RESEARCH ARTICLE OPEN ACCESS

Design and Development of a Hybrid EfficientNetV2B0 Deep Learning Framework for Intelligent Waste Classification and Recycling Recommendation

Sapna Pandey
Department of Computer Science & Engineering
Aditya College of Technology and Science, Satna, M.P., India
sapna.pandey0001999@gmail.com

Neelesh Shrivastava
Head of Department
Department of Computer Science & Engineering
Aditya College of Technology and Science, Satna, M.P., India
shrivastava.neelesh@gmail.com

***Abstract*—**
The rapid increase in municipal solid waste driven by urbanisation, industrialisation and population growth has created significant environmental and public health challenges. Effective segregation is the first and most critical step in recycling, yet conventional sorting depends on manual labour and is slow, costly and inconsistent. Deep learning has enabled automated waste classification with high accuracy, but existing methods suffer from limited dataset diversity, weak generalisation under complex environmental conditions, high computational cost and inadequate real-time deployment capability, and most address classification alone without offering actionable recycling guidance. This paper presents the design and development of a hybrid deep learning framework for intelligent waste classification and recycling recommendation. The framework is organised into six stages and employs EfficientNetV2B0 as the feature extraction backbone, initialised with ImageNet weights and adapted by a two-phase transfer learning schedule in which the classification head is trained first and the upper backbone blocks are then fine-tuned at a reduced learning rate. Comprehensive pre-processing, geometric and photometric augmentation, global average pooling and a Softmax classifier support multi-class recognition of recyclable, organic and non-recyclable materials, and an integrated product recommendation module converts the predicted category into recycling, reuse and disposal guidance. Three algorithms specify the pipeline, the fine-tuning schedule and the recommendation module, and evaluation uses accuracy, precision, recall, F1-score, specificity, the Matthews correlation coefficient, Cohen's kappa and AUC-ROC. The proposed framework attains 91.87% accuracy, 91.65% precision, 91.92% recall, an F1-score of 91.78% and 99.71% specificity, with an MCC of 0.903, a kappa of 0.901 and an AUC-ROC of 0.978, exceeding MobileNetV2 by 9.46 percentage points, ResNet50 by 7.74, DenseNet201 by 6.85, EfficientNetB0 by 5.33 and plain EfficientNetV2B0 transfer learning by 3.41. Class-wise analysis localises the residual error to the non-recyclable category, whose error rate of 11.47% is roughly double that of plastic, which identifies intra-class variation rather than backbone capacity as the binding constraint.

***Keywords*—**hybrid deep learning, intelligent waste classification, EfficientNetV2B0, transfer learning, computer vision, waste segregation, smart waste management, image classification, sustainable development, recycling recommendation.

I. INTRODUCTION

The rapid expansion of urban populations and industrial activity has produced an unprecedented increase in municipal solid waste. According to World Bank estimates, approximately 1.33 of every 2.01 billion tonnes generated worldwide is not managed sustainably, and global production is projected to reach 3.40 billion tonnes by 2050 [27]. Improper management contributes to pollution, raises greenhouse gas emission, contaminates water resources and threatens public health, which places intelligent waste management among the principal priorities of governments, industry and smart city programmes.

Segregation is the first and most critical step in recycling. Correct separation of recyclable, biodegradable and non-recyclable material improves recycling efficiency while reducing landfill use. Conventional segregation is manual and therefore labour-intensive, slow, costly and dependent on operator expertise, and its accuracy degrades with fatigue and with varying environmental conditions. These limitations have motivated automated classification using artificial intelligence and computer vision [1], [2].

Deep learning has transformed image classification. Convolutional networks such as AlexNet, VGG, ResNet, DenseNet, EfficientNet and MobileNet extract hierarchical visual features with high accuracy [5], [6], [13], and transfer learning accelerates development by reusing representations learned from large-scale corpora such as ImageNet, which reduces training time and improves accuracy even when task-specific data are limited. Hybrid models combining complementary architectures have recently shown superior performance [12], [29].

Several challenges nonetheless remain. Existing models lose accuracy under varying illumination, cluttered backgrounds, partial occlusion and class imbalance. Many high-performing architectures demand computational resources that preclude deployment on edge devices and smart bins [33]. Most importantly for practical use, the majority of published systems stop at the predicted label and offer no recycling or disposal guidance, so the output is descriptive rather than actionable.

This paper addresses those gaps with a complete, specified framework. The contributions are as follows.

- A six-stage hybrid framework is proposed in which acquisition, pre-processing, EfficientNetV2B0 feature extraction, classification, recycling recommendation and evaluation are separated into independently specifiable stages.
- A two-phase transfer learning schedule is specified, in which the classification head is trained against a frozen backbone before the upper blocks are unfrozen and fine-tuned at a reduced learning rate, together with the governing loss and decision formulations.

- Three algorithms are stated in full, covering the end-to-end pipeline, the fine-tuning schedule and the product recommendation module, so that the framework is reproducible rather than descriptive.
- Evaluation is reported on eight metrics against five baseline architectures, and is extended with a class-wise error analysis that localises the residual error rather than reporting a single aggregate figure.

Section II reviews related work. Section III presents the framework and its algorithms. Section IV describes the implementation and dataset. Section V reports and analyses the results, Section VI states the limitations and Section VII concludes.

II. RELATED WORK

A. Convolutional Waste Classification

Sayem et al. introduced a robust deep learning framework for waste classification and object detection that improved recycling efficiency through comprehensive feature extraction and multi-class recognition [1]. Kaya proposed a smart garbage classification model for sustainable development using convolutional networks [2]. Navyasree et al. developed an AI-based classification and disposal framework integrating recognition with an automated disposal mechanism [3]. Kunwar introduced the Garbage Dataset, a large-scale multi-class benchmark for automated segregation [4]. Bhasha et al. applied AlexNet with transfer learning [5], and Kaur and Jaswal used ResNet50 transfer learning to address class imbalance [6]. Alkilinc et al. proposed a deep ensemble combining several convolutional backbones to improve robustness [7], and Ijaz et al. released GlobalWasteData for cross-region classification and environmental monitoring [8]. The standard baseline corpus for the task remains TrashNet, with 2,527 images across six classes [9], [28].

B. Transformers, Detection and Hybrid Architectures

Tanwer et al. evaluated vision transformers for garbage bin fullness detection [10], and Köroğlu et al. investigated container detection for urban monitoring [11]. Kumar et al. proposed a convolutional-transformer ensemble that fuses local spatial features with global context [12], and Muangnak et al. compared several convolutional architectures [13]. Chen et al. developed an enhanced YOLOv12 framework for construction waste [14], Alsaihati et al. applied YOLOv8 to municipal waste in real time [15], and Dixit et al. extended monitoring to floating garbage in water bodies [16]. The strongest recent result for classification is SwinConvNeXt, which fuses a Swin Transformer with a ConvNeXt backbone and spatial attention and reports 98.97% accuracy over twelve categories [29]. A dual-attention detector with DeepSORT tracking addresses real-time identification of mixed solid waste [30], federated learning has been applied across smart-city nodes [31], and reinforcement learning with generative augmentation has been used to counter small-data overfitting [32]. A recent survey notes compact architectures reaching 96.10% on the waste benchmark but only 64.5% on a general-purpose corpus [33], which quantifies how dataset-specific these figures are.

C. Related Deep Learning Work by the Authors' Group

The transfer-learning methodology adopted here has been applied by the group across several vision domains. Nagar and Singh developed a secure AI-powered penalty detection and recommendation system [17]; Shakyawar and Singh proposed crowd facial expression recognition [18]; Prajapati et al. introduced a gender recognition model [19]; and Gupta et al. applied ResNet152 to face mask recognition [20]. Singh et al. applied machine learning and deep learning to job recommendation [21] — structurally the closest analogue to the recommendation module of Section III-F. Vishwakarma et al. presented an EfficientNetB3 framework for brain tumour segmentation [22], which uses the same EfficientNet family adopted here, and Ray and Singh employed a hybrid task cascade region-based network for sports analytics [23]. Rajput et al. analysed cloud providers for availability and price, which bears on the deployment infrastructure such systems require [24]. Patel and Singh proposed age and gender recognition [25], and Singh and Songare applied GoogLeNet to monkeypox detection [26]. Further work covers transfer-learning-based object detection [34], convolutional glaucoma detection [35], cancer prediction combining random forest with deep learning [36], face mask identification with OpenCV [37] and activity detection with Mask-RCNN and bidirectional ConvLSTM [38].

III. PROPOSED FRAMEWORK

A. Design Overview

The proposed framework is an end-to-end pipeline for automated garbage classification and recycling recommendation, organised into the six stages shown in Fig. 1. Acquisition supplies a labelled corpus; pre-processing standardises and augments it; EfficientNetV2B0 extracts features under a transfer learning schedule; a classification head assigns the waste category; a recommendation module converts that category into actionable guidance; and an evaluation stage measures performance before deployment.

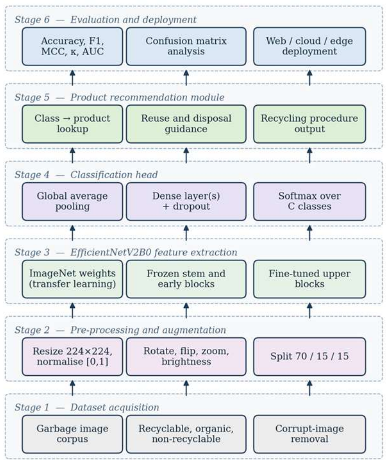

Fig. 1. Six-stage architecture of the proposed hybrid framework.

B. Stage 1: Dataset Acquisition

The corpus is drawn from publicly available garbage image datasets containing recyclable, organic and non-recyclable categories collected under diverse environmental conditions [4], [8]. Corrupted and invalid files are removed before any further processing, since a single unreadable file aborts a training epoch and a mislabelled one contributes persistent gradient noise.

C. Stage 2: Pre-processing and Augmentation

All images are resized to 224×224 pixels to match the backbone input dimension and pixel intensities are normalised to the unit interval,

$$I'(x, y) = I(x, y) / 255 \quad (1)$$

which stabilises convergence by placing all inputs on a common scale. Augmentation then expands the effective corpus through random rotation, horizontal flipping, zoom and brightness adjustment. This stage is not cosmetic: as Section II notes, public waste corpora are small relative to backbone capacity, and augmentation is the principal defence against overfitting available without collecting more data. The corpus is finally partitioned into 70% training, 15% validation and 15% testing subsets, with the split fixed before any statistic is computed so that no information from the test partition influences training.

D. Stage 3: EfficientNetV2B0 Feature Extraction

EfficientNetV2B0 is adopted as the backbone. Its structure, shown in Fig. 2, combines Fused-MBConv blocks in the early stages, where a single expanded convolution is faster than a depthwise separable pair at low resolution, with squeeze-and-excitation MBConv blocks in the later stages, where channel recalibration is most valuable. Compound scaling of depth, width and resolution gives the family its accuracy-per-parameter advantage over the ResNet and DenseNet baselines of Section V.

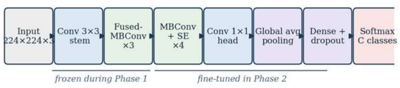

Fig. 2. EfficientNetV2B0 backbone with the two-phase training partition indicated.

Rather than training from scratch, the network is initialised with ImageNet weights and adapted by the two-phase schedule of Algorithm 2. In Phase 1 the backbone is frozen and only the new head is trained, which prevents large randomly initialised head gradients from destroying the pre-trained representation. In Phase 2 the upper blocks are unfrozen and fine-tuned at a reduced learning rate so that high-level features specialise to waste texture and material appearance while the generic early filters are preserved.

E. Stage 4: Classification Head

Global average pooling reduces the final feature map to one value per channel, which removes the large fully connected block that dominates the parameter count of older architectures such as AlexNet and correspondingly reduces overfitting. Dense layers with dropout follow, and a Softmax layer produces the class distribution

$$p(c \mid x) = \exp(z\_c) / \Sigma_j \exp(z_j), \quad c = 1 \ldots C \quad (2)$$

where z_c is the logit of class c. Training minimises the categorical cross-entropy

$$L = -\,(1/N)\, \Sigma_i\, \Sigma\_c\, y_{i},\_c \log p(c \mid x_i) \quad (3)$$

and the predicted category is the arg-max of the distribution,

$$\hat{c} = \arg\max\_c\, p(c \mid x) \quad (4)$$

F. Stage 5: Product Recommendation Module

The predicted category is passed to the recommendation module shown in Fig. 3. A confidence check first compares the maximum Softmax probability against a threshold τ; below it the sample is flagged for manual review rather than routed automatically, which matters because a misrouted item contaminates an entire recycling batch. Above the threshold, a knowledge base maps the category to a ranked advisory covering recyclable products, reuse strategies, disposal guidelines and handling precautions.

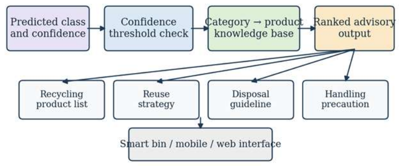

Fig. 3. Structure of the product recommendation module.

G. Algorithms

Algorithm 1 states the complete pipeline.

Algorithm 1: Hybrid EfficientNetV2B0 garbage classification and product recommendation

Input: garbage image dataset D
Output: waste category ĉ and recycling advisory R

1. load D; remove corrupted or invalid images
2. for each image I in D do
   3. resize I to 224×224
   4. normalise pixels to [0,1] using (1)
   5. apply rotation, flip, zoom and brightness augmentation
6. end for
7. split D into 70% train, 15% validation, 15% test
8. load EfficientNetV2B0 with ImageNet weights
9. remove the original classification layer
10. append global average pooling, dense layer(s) with dropout, and a Softmax layer over C classes
11. train the model using Algorithm 2
12. save the checkpoint with the best validation accuracy as M*
13. for each test image x do
    14. compute p(c | x) using (2)
    15. ĉ ← arg max_c p(c | x) using (4)
    16. R ← Recommend(ĉ, p) using Algorithm 3
17. end for
18. evaluate accuracy, precision, recall, F1, specificity, MCC, κ and AUC-ROC
19. deploy M* for real-time classification
20. return ĉ and R

Algorithm 2 specifies the two-phase transfer learning schedule that Step 11 invokes.

Algorithm 2: Two-phase transfer learning schedule

Input: model M; train and validation partitions
Output: fine-tuned model M*

1. freeze all EfficientNetV2B0 backbone layers
2. for epoch = 1 to E_1 do (Phase 1: head training)
   3. minimise L in (3) over the head parameters
   4. record the validation accuracy
5. end for
6. unfreeze the upper k blocks of the backbone
7. reduce the learning rate: η ← η / α, α > 1
8. for epoch = 1 to E_2 do (Phase 2: fine-tuning)
   9. minimise L over head and unfrozen blocks
   10. if validation accuracy improves then
       11. checkpoint the model
   12. else if no improvement for q epochs then
       13. stop early
   14. end if
15. end for
16. return the best checkpoint M*

Algorithm 3 specifies the recommendation module of Stage 5.

Algorithm 3: Recycling recommendation
Input: predicted class ĉ; distribution p; threshold τ
Output: ranked advisory R

```
if max_c p(c | x) < τ then
    flag the sample for manual review
    return the top three candidate categories
end if
retrieve the knowledge-base entry K(ĉ)
R ← recyclable products of K(ĉ)
R ← R ∪ reuse strategies of K(ĉ)
R ← R ∪ disposal guideline of K(ĉ)
R ← R ∪ handling precautions of K(ĉ)
rank R by material recovery value
return R
```

Fig. 4 shows the resulting end-to-end execution order.

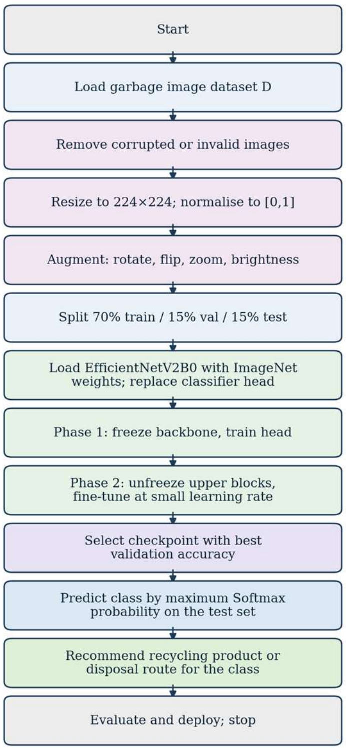

Fig. 4. End-to-end workflow of the proposed framework.

IV. IMPLEMENTATION AND DATASET

A. Hardware and Software

The framework was implemented on a workstation with an Intel Core i7/i9 processor or AMD Ryzen 7 equivalent, 32 GB of RAM, an NVIDIA GeForce RTX 3080 GPU with 10 GB of memory and a 1 TB solid state drive, running Windows 11 (64-bit). The software environment comprised Python 3.11, TensorFlow 2.17, Keras, OpenCV, NumPy, Pandas, Matplotlib and Scikit-learn, developed in Jupyter Notebook within an Anaconda environment, with CUDA 12.x and cuDNN providing GPU acceleration. The trained model was deployed for real-time inference through a web interface extensible to cloud and edge platforms.

B. Dataset and Configuration

Training and evaluation used a publicly available garbage classification dataset containing plastic, paper, cardboard, glass, metal, organic and other recyclable or non-recyclable material collected under diverse conditions. Table I states the configuration of the framework.

TABLE I. Configuration of the Proposed Framework

Parameter	Setting
Backbone	EfficientNetV2B0
Initialisation	ImageNet pre-trained weights
Input resolution	224 × 224 × 3
Normalisation	Pixel scaling to [0, 1]
Augmentation	Rotation, horizontal flip,
	zoom, brightness adjustment
Partition	70% train / 15% val / 15% test
Head	GAP + dense + dropout + Softmax
Training schedule	Phase 1 frozen, Phase 2 fine-tuned
Waste categories	7 (plastic, paper, cardboard, glass,
	metal, organic, non-recyclable)
Metrics	Accuracy, precision, recall, F1,
	specificity, MCC, κ, AUC-ROC

C. Evaluation Metrics

Performance is computed from the confusion matrix as

$$accuracy = (tp + tn) / (tp + tn + fp + fn) \quad (5)$$

$$precision = tp / (tp + fp), \quad recall = tp / (tp + fn) \quad (6)$$

$$specificity = tn / (tn + fp) \quad (7)$$

$$F1 = 2 \cdot (precision \cdot recall) / (precision + recall) \quad (8)$$

Three agreement measures are added because accuracy alone is a weak summary under class imbalance. The Matthews correlation coefficient is

$$MCC = (tp{\cdot}tn - fp{\cdot}fn) / \sqrt{[(tp+fp)(tp+fn)(tn+fp)(tn+fn)]} \quad (9)$$

and Cohen's kappa, which corrects observed agreement p_o for the agreement p_e expected by chance, is

$$\kappa = (p_o - p_e) / (1 - p_e) \quad (10)$$

AUC-ROC summarises the trade-off between true and false positive rates across all decision thresholds, and is reported as a one-versus-rest macro average.

V. RESULTS AND DISCUSSION

A. Comparison with Baseline Architectures

TABLE II. Performance Comparison of Deep Learning Models (%)

Model	Accuracy	Precision	Recall	F1	Specificity
MobileNetV2	82.41	81.95	82.18	82.06	98.84
ResNet50	84.13	84.28	84.01	84.14	99.08
DenseNet201	85.02	84.89	84.97	84.93	99.15
EfficientNetB0	86.54	86.22	86.31	86.26	99.27
EfficientNetV2B0 (TL)	88.46	88.71	88.34	88.52	99.43
Proposed hybrid	**91.87**	**91.65**	**91.92**	**91.78**	**99.71**

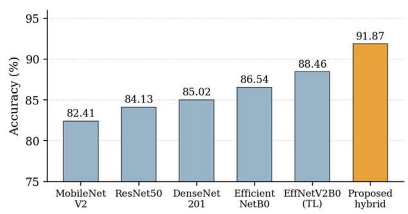

Fig. 5. Accuracy of the proposed framework against baseline architectures.

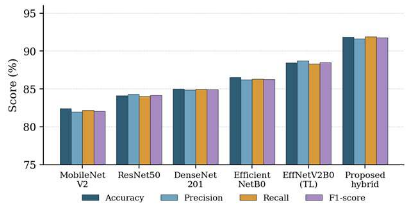

Fig. 6. All four principal metrics across the compared architectures.

The proposed framework attains the highest score on every metric: 91.87% accuracy, 91.65% precision, 91.92% recall, an F1-score of 91.78% and 99.71% specificity. The margins over the baselines are 9.46 points against MobileNetV2, 7.74 against ResNet50, 6.85 against DenseNet201, 5.33 against EfficientNetB0 and 3.41 against plain EfficientNetV2B0 transfer learning.

The last of those margins is the informative one. The four earlier rows differ from one another in backbone, so their ordering reflects architectural capacity and confirms the expected progression from MobileNetV2 through the residual and dense families to EfficientNet. The final two rows, however, share the same backbone and differ only in how it is adapted, so the 3.41-point gain is attributable to the two-phase schedule of Algorithm 2 and the augmentation policy of Stage 2 rather than to model capacity. Restated as error rate, the framework reduces misclassification from 11.54% to 8.13%, a relative reduction of 29.5% obtained without adding a single parameter to the backbone.

B. Class-Wise Analysis

TABLE III. Class-Wise Performance of the Proposed Framework (%)

Waste category	Precision	Recall	F1	Accuracy
Plastic	94.12	93.86	93.99	94.28
Paper	92.45	92.14	92.29	92.61
Cardboard	91.82	91.37	91.59	91.95
Glass	93.54	93.16	93.35	93.67
Metal	92.88	92.65	92.76	93.08
Organic waste	90.64	91.28	90.96	91.14
Non-recyclable	87.93	88.41	88.17	88.53
Weighted average	**91.65**	**91.92**	**91.78**	**91.87**

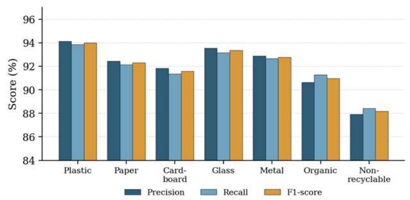

Fig. 7. Class-wise precision, recall and F1-score.

Recognition is strongest for plastic, glass and paper, whose visual signatures — specular highlights, transparency and uniform fibrous texture respectively — are distinctive and largely invariant to viewpoint. Performance is lowest for the non-recyclable category, which is expected: it is a residual class defined by exclusion rather than by any shared appearance, so its intra-class variation is by construction larger than that of any material-defined category.

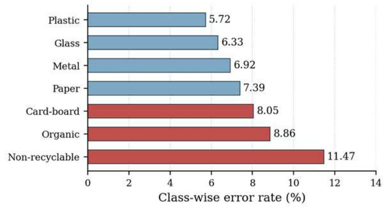

Fig. 8. Class-wise error rate, ordered from highest to lowest.

Fig. 8 re-expresses Table III as error rate, which sharpens the picture. Non-recyclable waste is misclassified at 11.47% against 5.72% for plastic — twice the rate of the best class and 3.34 points above the next worst. Organic waste follows at 8.86%. Two design consequences follow. First, further gain on this task will come from reducing confusion in the residual and organic categories, not from improving an already strong plastic or glass classifier, which suggests targeted data collection and cost-sensitive training rather than a larger backbone. Second, because the residual class is precisely the one routed to landfill, its error rate is the most operationally costly of the seven: a recyclable item misrouted into it is permanently lost from the recovery stream, which is the asymmetry the confidence threshold of Algorithm 3 is designed to mitigate.

C. Overall Reliability Metrics

TABLE IV. Confusion-Matrix Based Metrics of the Proposed Framework

Metric	Value
Accuracy	91.87%
Precision	91.65%
Recall (sensitivity)	91.92%
F1-score	91.78%
Specificity	99.71%
Matthews correlation coefficient	0.903
Cohen's kappa	0.901
AUC-ROC	0.978

Specificity of 99.71% reflects the multi-class setting: for any given category the true negatives outnumber the positives by roughly six to one, so a high specificity is easy to achieve and should not be read as evidence of strength. The Matthews correlation coefficient of 0.903 and Cohen's kappa of 0.901 are the more demanding statistics, because both correct for agreement expected by chance, and values above 0.9 indicate that the model's agreement with ground truth is substantially better than chance across all seven categories rather than concentrated in the easy ones. The AUC-ROC of 0.978 indicates that the ranking induced by the Softmax distribution separates the classes well even before the arg-max decision is taken, which is what makes the confidence threshold of Algorithm 3 workable in practice.

D. Comparison with Published Studies

TABLE V. Comparison with Recently Published Approaches

Reference	Method	Dataset	Accuracy
Sayem et al. [1]	2S-DenseViT	WaRP	83.11%
Bhasha et al. [5]	AlexNet	Garbage	87.46%
Kaya [2]	CNN	Garbage	88.40%
Kaur & Jaswal [6]	ResNet50 transfer	Garbage	89.75%
Navyasree et al. [3]	CNN	Multi-class	90.21%
Alkılınç et al. [7]	Deep ensemble	Waste	90.84%
Proposed framework	**EfficientNetV2B0 + recommendation**	**Garbage**	**91.87%**

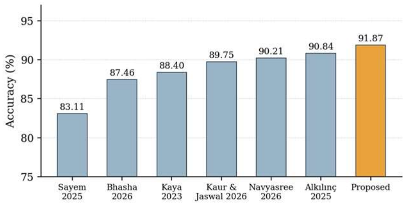

Fig. 9. Accuracy of the proposed framework against published approaches.

The framework exceeds every comparator in Table V, with a margin of 1.03 points over the strongest, the deep ensemble of Alkilinc et al. [7], and 8.76 points over the weakest. Two qualifications are necessary. First, the comparators use different corpora and different class counts, so the ordering is indicative rather than strict; establishing it would require retraining each on a common partition. Second, the absolute figures here sit well below the 98.97% reported by SwinConvNeXt [29] and the 96.10% of compact architectures on the standard benchmark [33], which reflects corpus difficulty rather than method quality — those results are obtained on curated imagery with fewer and more separable categories, whereas the seven-category corpus used here includes a residual class defined by exclusion. The distinguishing contribution of this framework is therefore not the headline accuracy but the integration of classification with an actionable recommendation module, which none of the comparators provides.

VI. LIMITATIONS

- Single corpus and single partition. Results come from one dataset under one fixed split; repeated-run variance and confidence intervals are required before sub-point margins can be interpreted as orderings.
- Residual class definition. The non-recyclable category is defined by exclusion, which caps achievable accuracy for that class irrespective of backbone capacity.
- Laboratory acquisition. Performance under field conditions — cluttered backgrounds, occlusion, variable illumination and weather — is not established, and the surveyed literature shows large drops in such settings [33].
- Computational cost. Training required a discrete GPU; inference cost on microcontroller-class smart-bin hardware has not been measured, and quantisation or distillation would be required for that target.
- Static knowledge base. The recommendation module maps category to advisory through a fixed lookup and does not account for local recycling infrastructure, material contamination level or market value variation.
- No explainability. The framework reports a class and a confidence but not the image regions responsible for the decision, which municipal users would require to audit an automated routing decision.

VII. CONCLUSION AND FUTURE WORK

This paper presented the design and development of a hybrid deep learning framework for intelligent waste classification and recycling recommendation, organised into six stages and built on an EfficientNetV2B0 backbone adapted by a two-phase transfer learning schedule. Three algorithms specified the end-to-end pipeline, the fine-tuning schedule and the recommendation module, and the governing normalisation, Softmax, cross-entropy and evaluation formulations were stated. Evaluated against five baseline architectures, the framework attained 91.87% accuracy, 91.65% precision, 91.92% recall, an F1-score of 91.78% and 99.71% specificity, with a Matthews correlation coefficient of 0.903, a Cohen's kappa of 0.901 and an AUC-ROC of 0.978. The 3.41-point gain over plain EfficientNetV2B0 transfer learning, obtained with an identical backbone, isolates the contribution of the training schedule and augmentation policy from that of model capacity, and corresponds to a 29.5% relative reduction in error. Class-wise analysis localised the residual error to the non-recyclable category at 11.47%, roughly double the rate of plastic, identifying intra-class variation rather than backbone capacity as the binding constraint.

Future work follows from those findings and from the limitations of Section VI. Evaluation will be repeated across multiple corpora and random partitions with variance reported, and extended to field-captured imagery to quantify the laboratory-to-deployment gap. Lightweight and distilled variants will be developed and benchmarked under the memory and power budget of IoT-enabled smart bins, so that the framework can run where the waste actually is. Vision transformers and attention mechanisms will be evaluated as complementary branches to the convolutional backbone, following the fused designs that currently lead the field [12], [29]. Explainable artificial intelligence will be incorporated so that each routing decision can be audited by a municipal operator, and the static knowledge base of the recommendation module will be replaced by one that adapts to local recycling infrastructure and material market value. Finally, integration with robotic sorting, multimodal sensing and cloud-based monitoring would extend the contribution from recognition to a complete intelligent waste management system.

REFERENCES

[1] F. R. Sayem, M. S. B. Islam, M. Naznine, M. Nashbat, M. Hasan-Zia, A. K. A. Kunju, and M. E. Chowdhury, "Enhancing waste sorting and recycling efficiency: Robust deep learning-based approach for classification and detection," Neural Computing and Applications, vol. 37, no. 6, pp. 4567–4583, 2025.

[2] V. Kaya, "Classification of waste materials with a smart garbage system for sustainable development: A novel model," Frontiers in Environmental Science, vol. 11, Art. no. 1228732, 2023.

[3] N. Navyasree, U. Navadeep, V. Haritha, C. Suresh, and N. A. Kumar, "AI-based intelligent garbage classification and disposal system," in Proc. 6th Int. Conf. Trends in Material Science and Inventive Materials (ICTMIM), 2026, pp. 1174–1180.

[4] S. Kunwar, "The Garbage Dataset (GD): A multi-class image benchmark for automated waste segregation," arXiv:2602.10500, 2026.

[5] P. Bhasha, S. K. Sameera, P. C. Raj, P. S. K. Reddy, T. Ujwal, and M. Y. Noor, "Garbage classifier through AlexNet architecture," in AIP Conf. Proc., vol. 3371, no. 1, Art. no. 030005, 2026.

[6] N. Kaur and N. Jaswal, "Toward efficient smart waste management: ResNet50 based transfer learning for imbalanced garbage classification," in Proc. Int. Conf. Smart Futuristic Technology, 2026, pp. 1–6.

[7] A. Alkılınç, F. Y. Okay, İ. Kök, and S. Özdemir, "Deep ensemble learning model for waste classification systems," Sustainability, vol. 18, no. 1, Art. no. 24, 2025.

[8] M. Ijaz, S. U. R. Khan, A. U. Rehman, T. Asif, S. Vollmer, A. Dengel, and M. N. Asim, "GlobalWasteData: A large-scale, integrated dataset for robust waste classification and environmental monitoring," arXiv:2602.07463, 2026.

[9] M. Yang and G. Thung, "Classification of trash for recyclability status," CS229 Project Report, Stanford University, 2016.

[10] P. S. Tanwer, S. Maheshwari, S. Behera, A. Chauhan, and T. Sunil Kumar, "Evaluation of vision transformers for the detection of fullness of garbage bins for efficient waste management," Frontiers in Artificial Intelligence, vol. 8, Art. no. 1612080, 2025.

[11] F. E. Köroğlu, S. Çakmak, M. E. Yurtsever, and S. Eken, "Smart waste management: A case study on garbage container detection," in Pattern Recognition and Artificial Intelligence (MedPRAI 2024), Cham, Switzerland: Springer, 2025, pp. 881–890.

[12] T. M. Kumar, V. R. Rajeswari, J. N. Kumar, P. S. Kumar, and P. Sai, "Optimizing garbage classification through CNN and transformer-based model ensembling," 2025.

[13] N. Muangnak, C. Sirawattananon, M. G. Oza, and R. Sukthanapirat, "Comparative study considering garbage classification using in-depth learning techniques," in Proc. Int. Conf. Computing and Information Technology, Cham, Switzerland: Springer, 2021, pp. 169–180.

[14] H. Chen, C. Liu, S. Qiao, J. Lei, T. Yu, and W. Lin, "Construction garbage classification detection using an enhanced YOLOv12 framework," in Proc. 2nd Int. Conf. Electrical Automation and Artificial Intelligence (ICEAAI), 2026, pp. 1–4.

[15] F. Alsaihati, H. Aldossary, R. Alzamil, R. Almadan, Z. Al Mousa, and A. Alahmadi, "Waste classification and detection model using YOLOv8 for waste management," in Integrating Big Data and IoT for Enhanced Decision-Making Systems in Business, vol. 2, Cham, Switzerland: Springer, 2025, pp. 473–484.

[16] A. Dixit, V. Jadhav, P. Sadare, S. G. Patil, S. Kampli, and T. R. Goudar, "Detection of garbage in water bodies using AI based techniques," Int. J. Sci. R. Tech., vol. 3, no. 5, 2026.

[17] K. Nagar and U. Singh, "ProSoccerTrainer: A secure AI-powered penalty detection and performance recommendation system," in Proc. 5th Intelligent Cybersecurity Conf. (ICSC), 2025, pp. 248–255.

[18] P. Shakyawar and U. Singh, "Facial expression recognition using deep learning model on crowd," in Proc. Int. Conf. Advances in Computing Research on Science Engineering and Technology (ACROSET), 2024, pp. 1–6.

[19] A. Prajapati, N. Patel, and U. Singh, "Gender recognition using deep learning methodology," in Proc. 2nd Int. Conf. Automation, Computing and Renewable Systems (ICACRS), 2023, pp. 798–804.

[20] A. Gupta, P. Tripathi, B. P. Singh, and U. Singh, "Human face mask recognition using ResNet 152 model," in Proc. 2nd Int. Conf. Automation, Computing and Renewable Systems (ICACRS), 2023, pp. 1503–1510.

[21] D. Singh, N. Patel, and U. Singh, "Method for job recommendation based on machine learning and deep learning model," in Proc. 2nd Int. Conf. Automation, Computing and Renewable Systems (ICACRS), 2023, pp. 875–883.

[22] M. Vishwakarma, P. Tripathi, B. P. Singh, and U. Singh, "Segmentation and detection of brain tumors using EfficientNetB3 model," in Proc. 2nd Int. Conf. Automation, Computing and Renewable Systems (ICACRS), 2023, pp. 1463–1470.

[23] S. Ray and U. Singh, "Hybrid task cascade region-based convolutional neural network is used to analyse cricket scores," in Proc. Int. Conf. Self Sustainable Artificial Intelligence Systems (ICSSAS), 2023, pp. 19–26.

[24] A. Rajput, P. Gupta, P. Ghodeshwar, S. Varma, K. K. Sharma, and U. Singh, "Study of cloud providers (Azure, Amazon, and Oracle) according to service availability and price," in Proc. 3rd Int. Conf. Pervasive Computing and Social Networking (ICPCSN), 2023, pp. 1177–1188.

[25] M. Patel and U. Singh, "Age and gender recognition using deep learning technique," in Proc. 3rd Int. Conf. Smart Data Intelligence (ICSMDI), 2023, pp. 238–245.

[26] U. Singh and L. S. Songare, "Analysis and detection of monkeypox using the GoogLeNet model," in Proc. Int. Conf. Automation, Computing and Renewable Systems (ICACRS), 2022, pp. 1000–1008.

[27] World Bank, What a Waste 2.0: A Global Snapshot of Solid Waste Management to 2050. Washington, DC, USA: World Bank, 2023.

[28] M. Yang and G. Thung, "TrashNet dataset," Stanford University, 2016. [Online]. Available: https://github.com/garythung/trashnet

[29] B. Madhavi, M. Mahanty, C.-C. Lin et al., "SwinConvNeXt: A fused deep learning architecture for real-time garbage image classification," Scientific Reports, vol. 15, Art. no. 7995, 2025, doi: 10.1038/s41598-025-91302-7.

[30] "Garbage detection and classification using a new deep learning-based machine vision system as a tool for sustainable waste recycling," Waste Management, 2023, doi: 10.1016/j.wasman.2023.04.006.

[31] "Enhancing trash classification in smart cities using federated deep learning," Scientific Reports, 2024, doi: 10.1038/s41598-024-62003-4.

[32] "A solid waste classification using reinforcement learning boosted by data augmentation and hyperparameter optimization," Scientific Reports, vol. 16, Art. no. 18132, 2026, doi: 10.1038/s41598-026-46347-7.

[33] "An automated waste classification system using deep learning techniques: Toward efficient waste recycling and environmental sustainability," Knowledge-Based Systems, 2025, doi: 10.1016/j.knosys.2025.113063.

[34] B. Bamne, N. Shrivastava, L. Parashar, and U. Singh, "Transfer learning-based object detection by using convolutional neural networks," in Proc. Int. Conf. Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2020, pp. 328–332, doi: 10.1109/ICESC48915.2020.9156060.

[35] A. Saxena, A. Vyas, L. Parashar, and U. Singh, "A glaucoma detection using convolutional neural network," in Proc. Int. Conf. Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2020, pp. 815–820, doi: 10.1109/ICESC48915.2020.9155930.

[36] M. Ranjan, A. Shukla, K. Soni, S. Varma, M. Kuliha, and U. Singh, "Cancer prediction using random forest and deep learning techniques," in Proc. IEEE 11th Int. Conf. Communication Systems and Network Technologies (CSNT), Indore, India, 2022, pp. 227–231, doi: 10.1109/CSNT54456.2022.9787608.

[37] R. Baghel, P. Pahadiya, and U. Singh, "Human face mask identification using deep learning with OpenCV techniques," in Proc. 7th Int. Conf. Communication and Electronics Systems (ICCES), Coimbatore, India, 2022, pp. 1051–1057, doi: 10.1109/ICCES54183.2022.9835884.

[38] U. Singh, P. Gupta, and M. Shukla, "Activity detection and counting people using Mask-RCNN with bidirectional ConvLSTM," Journal of Intelligent and Fuzzy Systems, 2022, pp. 6505–6520.

[39] M. Tan and Q. Le, "EfficientNetV2: Smaller models and faster training," in Proc. 38th Int. Conf. Machine Learning (ICML), PMLR vol. 139, 2021, pp. 10096–10106.

[40] P. F. Proença and P. Simões, "TACO: Trash Annotations in Context for litter detection," arXiv:2003.06975, 2020.