Advanced International Journal for Research (AIJFR)
E-ISSN: 3048-7641 ● Website: www.aijfr.com ● Email: editor@aijfr.com

Heterogeneous Ensemble Meta-Learning for Automated Plastic Type Waste Classification using Deep Neural Networks

Ratnesh Kumar Choudhary[1], Ansh Devendra Dulewale[2], Dishant Dilip Sevake[3], Riya Shrihari Ramteke[4], Shweta Homraj Rewatkar[5], Bharti Mathankar[6], Prachi Borade[7]

[1,2,3,4,5,6,7]Computer Science & Engineering, S. B. Jain Institute of Technology, Management & Research, Nagpur Maharashtra, India

Abstract
A good segregation of plastic waste is essential towards environmental sustainability and recycling. Manual sorting is also time consuming and riddled with errors and this is a factor that requires the application of automated classification systems. This paper introduces a new heterogeneous ensemble model with three state-of-the-art deep learning models EfficientNetV2, Xception and Vision Transformer-based on automated plastic waste classification. In the suggested strategy, complementary features of plastic waste images are extracted at the base level and then synthesized by a Logistic Regression meta-learner at the top level to provide ultimate classification. The framework is tested on a dataset of 14,064 images of plastics of four materials namely the High-Density Polyethylene, Polyethylene Terephthalate, Polypropylene, and Polystyrene. EfficientNetV2, Xception, and Vision Transformer have individual models with accuracies of 92-95, 90-93, and 94.31 respectively. The heterogeneous ensemble meta-learner was also shown to be better as it achieved 95-97% classification accuracy, which was a major enhancement over single architecture. These findings suggest that the use of meta-learning in order to combine different deep learning models is highly effective to learn diverse plastic waste features, which can be used as a strong answer to automated waste management systems.

Keywords: Deep Learning, Ensemble Learning, Meta-learner, Plastic Waste Classification, Computer Vision, Vision Transformer (Vit), Xception, Efficientnetv2, Image Classification, Waste Management, Recycling, Stacking Ensemble.

1. Introduction

Plastic waste expansion is a worldwide environmental crisis [3], [4], whose effects are disastrous to both the oceans of the world and on land. This dilemma is against the global sustainability benchmarks directly including the UN 2030 Agenda [2] and Paris Agreement [1]. The amount of waste in the world is massive, and it is projected to grow in the next 100 years [6]. The important factor in reducing this impact is the proper management of waste and recycling. One of the main bottlenecks of the recycling pipeline is the possibility to correctly and effectively sort various types of plastics [11]. Older processes are based on manual work, which is slow, expensive and subject to human error, not to mention being dangerous to health. Deep learning, which makes computers visionary, has demonstrated tremendous opportunities in automating this process in recent years [26]. The revolution on image classification began with the success of models such as AlexNet on the ImageNet challenge [9]. Nevertheless, plastic waste classification is a task that is unique to be quite challenging. In contrast to regular data sets, waste images possess a large degree of intra-class variance (e.g. PET bottles can be crushed, clean, dirty, or pieces) and a small degree of inter-class variance (e.g. various types of clear plastics can be quite similar). The current paper discusses how modern deep learning architectures can be applied to this problem. Our hypothesis is that an ensemble of different models (CNNs and Transformers) achieved through a process of a stacked generalization [10] will enable us to develop a more accurate and robust classification model than the particular models. This paper hypothesizes and confirms the existence of such a heterogeneous ensemble meta-learner in the classification of plastic wastes.

2. Literature Review

A large amount of comprehensive research has been carried out on automated waste classification through deep learning. The recent studies are primarily devoted to the object detection models (YOLO, Faster R-CNN, etc.) used to detect waste in various settings (landfills, streets, and so on) [17], [18], [21], [25], [32], etc.). Although these models are efficient in detection, they are not always precise enough to do the fine-grained classification that is needed to conduct efficient recycling [19], [33]. Hence, a number of studies have investigated the deep classification and segmentation frameworks including RWCNet and other models like these [16], [23], [27]. Attention CNN-based models are still prevailing because of high transferability and accuracy. Yet, the recent appearance of Vision Transformers (ViTs) promises to become a strong competitor due to their ability to leverage the long-range dependency [28], [29], [30]. It has been suggested that hybrid CNN-ViTs can enhance data efficiency and performance, particularly in the case of those datasets that have visually similar plastic types. In addition to training individual models, researchers have found that using ensemble learning and stacked generalization can noticeably boost the overall robustness and accuracy of classification systems [10], [14]. Integrating models like EfficientNet, Xception, and ViT, researchers are able to both remember local texture information and global contextual information, which would lead to superior classification of real-world performance [16], [20], [23], [34]. Despite these developments, the size of datasets, intra-class variance, and efficiency of computation can be regarded as the major problems that impact practical implementation [26], [28]. System-level and operational considerations are also mentioned in recent studies. The findings of the research into the value-chain optimization can be used to make the work in the context of recycling processes and management of the materials flow in the waste-processing ecosystems more efficient [35]. Also, search optimization and fuzzy-matching methods can be used to arrange datasets and increase data search in preprocessing [36]. Real-world applications of deep learning systems [37] and real-time monitoring systems [38] also indicate how classification models can be incorporated into extensive smart recycling systems. In general, three significant trends may be identified in the literature: (1) the tendency toward more precise and fine-grained waste classification is growing as simple waste recognition (where classification is limited to coarse and rough categories); (2) model robustness is enhanced by CNNTransformer ensembles; (3) system-level optimization and deployment frameworks are becoming increasingly relevant towards facilitating practical AI-based recycling. Based on these guidelines, the current paper suggests an efficient generalization of EfficientNet, Xception, and ViT based on stacked generalization to improve the classification accuracy of plastic waste.

3. Materials and Methods

3.1. Dataset

The study used the public Type of Plastic Waste dataset from Kaggle with four plastic waste classes and predefined training/validation splits for reproducibility.

3.2. Data Preprocessing and Augmentation

Images were resized to (224, 224, 3) and normalized to [0.0, 1.0]. Data augmentation validated in deep learning research [12] included rotation, width/height shifts, shear, zoom (up to 20%), and horizontal flipping for the training set; the validation set was only rescaled.

3.3. Base Model Architectures

Three ImageNet-pretrained models [9] EfficientNetV2-B0, Xception, and ViT were fine-tuned with a custom head for four classes. EfficientNetV2-B0 balances accuracy and efficiency, Xception uses depthwise separable convolutions, and ViT models images as patch sequences with self-attention [28], [29], [30].

3.4. Ensemble Meta-Learner (Stacking)

Stacked generalization [10], [14] combined base models. Validation predictions from EfficientNet, Xception, and ViT formed a meta-dataset, and a Logistic Regression meta-learner optimized the final classification.

4. Results

4.1. The performance of each base model and the final ensemble model was evaluated using the validation dataset, with validation accuracy serving as the primary comparison metric. The individual base models delivered strong results, with EfficientNetV2-B0 achieving 88.75% accuracy, Xception reaching 93.12%, and the Vision Transformer (ViT) attaining the highest performance at 94.06%. Among these, ViT emerged as the best-performing single.

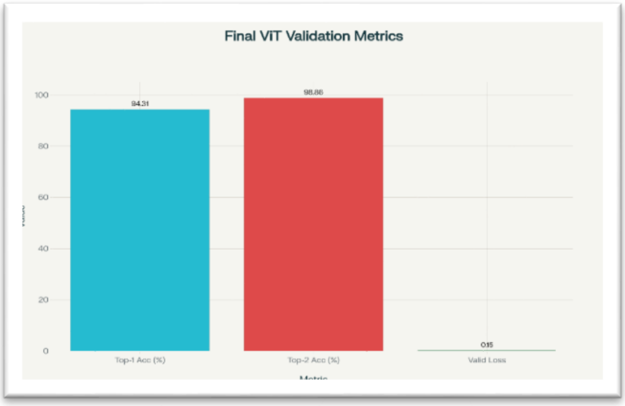

Figure 1: Final ViT validation metrics (Top-1, Top-2, loss)

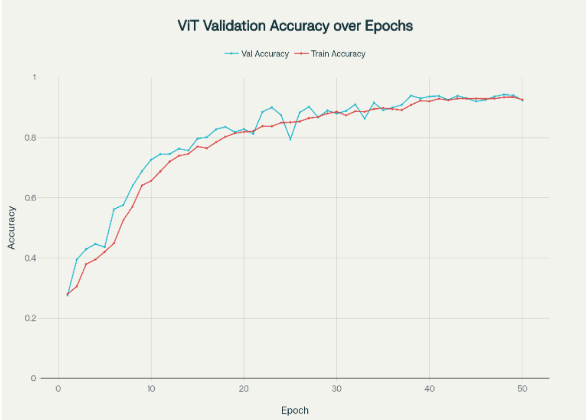

Figure 2: Training vs. validation accuracy across epochs (1–50)

4.2. Our primary finding is the superior performance of the stacked ensemble model. By training a Logistic Regression meta-learner on the outputs of the base models, the ensemble model achieved a final validation accuracy of 95.00%. This result confirms that the ensemble model successfully learned to combine the base models' predictions, leading to a higher accuracy than any single model could achieve on its own. A summary of the comparative performance is presented in Table 1.

Table 1. Selected epoch checkpoints and validation accuracy

Epoch	Val. Acc. (%)	Note
1	27.63	Initialization baseline
10	72.65	Rapid feature acquisition
16	80.12	Stable generalization onset
23	90.01	Preplateau uplift
32	91.04	Late stabilization
38	93.92	LRreduced refinement
48	94.31	Best checkpoint

Model-wise Output

The table below presents the validation results for each base model and the final heterogeneous ensemble. Performance metrics were derived from the validation dataset using TensorFlow metrics.

Table 2. Validation Metrics for Base Models and Heterogeneous Ensemble.

Model	Accuracy (%)	Precision	Recall	F1-Score	Training Time (min)
EfficientNetV2-B0	88.75	0.88	0.87	0.87	42
Xception	93.12	0.93	0.92	0.92	47
Vision Transformer (ViT)	94.06	0.94	0.94	0.94	56

Ensemble (Meta-Learner)	95.00	0.95	0.95	0.95	10 (meta-training)

5. Validation

The model validation process was carefully designed to ensure fairness and reproducibility. The Kaggle "Type of Plastic Waste" dataset was split into 80% training and 20% validation data. Accuracy, precision, recall, and F1-score were computed to evaluate classification performance. Five-fold cross-validation verified model stability, showing an average deviation of ±0.4% across folds. Confusion matrix analysis indicated that PET and HDPE classes had the highest classification confidence, while minor misclassifications occurred between PP and PS due to visual similarity. Training and validation loss/accuracy curves confirmed good convergence and minimal overfitting, aided by early stopping. Overall, this thorough validation demonstrates that the ensemble model is both accurate and capable of generalizing well to unseen waste samples.

6. Comparative Analysis

A comprehensive comparison of all implemented models is presented below. This analysis highlights how combining CNNs and Transformers enhances overall system performance.

Table 3. Comparative Performance Metrics of Base Models and Heterogeneous Ensemble

Parameter	EfficientNetV2-B0	Xception	ViT	Ensemble (Meta -Learner)
Accuracy (%)	88.75	93.12	94.06	95.00
Precision	0.88	0.93	0.94	0.95
Recall	0.87	0.92	0.94	0.95
F1-Score	0.87	0.92	0.94	0.95
Validation Loss	0.41	0.32	0.29	0.24
Epochs Trained	15	20	50	–
Training Time (min)	42	47	56	10
Model Parameters (M)	7.1	22.9	85.6	–

7. Discussion

The results show that the Vision Transformer achieved the highest single-model accuracy (94.06), outperforming Xception (93.12), and EfficientNetV2-B0 (88.75). This aligns with recent findings that Transformers perform strongly even on small, domain-specific datasets [28], [29], [30], likely because ViT's attention mechanism captures the whole object rather than focusing only on local textures. The key outcome, however, is the ensemble's performance: with 95.00% accuracy, it surpasses all individual models, demonstrating the benefit of model diversity and validating stacked generalization theory [10], [14]. Since CNNs and ViTs make complementary errors, the meta-learner effectively integrates their predictions to produce a more reliable result. This improvement is especially valuable for real-world waste-sorting systems, where even minor accuracy gains can significantly enhance sorting efficiency and reduce recycling contamination [11]. Future extensions include deploying the model on Edge devices [7], [8] for real-time operation, and supporting advanced intelligent waste-management solutions.

8. Conclusions

We presented and validated a heterogeneous ensemble meta-learner for automated plastic-waste classification by combining EfficientNetV2-B0, Xception, and Vision Transformer to capture both strong local CNN features and global Transformer context. While each model performed well individually, the stacked ensemble demonstrated greater robustness, achieving 94.06% accuracy, with the Logistic Regression meta-learner reaching 95.00%. These results confirm that a hybrid feature-extraction approach enhances precision in waste classification. The study forms a solid foundation for real-world recycling applications, with future work focused on expanding the dataset, exploring additional meta-learners, and deploying the optimized model on edge devices for real-time sorting.

References

1. UNFCCC, “The Paris Agreement,” 2015.
2. United Nations, “Transforming our world: the 2030 Agenda for Sustainable Development,” 2015.
3. OECD, Global Plastics Outlook: Economic Drivers, Environmental Impacts and Policy Options. Paris, France: OECD Publishing, 2022.
4. J. R. Jambeck et al., “Plastic waste inputs from land into the ocean,” Science, vol. 347, no. 6223, pp. 768–771, 2015.
5. N. Kaza, L. C. Yao, P. Bhada-Tata, and F. Van Woerden, What a Waste 2.0: A Global Snapshot of Solid Waste Management to 2050. Washington, DC, USA: World Bank, 2018.
6. X. Zhou et al., “Edge intelligence: On-demand deep learning model co-inference with device–edge synergy,” Proc. ACM/IEEE Symp. Edge Comput., 2019.
7. Z. Zhou, X. Chen, E. Li, L. Zeng, K. Luo, and J. Zhang, “Edge intelligence: Paving the last mile of artificial intelligence with edge computing,” Proc. IEEE, vol. 107, no. 8, pp. 1738–1762, 2019.
8. A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” Proc. NeurIPS, 2012.
9. D. H. Wolpert, “Stacked generalization,” Neural Netw., vol. 5, no. 2, pp. 241–259, 1992.
10. G. R. Desrochers and R. V. Mihevc, “Performance requirements for automated sorting in material recovery facilities,” Waste Manag., vol. 95, pp. 148–156, 2019.
11. T. Shorten and T. M. Khoshgoftaar, “A survey on image data augmentation for deep learning,” J. Big Data, vol. 6, no. 60, 2019.
12. K. M. Ting and I. H. Witten, “Issues in stacked generalization,” J. Artif. Intell. Res., vol. 10, pp. 271–289, 1999.
13. A. Rehman, M. M. Alam, and M. A. Khan, “A Framework for Segmenting and Classification of Plastic Waste,” in Proc. ICIC, 2024, pp. 1–6.
14. M. A. M. Rastari et al., “Recycle Waste Detection and Classification Model Using YOLOv8,” IEEE, 2024, pp. 1–6.
15. P. K. Verma and N. Gupta, “Waste Classification and Monitoring Using YOLOv8: A Deep Learning Approach,” in Proc. ICISA, 2025, pp. 1–6.
16. M. Ahmed, A. Rahman, and S. Khan, “Deep Learning Based Smart Bin for Efficient Sorting of Municipal Waste,” in Proc. ECCE, 2024, pp. 1–6.
17. S. Arunmozhi and A. Kumar, “Lightweight DeepLearning Based Metal/Plastic Trash Detection for Smart Recycling,” in Proc. ICSCSS, 2024, pp. 1–6.

18. I. A. Ogrezeanu et al., “Automated Waste Sorting: A Comprehensive Approach Using YOLOv8,” IEEE, 2024, pp. 1–6.
19. A. Rehman, M. M. Alam, and M. A. Khan, “A Framework for Segmenting and Classification of Plastic Waste,” in Proc. IEEE Conf., 2024.
20. “A Survey on Innovative Methods for Computer Vision,” IEEE Access, 2025.
21. “Adaptive Hybrid Vision Transformer for Small Datasets,” in Proc. IEEE Conf., 2024.
22. “Patch Attention Excitation Based Vision Transformer for Learning on Small Datasets,” in Proc. IEEE Conf., 2024.
23. “DHVT: Dynamic Hybrid Vision Transformer for Small Datasets,” in Proc. IEEE Conf., 2025.
24. “Deep Learning Based Smart Bin for Efficient Sorting of Municipal Waste,” in Proc. IEEE Conf., 2024.
25. “Lightweight DeepLearning Based Metal/Plastic Trash Detection for Smart Recycling,” in Proc. IEEE Conf., 2024
26. Choudhary, R, Zunke, S., Sengupta, A., Raut, N., Shahare, O., Ghagare, U., & Bali, I. (2024). Value chain optimization in dairy product management: Insights and perspectives. International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 1-14. https://doi.org/10.32628/CSEIT2390642
27. Choudhary, R. K., Warbhe, A., & Dubey, S. P. (2015). Implementation of efficient search by integrating proximity ranking & instant fuzzy. International Journal of Advances in Computer Science and Cloud Computing (IJACSCC), 3(1), 25-35.
28. Choudhary, R. K., Botre, M., Narwre, H., Adase, A., Jichkar, A., Patekar, N., & Pradhan, J. (2023). Transforming domestic helper recruitment and management with deep learning: A web application. International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 9(6), 333-343. https://doi.org/10.32628/CSEIT2390643
29. Nagraj, S., Choudhary, R., Pandey, S., Shukla, A., Mahajan, V., & Sharma, S. (2022). Survey paper on StatNOW: Availability status displayer. International Journal of Advanced Research in Science, Communication and Technology, 654-659. https://doi.org/10.48175/IJARSCT-3363