to environmental and recycling tasks has been extensively studied. Pan and Yang [5] provided a comprehensive survey establishing theoretical foundations for transfer learning across domains, which has become critical for waste classification where labeled data is limited. Vo et al. [8] developed DNN-TC, an improved ResNext model using deep transfer learning that achieved 98% accuracy on their VN-trash dataset, demonstrating that leveraging pretrained ImageNet weights significantly outperforms random initialization. Zhang et al. [9] further validated this approach using DenseNet169 with transfer learning on a larger recyclable waste dataset of nearly 20,000 images across five categories, achieving strong generalization despite dataset diversity challenges. This body of work validates the importance of transfer learning in scenarios with limited labeled waste data.

Dataset Characteristics and Model Generalization: Recent research has emphasized that dataset properties significantly impact model performance in real-world deployment. Zhang et al. [9] noted that existing public datasets often have small sample sizes, uneven distribution, single backgrounds, and obvious features, limiting model generalization ability. Their construction of more diverse datasets with balanced distributions and varied backgrounds proved essential for training robust models. This finding aligns with our cross-dataset experiments showing that identical architectures achieve substantially different performance on datasets with varying realism and diversity.

Our work builds upon these foundations by providing a systematic comparison of modern architectures (including the recent ConvNeXt family) across multiple datasets with varying characteristics. Unlike previous studies that focus on single architectures or controlled datasets, we evaluate how dataset realism and diversity affect model performance, providing practical insights for deploying waste classification systems under real-world conditions.

III. Methods

A. Architecture Comparison

To identify the optimal CNN architecture for waste classification, we conducted a systematic evaluation of four modern architectures with varying complexity and design philosophies. We additionally trained one architecture from random initialization to quantify the benefit of transfer learning on the limited dataset.

B. Dataset

We used the TrashNet dataset, a widely-benchmarked collection of 2,527 images across six waste categories: cardboard (403), glass (501), metal (410), paper (594), plastic (482), and trash (137). Images were captured against uniform backgrounds at 512×384 resolution. We applied stratified sampling to create 70/15/15 train/validation/test splits, preserving class proportions across all partitions. One notable case of class imbalance was the trash category, constituting only 5.4% of samples versus eg. paper at 23.5%.

C. Architectures

We evaluated four architectures representing different CNN design paradigms:

TABLE I
Architecture Specifications

Architecture	Parameters	Design Rationale
ResNet-50	23.5M	Residual connections baseline; well-understood benchmark
ConvNeXt-Tiny	28.6M	Modernized ConvNet with transformer-inspired design
EfficientNet-B3	12.2M	Compound-scaled architecture optimized for parameter efficiency
MobileNetV3-Large	5.5M	Lightweight architecture for edge deployment

All models were initialized with ImageNet-1K pretrained weights via the timm library. ResNet-50 [1] serves as our baseline with its well-understood residual connection architecture. ConvNeXt-Tiny [2] modernizes traditional CNNs with transformer-inspired design choices including LayerNorm and larger kernel sizes. EfficientNet-B3 [4] employs compound scaling to optimize depth, width, and resolution simultaneously. MobileNetV3-Large [3] utilizes depthwise-separable convolutions and neural architecture search for efficient edge deployment.

D. Training Configuration

These were the hyperparameters across this experiments to ensure a fair comparison:

TABLE II
Training Hyperparameters

Parameter	Value
Optimizer	AdamW
Learning Rate	1×10^{-4}
Weight Decay	0.01
LR Schedule	Cosine annealing ($\eta_{min} = 1 \times 10^{-6}$)
Batch Size	32
Max Epochs	50
Early Stopping	10 epochs (patience)

We applied a minimal form of data augmentation which consisted of random resized crops (scale 0.8–1.0), horizontal flips (p=0.5), and color jitter (brightness, contrast, saturation: ±0.2; hue: ±0.1). All images were resized to 224×224 and normalized using ImageNet statistics.

E. Evaluation Metrics

We assessed model performance using:

- **Accuracy**: Overall classification correctness
- **Macro F1-Score**: Harmonic mean of precision and recall, averaged across classes without weighting (important given the class imbalance in the dataset)
- **Per-class F1**: Individual class performance to identify category-specific weaknesses
- **Inference Latency**: Mean inference time (ms) over 100 runs with GPU synchronization, measured on Apple M1 Pro
- **Model Size**: Total parameters and disk footprint (MB)

F. Experimental Procedure

Each architecture was trained with identical random seeds for reproducibility. The best model checkpoint (by validation F1) was selected for final test evaluation. Inference benchmarking included 10 warmup iterations to account for GPU compilation overhead before timing 100 inference passes.

IV. Experiments and Results

A. TrashNet Dataset Results

Table III summarizes performance across all evaluated architectures.

TABLE III
PERFORMANCE COMPARISON ACROSS ARCHITECTURES

Architecture	Pre-trained	Test Acc	Test F1	Inference (ms)	Params (M)
ResNet-50	Yes	0.916	0.892	9.19	23.5
ConvNeXt-Tiny	Yes	0.966	0.955	10.60	28.6
EfficientNet-B3	Yes	0.929	0.915	19.45	12.2
MobileNetV3-Large	Yes	0.897	0.873	8.51	5.5
ResNet-50	No	0.776	0.750	11.22	23.5

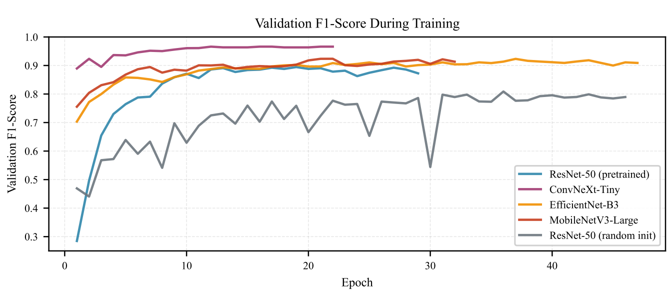

Fig. 1. Training curves showing validation F1-score over epochs for all architectures.

1) Transfer Learning Impact: Comparing pretrained versus randomly-initialized ResNet-50 revealed a 14.3 percentage point improvement in F1-score (0.892 vs 0.750), representing a 19% relative improvement from transfer learning. This demonstrates the critical advantage of using pretrained weights when training on small datasets like TrashNet, where the model benefits substantially from features learned on ImageNet's diverse visual patterns.

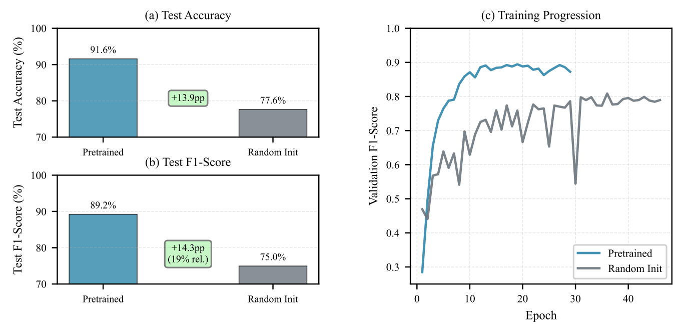

Fig. 2. Transfer learning impact comparison - ResNet-50 pretrained vs random initialization showing (a) Test Accuracy, (b) Test F1-Score, and (c) Training progression curves.

2) Accuracy-Efficiency Tradeoff: ConvNeXt-Tiny achieved the highest F1-score of 0.955 and required 10.60 ms per inference. In contrast, MobileNetV3-Large provided the fastest inference at 8.51 ms (24% faster) with only an 8.6 percentage point reduction in F1-score (0.873). For deployment scenarios prioritizing real-time performance, MobileNetV3-Large offers an attractive balance, achieving 117 images/second throughput while maintaining competitive accuracy. EfficientNet-B3, despite its parameter efficiency (12.2M parameters), exhibited unexpectedly slow inference (19.45 ms) on M1 Pro hardware, likely due to suboptimal operator implementations for its compound-scaled architecture.

3) Per-Class Analysis: All models struggled most with the minority "trash" class (mixed waste), which represents real-world ambiguous items. ConvNeXt-Tiny achieved the best trash class F1 of 0.872, while the randomly-initialized ResNet-50 managed only 0.581. This disparity highlights how pretrained features help disambiguate challenging categories. The confusion matrices (Figure 4) reveal consistent error patterns: models frequently misclassified metal as glass or cardboard, suggesting visual similarity in reflective and textured surfaces. Paper and cardboard also showed mutual confusion, likely due to their similar textures.

Table IV shows per-class F1-scores for the best-performing model (ConvNeXt-Tiny):

B. Garbage Classification Dataset Experiments

1) Dataset: We also experimented with the Garbage Classification v2 dataset, a comprehensive collection of waste images designed for multi-class classification tasks. The dataset comprises 10 distinct waste categories: battery, biological waste, cardboard, clothes, glass, metal, paper, plastic, shoes,

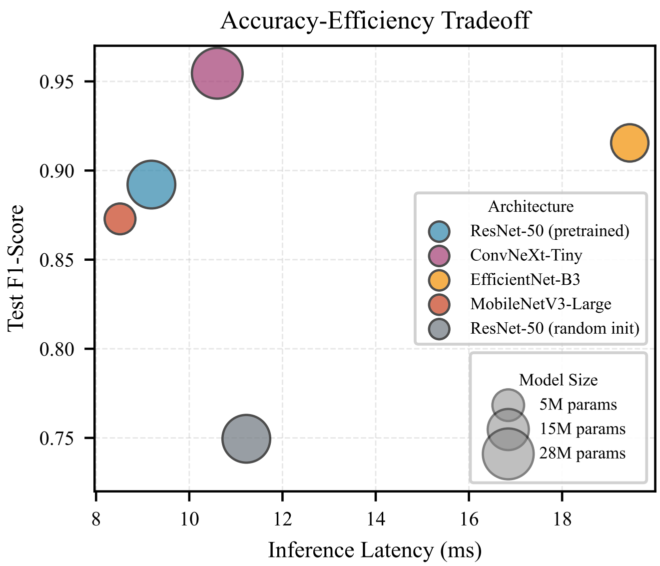

Fig. 3. Scatter plot showing F1-score vs. inference latency tradeoff across architectures.

TABLE IV
PER-CLASS F1-SCORES FOR CONVNEXT-TINY

Class	**F1-Score**
Cardboard	0.974
Glass	0.980
Metal	0.976
Paper	0.972
Plastic	0.953
Trash	0.872

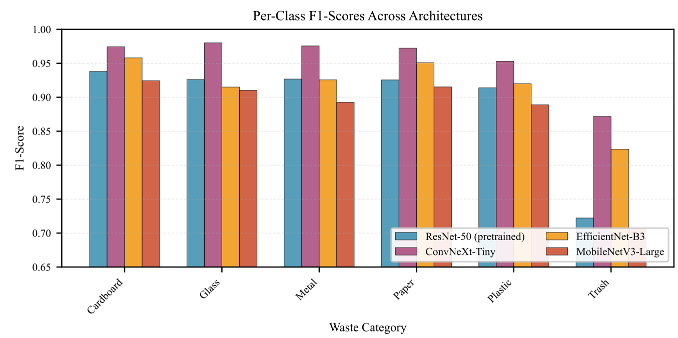

Fig. 5. Bar chart comparing per-class F1-scores across all pretrained architectures.

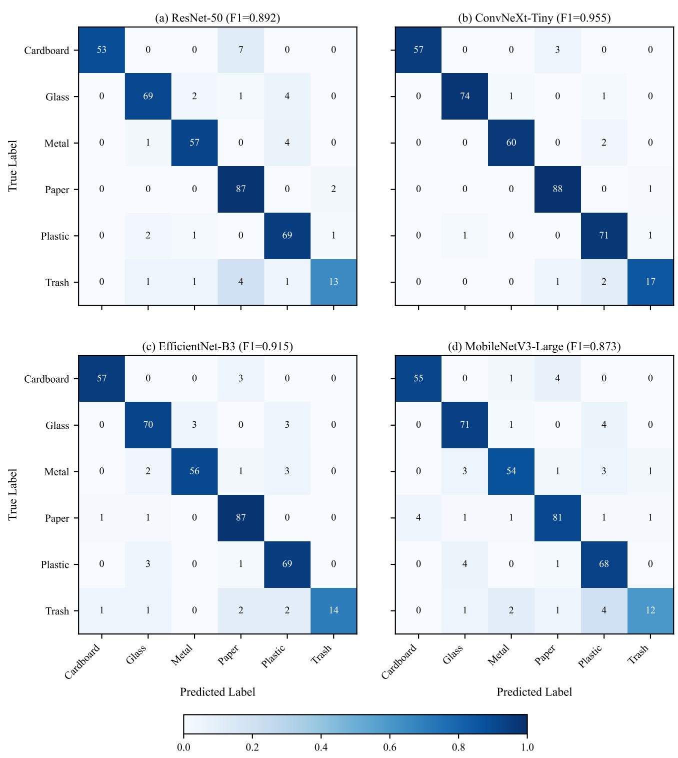

Fig. 4. Confusion matrices for all 4 tested architectures.

and trash. This expanded taxonomy extends beyond the traditional six-class waste classification by incorporating additional categories relevant to comprehensive waste management systems—batteries for hazardous waste handling, biological waste for composting operations, and clothing/shoes representing textile waste streams. Images feature real-world variability including diverse backgrounds, lighting conditions, and object states (clean versus contaminated, whole versus damaged), making the dataset valuable for evaluating model robustness under practical deployment conditions.

For this expanded classification task, we evaluated three architectures: ResNet-50, ConvNeXt-Tiny, and MobileNetV3-Small. The increased number of classes (10 versus 6) tests architectural capacity to learn fine-grained inter-class boundaries, particularly for semantically related categories. This experiment complements our TrashNet evaluation by assessing how models scale to more complex waste taxonomies encountered in real-world sorting facilities.

For this experiment with the Garbage Classification dataset, we used the following hyperparameters:

TABLE V
GARBAGE CLASSIFICATION DATASET HYPERPARAMETERS

Parameter	**Value**
Optimizer	AdamW
Initial Learning Rate	0.001
LR Schedule	LR mutliplieer on plateau 0.5
Batch Size	64
Number of epochs	25
Loss	Cross-entropy
Dropout rate	20%

2) Results: On the Garbage Classification v2 dataset, all three architectures achieved strong performance despite the increased complexity of 10 classes. ConvNeXt-Tiny achieved the highest accuracy at 96.73%, followed closely by ResNet-50 at 96.56%, while MobileNetV3-Small reached 95.61% with significantly lower computational requirements. The minimal performance gap between architectures (1.12 percentage points) suggests that all three effectively learned discriminative features for the expanded waste taxonomy. Per-class analysis (Figures 8 and 10) reveals challenges with specific categories across all models, particularly trash and metal, suggesting these categories present visual ambiguity.

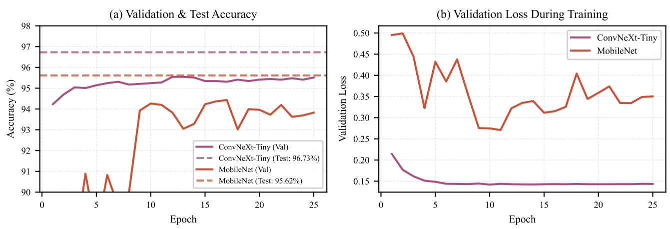

Fig. 6. Test and validation accuracy for Garbage Classification datasets across different models.

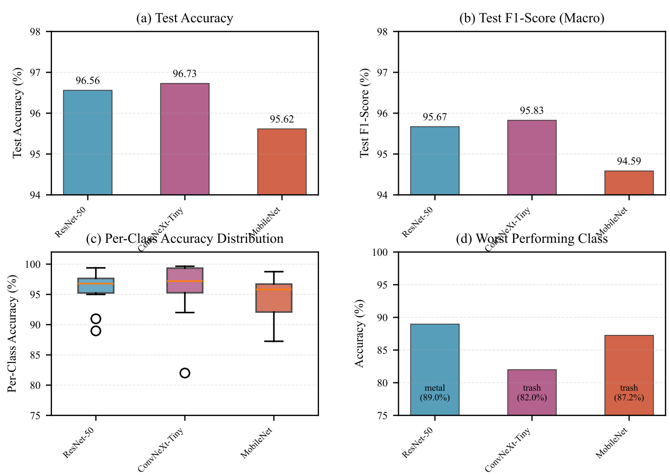

Fig. 7. Test accuracy, F1-Score and worst performing classes across differnet models for Garbage Classification dataset.

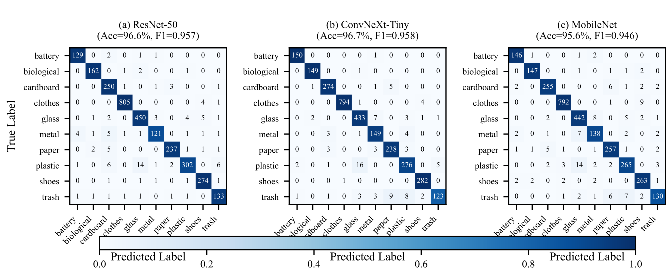

Fig. 8. Confusion matrix for 10 classes in Garbage Classification dataset across 3 models.

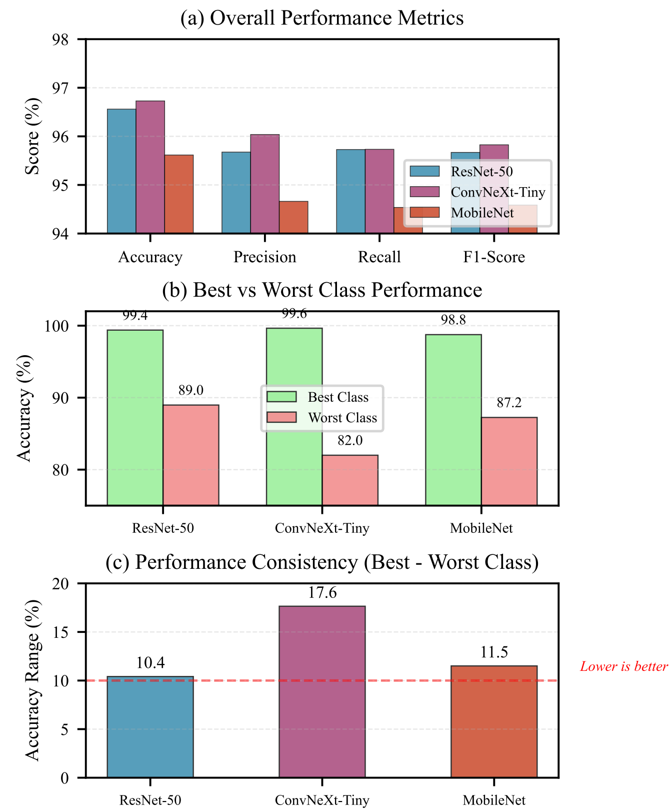

Fig. 9. Overall performance metrics, performance consistency across all three models.

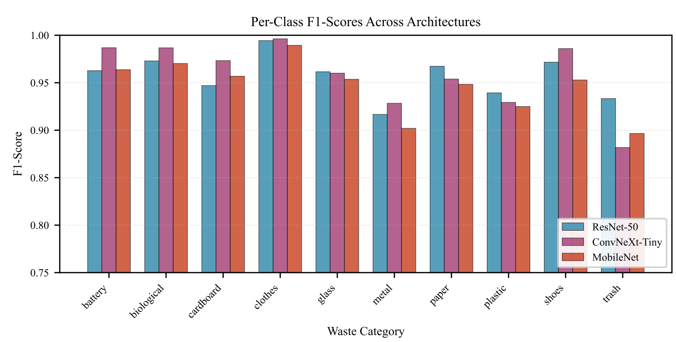

Fig. 10. Per class F1-Score across models for Garbage Classification dataset.

C. Cross-Dataset Comparison: TrashNet and Kaggle

To isolate the impact of dataset characteristics on model performance, we trained identical ResNet-50 models on two datasets with the same six waste categories but differing in size, visual characteristics, and collection conditions. This experiment controls for architecture and hyperparameters, allowing direct assessment of how dataset properties, rather than model sophistication, influence classification accuracy.

1) Dataset Descriptions: **TrashNet Dataset:** Comprises 2,527 images across six waste categories (cardboard, glass, metal, paper, plastic, and trash) with approximately 400-500 images per class. Images feature isolated objects against clean white backgrounds with consistent lighting and orientation, providing controlled conditions for initial model evaluation.

Kaggle Waste Dataset (Subset): A curated subset of 9,428 images representing the same six waste categories extracted from a larger multi-class waste dataset. This subset presents real-world challenges including heterogeneous backgrounds, variable lighting conditions, diverse object orientations, and inconsistent image quality.

2) Experimental Configuration: We trained ResNet-50 with identical hyperparameters on both datasets: 8 epochs, batch size 128, learning rate 0.0001, weight decay 0.01, early stopping patience of 10 epochs, and 224×224 pixel images. Table VI and Table VII summarize the training configuration and performance comparison.

TABLE VI
TRAINING CONFIGURATION FOR CROSS-DATASET COMPARISON

Model	Epoch	Batch Size	Learning Rate	Weight Decay	Pixel Dimensions
ResNet50	8	128	1.350	81.67%	224x224

TABLE VII
CROSS-DATASET PERFORMANCE COMPARISON

Summarized Statistics	Test Acc	Test F1	Best Class	Worst Class
TrashNet	70.3%	0.682	Cardboard 81.67%	Metal 54.80%
Kaggle	80.4%	0.809	Metal 91.11%	Plastic 69.17%

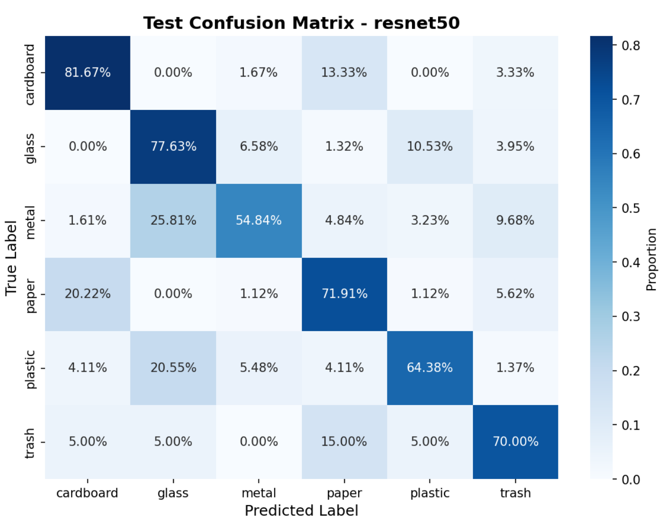

Fig. 11. Confusion Matrix for TrashNet in 6 type classification.

3) Overall Performance: The models demonstrated fundamentally different performance characteristics on the two datasets. On the Kaggle dataset, the model achieved 80.4% accuracy with a macro-averaged F1-score of 0.809, indicating robust learning across all classes. In contrast, the TrashNet model attained only 70.3% accuracy with a 0.682 F1-score,

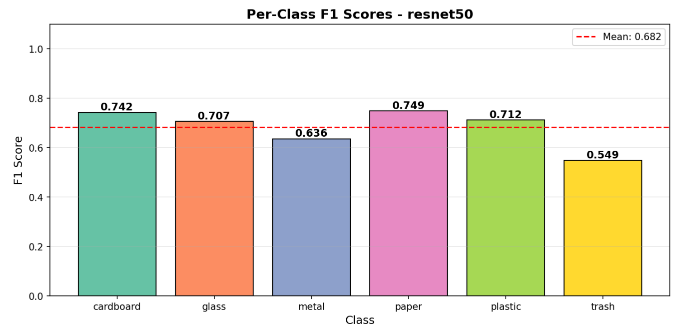

Fig. 12. F1 Scores per 6 classes with TrashNet and mean F1 Score.

Fig. 13. Training History of ResNet over 8 epochs as it trains off TrashNet data.

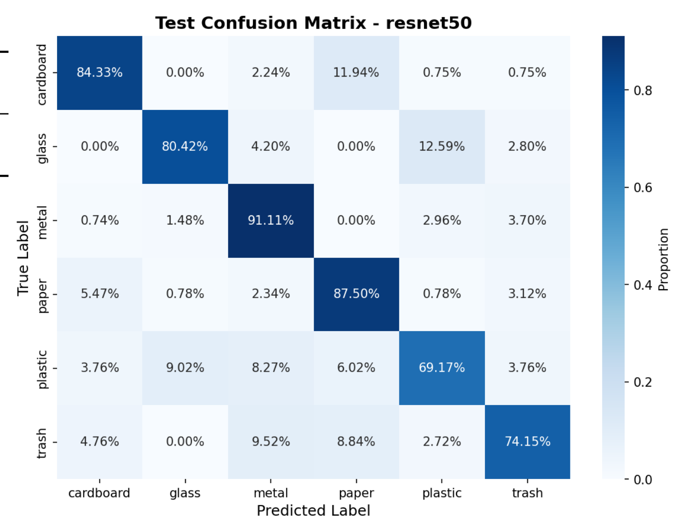

Fig. 14. Confusion Matrix for Kaggle in 6 type classification.

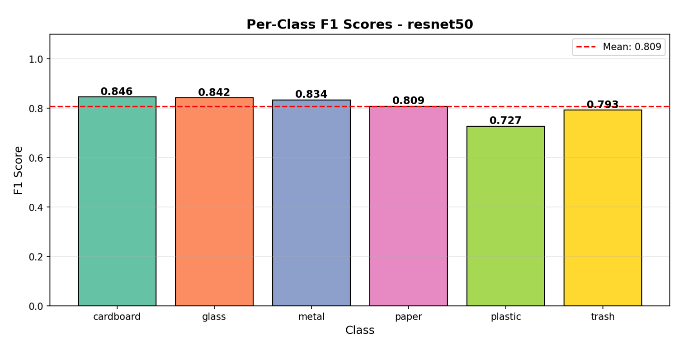

Fig. 15. F1 Scores per 6 classes with Kaggle and mean F1 Score.

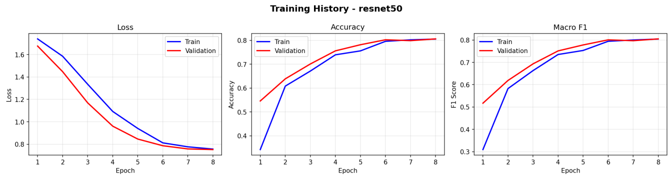

Fig. 16. Training History of ResNet over 8 epochs as it trains off Kaggle data.

suggesting significant learning challenges despite the dataset's controlled conditions.

4) Training Dynamics Analysis: Both models exhibited healthy learning dynamics with consistent training and validation loss reduction throughout the eight epochs. The Kaggle model showed particularly stable convergence, with training loss decreasing from 1.76 to 0.25 (85.8% reduction) and validation loss from 1.75 to 0.36 (79.4% reduction). The parallel decrease in both curves with minimal divergence (final gap: 0.11) indicates effective regularization and no overfitting.

The TrashNet model displayed similar directional trends but with higher final loss values (training: 1.34, validation: 1.35), suggesting incomplete convergence. Both models continued improving at epoch termination, indicating additional training epochs would likely yield further gains.

Accuracy progression revealed distinct learning phases. The Kaggle model showed rapid improvement in early epochs (22% to 71% train accuracy in epochs 1-3) followed by refinement (71% to 78% in epochs 4-8). The TrashNet model progressed more slowly (22% to 75% over eight epochs), potentially indicating greater intrinsic difficulty or suboptimal hyperparameters for this dataset.

5) Class wise performance: This performance hierarchy aligns with expected classification difficulties. Cardboard and plastic present distinctive visual textures and shapes, while glass exhibits transparency variations and background dependencies that complicate feature extraction. The "trash" category's heterogeneity inherently challenges classification consistency.

Kaggle: Per-class F1-scores revealed significant variation in model proficiency across waste categories:

- High performance: Cardboard (0.894) and plastic (0.873)
- Moderate performance: Paper (0.795) and metal (0.812)
- Lower performance: Glass (0.761) and trash (0.743)

TrashNet: The TrashNet model showed different performance patterns:

- Highest performance: Paper (0.749) and cardboard (0.742)
- Moderate performance: Plastic (0.712) and glass (0.707)
- Lowest performance: Metal (0.636) and trash (0.549)

6) Error Pattern Analysis: The confusion matrices (implied by F1-score disparities) suggest systematic rather than random errors. On both datasets, the lowest-performing classes (glass and trash) were likely misclassified into visually or semantically similar categories: glass confused with plastic or metal (reflective surfaces), and trash misclassified as various other categories depending on specific object characteristics.

This pattern indicates that the models learned meaningful feature representations but struggled with borderline cases and category ambiguities inherent in waste classification.

7) Limitations and Methodological Constraints: Several limitations warrant consideration in interpreting these results:

- Training duration: Eight epochs may be insufficient for complete convergence, particularly for the TrashNet model
- Hardware constraints: CPU-only training limited batch size and architectural experimentation
- Class imbalance: Unquantified class distribution differences may affect performance metrics
- Hyperparameter uniformity: Identical hyperparameters across datasets may be suboptimal for one or both
- Single architecture: Results limited to ResNet-50 without comparison to alternative architectures

8) Implications for Practical Deployment: The performance differential (80.4% vs 70.3%) highlights the critical importance of dataset characteristics in waste classification systems. Models trained on controlled laboratory-style datasets may not generalize effectively to real-world conditions. The superior performance on the more challenging but realistic Kaggle dataset suggests that diversity and realism in training data may outweigh perfect labeling and controlled conditions for practical applications.

The consistent difficulty with glass and trash classification across both datasets indicates these categories require specialized attention in waste sorting systems, potentially through ensemble methods, targeted data collection, or multi-stage classification approaches.

V. Discussion and Summary

This study systematically evaluated modern CNN architectures for automated waste classification across multiple datasets, revealing insights that extend beyond raw accuracy comparisons. While ConvNeXt-Tiny consistently achieved the highest performance across datasets, the results demonstrate that architectural sophistication alone does not determine success in practical waste classification systems. Instead, performance emerges from the interaction between architecture choice, transfer learning, and, most critically, dataset characteristics.

Transfer learning proved essential across all experiments. Pretrained models significantly outperformed randomly initialized counterparts, reinforcing that leveraging large-scale visual priors is non-negotiable for waste classification tasks, where labeled data is often limited. Across both TrashNet and Garbage Classification v2, all pretrained architectures converged reliably and achieved strong performance, suggesting that modern CNNs possess sufficient representational capacity for this domain when appropriately initialized.

A key finding is that dataset properties had a greater impact on performance than architectural differences. Identical ResNet-50 models achieved substantially higher accuracy and macro F1-score when trained on the larger, more diverse Kaggle dataset compared to the more controlled TrashNet dataset. This indicates that visual diversity and real-world variability are more valuable for generalization than clean backgrounds or uniform capture conditions. Models trained on controlled datasets may therefore overestimate real-world performance if dataset realism is not prioritized.

Despite overall strong results, all architectures consistently struggled with ambiguous categories, particularly mixed ”trash” and glass. These errors were systematic rather than random, with frequent confusion between visually similar materials such as metal and glass or paper and cardboard. The persistence of these patterns across datasets and architectures suggests that these challenges are inherent to the waste taxonomy itself rather than deficiencies in specific models.

From a deployment perspective, the results highlight clear trade-offs. ConvNeXt-Tiny is best suited for accuracy-critical applications, while MobileNetV3 variants provide a strong balance of performance and efficiency for edge and real-time systems. More broadly, the findings suggest that improving waste classification systems will require advances beyond model selection alone, through richer datasets, targeted handling of ambiguous categories, and potentially hierarchical or multi-stage classification strategies.

In conclusion, effective waste classification depends as much on data realism and task formulation as on architectural choice. Future work should explore domain adaptation across facilities, ensemble and hierarchical models for ambiguous classes, and multi-modal sensing to overcome inherent visual ambiguities. By aligning model design with realistic deployment conditions, automated waste sorting systems can move closer to reliable, scalable real-world adoption.

References

[1] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.

[2] Z. Liu *et al.*, “A ConvNet for the 2020s,” in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 11976–11986.

[3] A. Howard *et al.*, “Searching for MobileNetV3,” in *Proc. IEEE Int. Conf. Computer Vision (ICCV)*, 2019, pp. 1314–1324.

[4] M. Tan and Q. Le, “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,” in *Proc. Int. Conf. Machine Learning (ICML)*, 2019, pp. 6105–6114.

[5] S. J. Pan and Q. Yang, “A Survey on Transfer Learning,” *IEEE Trans. Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.

[6] C. Bircanoğlu, M. Atay, F. Beşer, O. Genç, and M. A. Kızrak, “RecycleNet: Intelligent Waste Sorting Using Deep Neural Networks,” in *Proc. IEEE Int. Conf. Innovations in Intelligent Systems and Applications (INISTA)*, 2018, pp. 1–7.

[7] V. Ruiz, Á. Sánchez, J. F. Vélez, and B. Raducanu, “Automatic Image-Based Waste Classification,” in *Proc. Int. Work-Conf. Interplay Between Natural and Artificial Computation (IWINAC)*, Lecture Notes in Computer Science, vol. 11487, Springer, 2019, pp. 422–431.

[8] A. H. Vo, L. H. Son, M. T. Vo, and T. Le, “A Novel Framework for Trash Classification Using Deep Transfer Learning,” *IEEE Access*, vol. 7, pp. 178631–178639, 2019.

[9] Q. Zhang, Q. Yang, X. Zhang, Q. Bao, J. Su, and X. Liu, “Waste Image Classification Based on Transfer Learning and Convolutional Neural Network,” *Waste Management*, vol. 135, pp. 150–157, 2021.