ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

AN ATTENTION-BASED HYBRID DEEP LEARNING APPROACH FOR SOLID WASTE CLASSIFICATION

Shruti Handa

School of Engineering and Technology, CT University, Ludhiana, Punjab, India

Mandeep Kaur

School of Engineering and Technology, CT University, Ludhiana, Punjab, India

ABSTRACT

With the expansion of urban and economic landscapes, the volume of solid waste generated globally surges, posing significant environmental and public health challenges. Sustainable waste segregation is essential for proper disposal, promoting recycling, and reducing landfill accumulation, thereby supporting ecological balance. Existing studies leverage deep learning for solid waste classification, but mostly datasets consist of single-object images on plain backgrounds, which limits real-world applicability. To address this gap, a diverse dataset of 22,000 images spanning 12 waste categories is compiled from multiple public sources. Six state-of-the-art pre-trained convolutional neural networks—DenseNet201, ResNet101, EfficientNetB7, ConvNeXtBase, MobileNetV2, and InceptionV3—are fine-tuned using transfer learning. Among these, ConvNeXtBase achieves the highest individual test accuracy of 98.13%. To further improve performance, a hybrid model combining DenseNet201 and ConvNeXtBase is developed using an attention-based fusion mechanism. This model achieves a test accuracy of 98.45%, outperforming all single models. The results demonstrate the effectiveness of attention-driven ensemble learning in complex waste classification tasks. Future research emphasizes real-time deployment, adaptability across diverse waste streams, and integration with edge devices while promoting sustainable waste management practices. To further enhance accuracy, the study suggests expanding datasets, optimizing attention mechanisms, and experimenting with architectures such as Vision Transformers.

Keywords: deep learning, hybrid model, transfer learning, ensemble learning, solid waste, sustainable waste management

1. INTRODUCTION

Solid waste refers to the wide range of discarded materials generated as a result of various human activities in everyday life. This includes municipal household waste, medical and biomedical waste, construction and demolition debris, as well as industrial by-products. Amid growing human development, global solid waste generation has escalated, creating significant environmental and health concerns. Municipal solid waste generation is anticipated to escalate significantly, growing from 2.1 billion tonnes in 2023 to an estimated 3.8 billion tonnes by 2050, reflecting the mounting pressure of global consumption and urbanization [1]. A study by The Energy and Resources Institute (TERI) reveals that India produces approximately 62 million tons (MT) of waste annually. Of this, only 43 MT are collected, 12 MT are processed before disposal, while the remaining 31 MT end up in landfills. Projections by the Central Pollution Control Board (CPCB) suggest that the country's yearly waste generation could escalate to 165 million tons by 2030, underscoring a pressing need for improved waste management systems [2]. Despite the massive quantities of waste produced daily, only a small fraction is effectively recycled or processed in an environmentally responsible manner. A large portion of solid waste ends up in landfills or is improperly disposed of in open spaces and water bodies. This mismanagement leads to serious environmental issues, including soil contamination, air and water pollution, and the proliferation of disease-causing pathogens. Leachate from landfills contaminates groundwater, while the incineration of waste without proper controls releases harmful gases, contributing to air quality degradation and climate change.

The consequences of inadequate solid waste management extend beyond environmental damage. It directly impacts public health. Exposure to improperly disposed waste can cause respiratory ailments, vector-borne diseases, skin infections, and long-term health complications, particularly among vulnerable populations such as children, the elderly, and waste workers [3].

Given these far-reaching implications, effective solid waste management is critical to achieving urban sustainability and protecting ecological systems. It plays a pivotal role in conserving natural resources through recycling and reuse, reducing environmental pollution, and safeguarding human health. Therefore, developing a scalable, efficient, and intelligent waste classification system is essential to support responsible waste management strategies and promote a cleaner, healthier, and more sustainable urban environment [4].

Conventional methods for waste classification often rely on manual sorting, which is not only slow and resource-intensive but also prone to human error. This inefficiency hampers effective recycling efforts and contributes to the growing burden on landfills. Advancements in Artificial Intelligence (AI) have emerged as powerful tools for revolutionizing the waste management industry. AI technologies are increasingly being leveraged to optimize waste collection routes, predict waste generation trends, detect bin fill levels, and forecast waste characteristics with greater accuracy [5]. In this context, deep learning has also emerged as a promising approach to automate waste classification by leveraging its ability to learn complex patterns from visual data.

Recent advances in deep Convolutional Neural Networks (CNNs) have demonstrated impressive performance across numerous image classification tasks, including object detection, industrial quality control, and medical imaging. These successes have inspired a growing body of research focused on applying CNNs to solid waste classification. However, a critical limitation persists in most existing approaches: they are trained and evaluated on simplistic datasets composed primarily of single objects captured against uniform, white backgrounds [6]. Such datasets fail to capture the complexities of real-world waste disposal environments, where waste items often appear in varied orientations, under different lighting conditions, with cluttered or textured backgrounds, and frequently alongside other waste types.

To address this gap, we propose a robust, real-world-oriented solution for solid waste classification. Our research introduces two key contributions. First, we present a curated dataset comprising 22,000 images across 12 diverse waste categories: glass, e-waste, clothes, shoes, plastic, cardboard, paper, construction waste, medical waste, metal, organic waste, and trash. Unlike conventional datasets, this dataset includes both single and multiple objects per image, captured in varied indoor and outdoor environments, thereby significantly enhancing the dataset's representativeness and realism.

Second, we developed a hybrid deep learning model that leverages the complementary strengths of two high-performing CNN models: DenseNet201 and ConvNextBase. By fusing the feature representations of these pretrained networks, our hybrid model captures diverse spatial hierarchies and contextual information that individual models may overlook. This

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

ensemble-like fusion approach, combined with transfer learning and data augmentation strategies, substantially improved classification accuracy and generalizability under real-world conditions.

Through extensive experimentation, we have demonstrated that our hybrid model consistently outperforms individual CNN architectures, achieving high accuracy and robustness even in challenging scenarios involving lighting variations and object clutter. By integrating a rich dataset with a hybrid classification framework, our research aims to move closer to deployable, real-time waste classification systems supporting automated sorting in recycling plants and municipal waste management systems.

This research paper is organized as follows: Section 2 reviews the existing literature related to solid waste classification using deep learning techniques. Section 3 provides a detailed description of the curated dataset, the pretrained models employed, and the architecture of the proposed attention-based hybrid model. Section 4 presents a comprehensive evaluation of the proposed approach using various performance metrics to validate the effectiveness. Finally, Section 5 summarizes the findings and suggests potential paths for future research.

2. LITERATURE REVIEW

This literature review highlights the emerging importance of deep learning, namely CNNs, as a powerful tool for accurate and automated solid waste classification. The review demonstrates that researchers have used transfer learning and hybrid model architectures to improve classification performance on solid waste datasets. The review also covers a broad spectrum of state-of-the-art techniques ranging from standalone CNN-based models to sophisticated ensemble and hybrid models, revealing significant advances in the field while also identifying ongoing challenges and research gaps.

The study [7] utilizes CNN, specifically a pretrained ResNet-50 model with transfer learning, to categorize waste into six types. The approach handles images with multiple objects by employing a sliding-window technique for initial segmentation and Gaussian clustering to pinpoint the location of classified items. While the model achieved a high classification accuracy of 92.4% on individual objects, the overall detection rate in simulation for mixed waste piles was 48.4%, highlighting challenges in object detection and segmentation under varying conditions.

The authors [8] evaluated various deep CNN models for image-based waste segregation, combining existing datasets with newly collected images. The study utilized a dataset of 3102 waste images organized into four categories: paper, plastic, metal, and glass. Through transfer learning and fine-tuning, the researchers found that the ResNet18 architecture performed best, achieving a validation accuracy of 87.8% in classifying waste materials. Plastic had a lower correct prediction rate (82.8%) compared to other waste types because of frequent changes in shape, size, and color.

The authors [9] proposed utilizing transfer learning, leveraging three distinct pretrained CNN models (VGG19, DenseNet169, and NASNetLarge), to improve classification accuracy. By constructing a candidate classifier from each pretrained model and selecting the optimal output, their approach enhances the overall classification accuracy across different waste subcategories. The proposed model demonstrates superior performance compared to existing methods on two distinct waste image datasets, achieving an accuracy of 96.5% and 94%, respectively. Although the dataset consists of 5000 waste images divided into six main categories (glass, paper, plastic, metal, textile, and organic waste) and 12 subcategories,

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

accurate classification and identification of these is difficult because the shape of objects changes dramatically as they are discarded. To learn feature parameters, a dataset with a clear classification is necessary.

The study [10] compares the performance of four pre-trained CNN models (ResNet50, DenseNet169, VGG16, and AlexNet) trained on an augmented dataset of six waste categories: cardboard, glass, metal, paper, plastic, and trash, with about 4,163 total images. The experimental results indicate that DenseNet169 performed the best overall, with an accuracy of 94.9%, while ResNet50 achieved a similar performance, with 93.4% accuracy. Although DenseNet169 was the best-performing model, trash and glass were found to be the most frequently misclassified images. As a result, additional clear images of these categories are required.

The authors [11] presented a systematic approach to improve the accuracy of image-based waste classification models using transfer learning and data augmentation techniques. It focuses on building upon the IBM WasteNet project, which aims to enhance recycling by employing AI for waste sorting. The research provides detailed insights into model training decisions and demonstrates significant improvements in classification accuracy, with test accuracy reaching 95.40%. However, deep architectures, such as EfficientNets, can be explored to implement real-time classification on embedded devices.

The authors [12] introduced a framework employing the EfficientNet-B0 model, fine-tuned with region-specific image datasets, to improve the accuracy and efficiency of municipal solid waste categorization. They compare their approach to existing techniques and highlight the reduced computational resources required by their method while achieving comparable accuracy to more complex models, ultimately aiming to automate and optimize waste management processes. The proposed model achieved an accuracy of 85%. This study uses only EfficientNet B0 to classify solid waste. However, EfficientNet-B0 to B7 can be used to explore transfer learning for even greater accuracy.

The study [13] explores the effectiveness of an ensemble approach that combines multiple pretrained CNN models—InceptionResNetV2, EfficientNetB3, and DenseNet201—for classifying landfill waste into nine categories. The ensemble model outperformed individual CNNs, achieving higher prediction precision (90% compared to 86–88%) while maintaining comparable or lower computational costs. Despite its strong overall performance, the model faced challenges in accurately distinguishing between waste classes with visually similar features.

The authors [14] highlight the inefficiencies of traditional manual methods and propose a smart system, Learning Approach with a Deep Neural Network for Smart systems (LADS), which leverages deep neural networks, specifically a CNN, for automated waste classification. The LADS model demonstrated superior accuracy (94.53%) in categorizing waste into organic and recyclable compared to existing pre-trained models like AlexNet, VGG16, and ResNet34. The study [15] addresses the critical global issue of poor waste segregation by proposing an improved Deep Convolutional Neural Network (DCNN) for automated classification of organic and recyclable waste. The improved DCNN model, utilizing Leaky-ReLU as an activation function and dropout for regularization, achieved a classification accuracy of 93.28%. This performance surpasses that of several other popular deep learning models, including VGG16, VGG19, MobileNetV2, DenseNet121, and EfficientNetB0, highlighting its effectiveness for binary waste classification. The authors [16] proposed a hybrid CNN-LSTM model with transfer learning for smart waste classification, aimed at promoting sustainable development. Waste was categorized into recyclable and organic classes by combining spatial feature extraction using CNN with temporal sequence learning through LSTM. Transfer learning with ImageNet was employed to enhance classification performance, and an improved data augmentation technique was used to address overfitting and data imbalance. The model was evaluated using a sample from the TrashNet dataset consisting of images labeled as organic and recyclable. The hybrid model, optimized with the Adaptive Moment Estimator (AME), outperformed existing CNN models (VGG-16, ResNet-34, ResNet-50, and AlexNet), achieving the highest precision of 95.45% and the lowest training, validation, and testing losses. While these studies demonstrate notable advancements in solid waste classification, they are limited to only two waste categories, which constrains their applicability in real-world scenarios where mixed waste streams are more complex and diverse.

The research [17] conducts a comparative analysis of different deep learning models, specifically CNNs.The ResNeXt-101 model consistently outperformed others, achieving the highest test accuracy of 89.62% using the TrashBox dataset. The study proposes deploying a federated framework for visual detection of 7 trash classes at waste management facilities, using a combination of four models: ResNeXt-101, ShuffleNetV2, ResNet-34, and MobileNetV3-Large, to improve the waste classification accuracy in varied environmental settings.

The review indicates that CNN-based deep learning models, particularly those employing transfer learning, hybrid architectures, and ensemble techniques, have significantly advanced the field of automated solid waste classification. While such models demonstrate promising classification performance, challenges remain in handling visually similar waste types and ensuring generalizability across diverse real-world environments. Future research should focus on enhancing model accuracy by utilizing larger and more varied waste datasets.

3. METHODOLOGY

This section outlines the curated dataset and preprocessing strategies, details the pre-trained models utilized, and describes the architecture of the proposed attention-based hybrid model.

3.1. Dataset Description

To develop and evaluate the proposed solid waste classification model, a comprehensive dataset was curated by aggregating high-quality waste images from multiple sources. Data collection included publicly available datasets from platforms such as Kaggle, Mendeley Data, UC Irvine Machine Learning Repository, and Roboflow to ensure greater variability. **Table 1** lists the public datasets, their respective sources, and the corresponding waste categories from which the dataset for this research was curated.

Table 1. Summary of Public Datasets Utilized

Dataset	Waste Categories
Trashnet [18]	Cardboard, Metal, Trash, Paper, Glass, Plastic
RealWaste [19]	Cardboard, Food organics, Glass, Metal, Miscellaneous trash, Paper, Plastic, Textile trash, Vegetation
Construction and demolition waste object detection dataset [20]	Bricks, Concrete, Tiles, Wood, Pipes, Plastics, General waste, Foaming insulation, Stones, Plaster boards
E-Waste Dataset [21]	Battery, Keyboard, Microwave, Mouse, Television, Printer, Washing Machine

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

Garbage Classification [22]	Paper, Cardboard, Biological, Metal, Plastic, Green-glass, Brown-glass, White-glass, Clothes, Shoes, Batteries, Trash
Clase Congestion Classification [23]	Wood
Metal 4 Classification [24]	Corroded Metal

The dataset comprises 22,000 images categorized into 12 types, representing major forms of solid waste: glass, e-waste, clothes, shoes, plastic, cardboard, paper, construction waste, medical waste, metal, organic waste, and trash. **Table 2** presents the distribution of images per category, highlighting the scale and structure of the dataset used in this study.

Table2. Waste image classes and their counts

Waste Classes	**Image Count**
Construction	1480
Glass	3039
Clothes	2645
Trash	930
Ewaste	3000
Metal	1045
Organic	996
Medical	1380
Paper	1668
Shoes	2009
Cardboard	1825
Plastic	1983

Unlike many existing datasets that primarily consist of single-object images on plain white backgrounds, the curated dataset used in this study incorporates a much broader range of image types. It includes not only single-object images with clean, white backgrounds but also real-world waste images captured in actual landfill environments. Additionally, the dataset features images taken in diverse settings with varied lighting conditions and complex backgrounds, often containing multiple waste objects. **Figure 1** illustrates sample images from each waste category, providing an overview of the visual diversity within the dataset.

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

Figure 1. Sample images of waste categories from the dataset

This diverse and extensive dataset forms the foundation for training and testing the hybrid deep learning model, enabling more accurate classification of waste in realistic, unstructured environments.

3.2. Data Preprocessing

To prepare the dataset for training the deep learning models, a comprehensive data preprocessing pipeline was implemented. The dataset, consisting of 22,000 images across 12 solid waste categories, was initially organized in a dataframe linking each image file path with its corresponding label. A method called stratified sampling was used to divide the dataset into training and test sets in an 80:20 ratio, ensuring that each category was fairly represented in both groups. The training set was further divided into training and validation subsets using a stratified 70:10 split, maintaining balanced class distributions across all three sets.

To enhance the model's ability to generalize and perform well in varied real-world scenarios, data augmentation techniques were applied to the training set. The augmentation strategy included random rotations up to 60 degrees, horizontal and vertical shifts of up to 15%, zoom variations of ±20%, horizontal and vertical flips, slight shearing, brightness adjustments ranging from 90% to 110%, and random channel shifts to simulate changes in lighting. Any pixels introduced during these transformations were filled using the nearest neighbor method. Additionally, all images were normalized using a model-specific preprocessing function to ensure consistency with the pretrained architectures used.

Images were loaded in batches, which facilitated efficient real-time data augmentation and ensured compatibility with large datasets stored on disk. All images were resized to 224×224

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

pixels, and a batch size of 32 was used throughout. Labels were one-hot encoded using the categorical mode, and shuffling was disabled to maintain a consistent order across evaluations. While the training data was augmented, the validation and test sets were only normalized, without any augmentations, to ensure fair and unbiased evaluation. This preprocessing strategy contributed significantly to the accuracy of the hybrid deep learning models in classifying diverse types of solid waste under varying conditions.

3.3. Pretrained Models

In this study, six state-of-the-art deep convolutional neural network architectures were employed to develop an accurate and generalizable solid waste classification system: DenseNet201, ResNet101, EfficientNetB7, MobileNetV2, InceptionV3, and ConvNeXtBase. These architectures were chosen due to their proven effectiveness in various computer vision tasks and their complementary design characteristics, which were later utilized to construct a hybrid model.

DenseNet201 is designed with a densely connected architecture in which every layer is directly linked to all earlier layers, allowing it to reuse features and enhance information flow throughout the network. This helps to relieve the vanishing gradient problem, stimulate feature reuse, improve feature propagation, and significantly reduce the number of parameters [25]. ResNet101 uses deep residual learning with identity shortcut connections, allowing very deep networks to be trained effectively by mitigating the degradation problem [26]. EfficientNetB7 utilizes a compound scaling strategy that proportionally increases the input resolution, width, and depth. This balanced approach enables the model to deliver high accuracy while maintaining computational efficiency [27].

MobileNetV2 is designed for mobile and resource-constrained environments. It uses inverted residual blocks and depthwise separable convolutions, offering a good trade-off between accuracy and computational cost [28]. Compared to its predecessor, MobileNetV2 is more efficient and uses far fewer parameters. InceptionV3 utilizes factorized convolutions, auxiliary classifiers, and label smoothing, making it efficient in terms of computation and effective for diverse visual tasks [29]. ConvNeXtBase, a more current architecture, modernizes convolutional networks with design decisions inspired by vision transformers, enhancing its scalability and performance on huge datasets. It focuses on leveraging depth-wise convolution and the ResNext family of convolutional neural networks to classify images efficiently [30].

All models were initialized with ImageNet-pretrained weights and then tailored for waste classification by replacing the top layers with a new classifier. To ensure uniformity, all models used a fixed input size of 224 × 224 × 3. To balance generalization with task-specific adaptation, only the top 20% of each model's layers were made trainable, with the remaining 80% frozen. This method allowed the models to keep core visual elements while adjusting higher-level features for the solid waste dataset. Each model consisted of a Global Average Pooling (GAP) layer, a dropout layer for regularization, and a dense output layer with 12 units corresponding to the 12 waste categories and softmax activation. To mitigate overfitting, L2 regularization was incorporated into the dense layer.

The models were built and optimized using the Adam optimizer, configured with a learning rate of 0.0001. The model was trained using the categorical cross-entropy loss function, an appropriate choice for addressing classification tasks involving multiple categories. To address the class imbalance in the dataset, class weights were computed and applied based on the training label distribution. All models were trained over a maximum of 30 epochs, with

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

early stopping and ReduceLROnPlateau callbacks to dynamically regulate the learning process and prevent overfitting. This rigorous and consistent training framework allowed for a reliable comparison across models and provided a strong foundation for developing a hybrid model for solid waste classification.

3.4. Proposed Hybrid Model

A hybrid deep learning architecture was created by integrating two high-performing convolutional neural networks, DenseNet201 and ConvNeXtBase, to improve the solid waste classification accuracy and generalization. These two models were chosen for their higher individual performance on the curated waste dataset, surpassing the other pretrained models used in this work, such as ResNet101, EfficientNetB7, MobileNetV2, and InceptionV3. DenseNet201 was chosen for its densely connected architecture, which encourages feature reuse and mitigates the vanishing gradient problem, allowing deeper networks to be trained effectively. Its ability to preserve feature information across layers proved valuable in identifying subtle patterns in waste images. ConvNeXtBase, a modernized convolutional network inspired by Vision Transformers, was selected for its scalability and high representational strength, particularly in diverse and complex image environments. **Figure 2** presents a schematic overview of the proposed hybrid model, visually outlining the sequential process followed in the methodology.

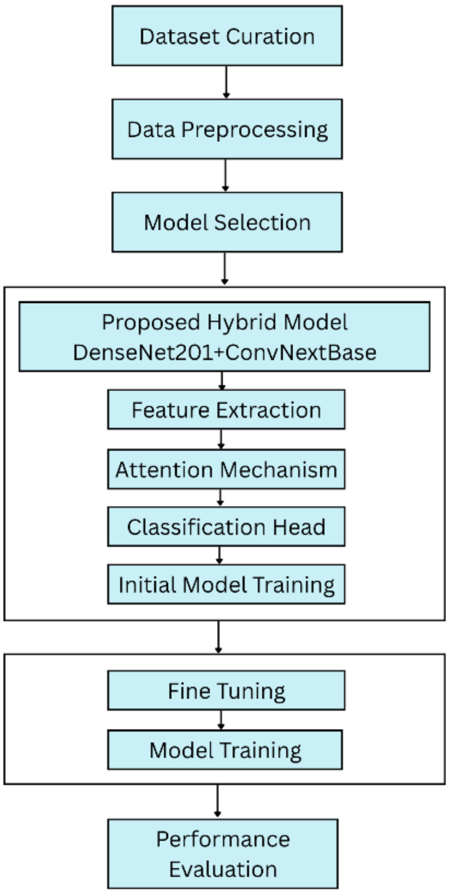

Figure 2. Workflow diagram of the proposed Hybrid Model

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

The proposed hybrid model integrates DenseNet201 and ConvNeXtBase using an attention-based fusion mechanism. Both models were initialized with pretrained ImageNet weights, and their convolutional backbones served as fixed feature extractors by freezing all their layers during initial training. Features from both networks were extracted using GAP to reduce spatial dimensions while preserving important features. The GAP outputs from both models were concatenated and passed through a dense attention layer that learned to assign importance weights to each feature stream. The attention mechanism used a softmax activation to generate a pair of weights, which were then applied to the DenseNet and ConvNeXt feature vectors, respectively. The weighted features were concatenated again to form a fused representation that combines complementary information from both models.

The fused feature vector was passed through a classification head comprising a dropout layer to prevent overfitting, a dense layer with 512 units and ReLU activation to introduce non-linearity, a second dropout layer for further regularization, and a final dense layer with softmax activation for classifying images into 12 waste categories. The model was trained in two distinct phases to maximize stability and adaptability. The first phase was the initial training or feature extraction phase, where all layers except the final classification layers were frozen. Only the last 10 layers (mostly fully connected and dropout layers) were made trainable. The model was compiled with the Adam optimizer (learning rate = 1e-3) and trained for 20 epochs with categorical cross-entropy loss. Early stopping and ReduceLROnPlateau callbacks were used to monitor validation loss and dynamically adjust learning rates. The second phase was fine-tuning, where all layers of the model were unfrozen to allow end-to-end learning. A lower learning rate (1e-5) was used for stable gradient updates. Training was conducted for an additional 30 epochs with the same early stopping and learning rate reduction strategies.

This two-phase training strategy ensured that the model first stabilized on general visual features before fine-tuning all layers to the specific characteristics of the waste classification task.

4. RESULTS AND DISCUSSION

This section presents the performance evaluation of the proposed models against individual models, followed by an analysis of inference time and a comparison with state-of-the-art models.

4.1. Performance Analysis of Proposed Model

The accuracy results from different models show how well various deep learning architectures perform and how well they can apply what they have learned to new data for solid waste classification. **Table 3** compares the training, validation, and testing accuracies of DenseNet201, ResNet101, EfficientNetB7, ConvNextBase, MobileNetV2, InceptionV2, and the proposed hybrid model.

Table 3. Accuracy Comparison of Individual and Hybrid Models

Model	Training Accuracy (%)	Validation Accuracy (%)	Test Accuracy (%)
DenseNet201	99.91	97.05	96.45
ResNet101	99.83	96.32	95.99

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

EfficientNetB7	99.56	96.27	96.29
ConvNeXtBase	99.94	98.64	98.13
MobileNetV2	99.14	93.91	94.02
InceptionV3	99.75	95.18	94.97
Proposed Hybrid	99.62	98.32	**98.45**

Among individual models, ConvNeXtBase achieved the highest test accuracy (98.13%), closely followed by DenseNet201 (96.45%) and EfficientNetB7 (96.29%), indicating strong generalization from training to unseen data. While MobileNetV2 and InceptionV3 showed slightly lower test accuracies (94.02% and 94.97%, respectively), they still performed reasonably well considering their computational efficiency. Notably, the proposed hybrid model outperformed all individual models, achieving a test accuracy of 98.45%, validating the effectiveness of ensemble learning in capturing diverse feature representations and reducing generalization error. This suggests that hybrid architectures can significantly enhance classification robustness in complex, multi-class tasks, such as waste categorization.

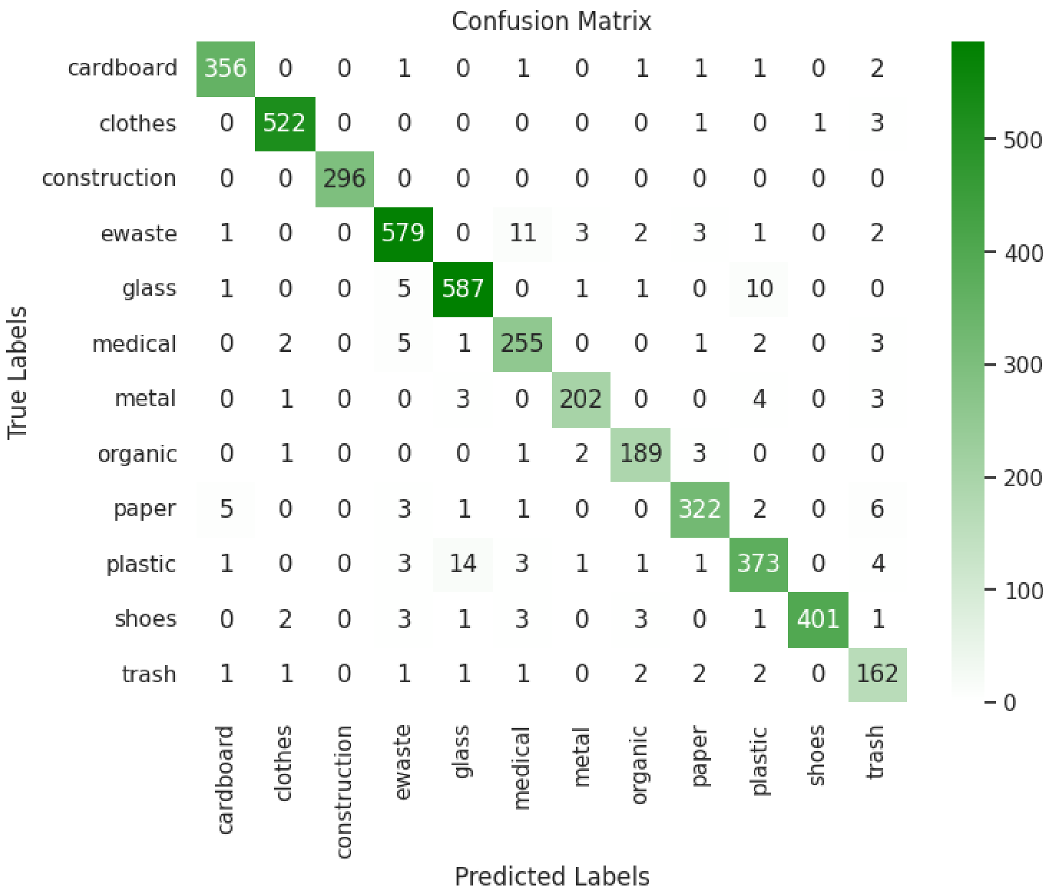

Figure 3. Confusion matrix for DenseNet201 model

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

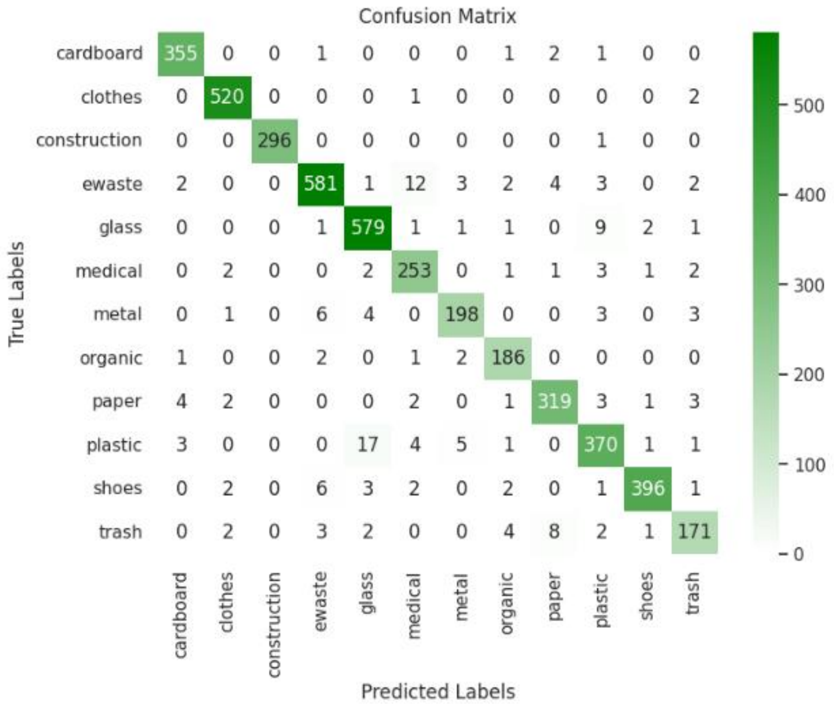

Figure 4. Confusion matrix for ResNet101 model

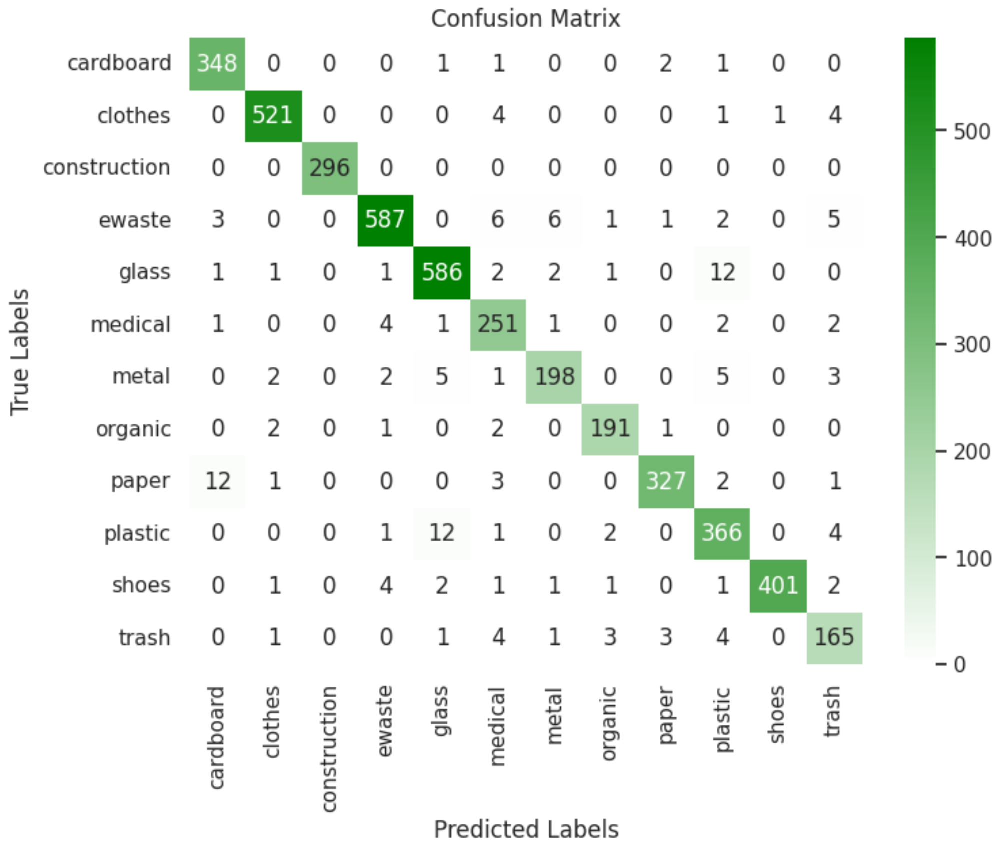

Figure 5. Confusion matrix for EfficientNetB7 model

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

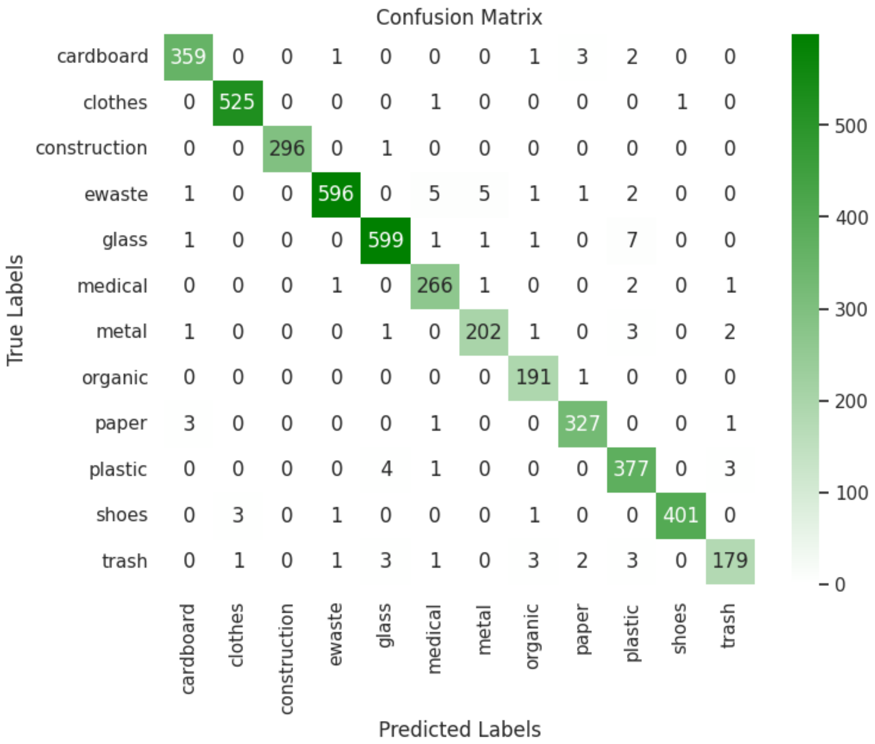

Figure 6. Confusion matrix for ConvNextBase model

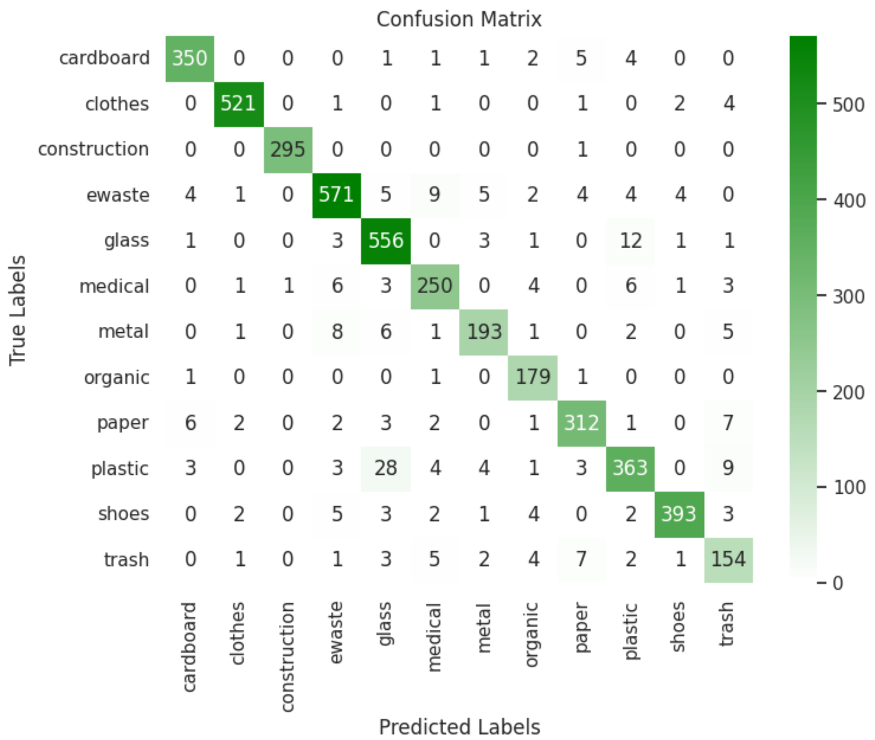

Figure 7. Confusion matrix for MobileNetV2 model

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

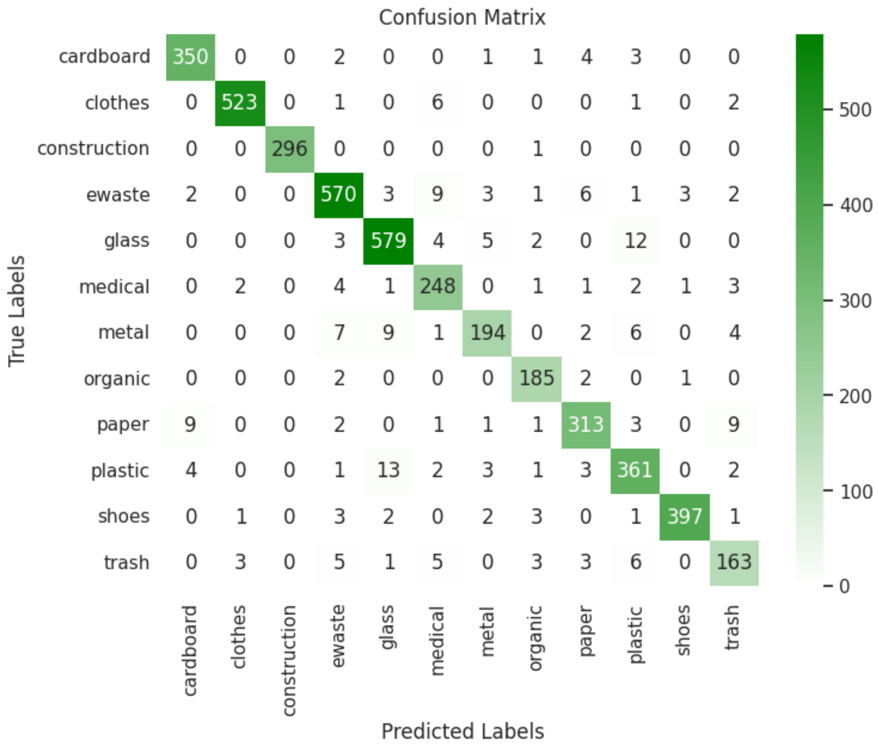

Figure 8. Confusion matrix for InceptionV3 model

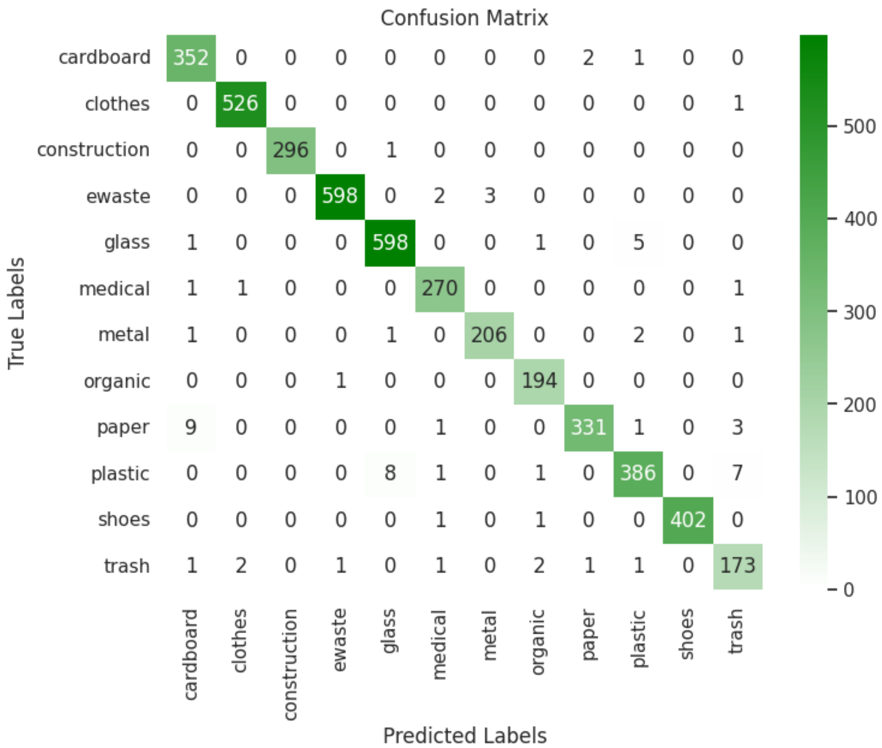

Figure 9. Confusion matrix for the Proposed Hybrid model

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

Figures 3-9 represent confusion matrices for all the individual and hybrid models evaluated on the test dataset. DenseNet201 model achieves exceptionally high correct predictions in major categories, including construction, shoes, clothes, organic, paper, and cardboard, showcasing the model's strong feature extraction capability and robustness in handling visually distinct materials. This highlights DenseNet201's ability to learn discriminative patterns, especially for categories with unique textures and structural features. The model also performs reliably in more challenging categories. The confusion matrix also reveals that the plastic category has 14 instances misclassified as glass, while the glass category has 10 instances misclassified as plastic, indicating significant visual overlap between these materials. A few minor misclassifications appear in trash, which is inherently difficult due to its heterogeneous composition and overlapping visual traits with other classes, such as organic, paper, or plastic. Nonetheless, even in these complex classes, misclassification rates remain low.

ResNet101 model shows high precision in categories such as construction, clothes, cardboard, medical, organic and paper, reflecting the model's effectiveness in learning deep spatial features for visually distinctive classes. The diagonal dominance in these rows indicates a reliable prediction rate. Notably, the ResNet101 confusion matrix highlights that 17 plastic instances were misclassified as glass and 9 glass instances as plastic, indicating a strong visual resemblance that challenges accurate classification between these categories. However, minor confusion is visible in more ambiguous categories. For example, plastic is sometimes misclassified as glass or metal (17 and 5 instances, respectively), likely due to shared visual textures such as shine or transparency. Similarly, inherently heterogeneous trash category experience limited but noticeable confusion, with a few samples mispredicted as glass, paper, or e-waste. The model also struggles slightly with metal, occasionally confusing them with e-waste.

The confusion matrix for EfficientNetB7 shows high accuracy in classes such as construction, clothes, medical, metal, shoes, and e-waste. The results indicate that the model effectively captures spatial hierarchies and subtle visual features, benefiting from its compound scaling architecture. Misclassifications are relatively few but occur primarily in visually overlapping or heterogeneous classes. For instance, plastic waste is occasionally misidentified as glass or paper (12 and 4 instances, respectively), likely due to visual similarities such as color or transparency. Similarly, paper is misclassified as cardboard in 12 instances, likely due to textural overlap. E-waste, while accurately classified in most cases, shows minor confusion with medical, metal, and trash, reflecting shared metallic and plastic components in these materials.

ConvNeXtBase model shows good classification accuracy across all 12 solid waste types, demonstrating the model's ability to extract and generalize features. Notably, categories such as construction, clothes, medical, metal and organic exhibit nearly perfect classification, showing the model's ability to distinguish clear, class-specific patterns in the data. Minimal confusion is observed in visually overlapping categories. For instance, glass is occasionally misclassified as plastic (7 instances), likely due to transparency or reflective surface similarities. E-waste shows minor confusion with metal and medical, reflecting shared textures or components in such materials. Trash is one of the most diverse categories and sees a few misclassifications as glass, organic, paper and plastic, highlighting the inherent ambiguity in mixed or poorly segregated waste. The model also maintains high precision in categories with subtle visual differences, such as metal, medical, and organic, where the

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

correct classifications are dominant and misclassifications are limited to 1–3 instances across a few neighboring classes.

The confusion matrix for the MobileNetV2 model reveals reasonably good classification performance across most waste categories but indicates relatively more confusion compared to heavier models like ConvNeXtBase or EfficientNetB7. MobileNetV2's lightweight architecture offers speed and efficiency, but at the cost of slightly reduced precision, especially in visually similar categories. Most classes, such as construction, organic clothes, and cardboard, are well classified with minimal errors. However, e-waste and glass show increased confusion. Both glass and plastic are highly confused with 28 instances of plastic misclassified as glass, showing the model's difficulty identifying transparent or reflective materials.

The InceptionV3 model indicates that it effectively distinguishes several waste categories with high accuracy, particularly construction, clothes, cardboard, and organic. Shoes and medical also show reliable predictions. However, some confusion is evident between visually similar categories. Notably, 13 plastic samples are misclassified as glass, and 12 glass samples are confused with paper, likely due to overlapping visual characteristics such as transparency and reflective surfaces. E-waste is occasionally misclassified as medical waste, which can be attributed to the plastic material used in both. Metal is also confused with e-waste and glass, possibly due to similar shiny textures. Overall, InceptionV3 performs well, with minor misclassifications primarily between categories that share visual features, suggesting potential improvements through targeted data augmentation or enhanced feature discrimination techniques.

The confusion matrix for the new hybrid model that combines DenseNet201 and ConvNeXtBase shows great classification results for almost all types of waste. The model accurately classifies a high number of instances for classes such as cardboard, clothes, ewaste, shoes, construction, paper, medical, metal, organic, plastic, and glass, indicating strong discriminative ability for these types. For instance, categories such as clothes and organic show excellent performance with only 1 instance of clothes misclassified as trash and 1 instance of organic misclassified as ewaste. Misclassifications are generally low and scattered, often involving very few instances, which suggests that the model generalizes well without significant class confusion. Overall, the hybrid approach has led to improved robustness and precision, especially in differentiating between visually similar waste types.

The comparative analysis of the confusion matrices across various deep learning models highlights the strengths and limitations of each in classifying 12 categories of solid waste. Among the models, the hybrid model combining DenseNet201 and ConvNeXtBase demonstrates the most robust and balanced performance, achieving high accuracy in almost all categories with minimal confusion, particularly excelling in complex classes like trash, organic, and medical. DenseNet201 also performs exceptionally well, especially in categories with distinct textures such as clothes and construction, due to its dense connectivity and strong feature extraction capabilities. ConvNeXtBase, despite being slightly lighter, achieves near-perfect classification in categories such as glass, plastic, construction, cardboard, and clothes due to its efficient feature generalization and modern architecture.

EfficientNetB7 leverages its compound scaling strategy to deliver reasonable classification performance, particularly in visually complex categories like plastic and glass, due to their similar textures and overlapping visual features. ResNet101 exhibits reliable performance across most categories but suffers slightly more from confusion in classes such as plastic,

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

metal, and trash, where reflective or heterogeneous features pose a challenge to the model. InceptionV3 maintains solid accuracy in many categories but displays noticeable confusion between glass, plastic, and metal, most likely due to common visual characteristics such as glossiness and transparency. Finally, MobileNetV2, while efficient and lightweight, exhibits more misclassifications in visually comparable classes, indicating a restricted representational capability when compared to deeper architectures. Overall, the proposed hybrid model emerges as the most effective, combining the strengths of both architectures to deliver high accuracy and minimal misclassification across both distinct and ambiguous waste categories.

Table 4. Performance Metrics of DenseNet201 Model

Class	Precision	Recall	F1-Score
Cardboard	0.98	0.98	0.98
Clothes	0.99	0.99	0.99
Construction	1.00	1.00	1.00
Ewaste	0.96	0.96	0.96
Glass	0.97	0.97	0.97
Medical	0.95	0.92	0.94
Metal	0.95	0.97	0.96
Organic	0.96	0.95	0.96
Paper	0.95	0.96	0.96
Plastic	0.93	0.94	0.94
Shoes	0.97	1.00	0.98
Trash	0.94	0.87	0.90

Table 5. Performance Metrics of ResNet101 Model

Class	Precision	Recall	F1-Score
Cardboard	0.99	0.97	0.98
Clothes	0.99	0.98	0.99
Construction	1.00	1.00	1.00
Ewaste	0.95	0.97	0.96
Glass	0.97	0.95	0.96
Medical	0.95	0.92	0.94
Metal	0.92	0.95	0.93
Organic	0.97	0.93	0.95
Paper	0.95	0.96	0.95
Plastic	0.92	0.93	0.93
Shoes	0.96	0.99	0.97
Trash	0.89	0.92	0.90

Table 6. Performance Metrics of EfficientNetB7 Model

Class	Precision	Recall	F1-Score
Cardboard	0.99	0.95	0.97
Clothes	0.98	0.98	0.98
Construction	1.00	1.00	1.00
Ewaste	0.96	0.98	0.97
Glass	0.97	0.96	0.97

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

Medical	0.96	0.91	0.93
Metal	0.92	0.95	0.93
Organic	0.97	0.96	0.96
Paper	0.95	0.98	0.96
Plastic	0.95	0.92	0.94
Shoes	0.97	1.00	0.98
Trash	0.91	0.89	0.90

Table 7. Performance Metrics of ConvNextBase Model

Class	Precision	Recall	F1-Score
Cardboard	0.98	0.98	0.98
Clothes	1.00	0.99	0.99
Construction	1.00	1.00	1.00
Ewaste	0.98	0.99	0.98
Glass	0.98	0.99	0.98
Medical	0.98	0.96	0.97
Metal	0.96	0.97	0.96
Organic	0.99	0.96	0.98
Paper	0.98	0.98	0.98
Plastic	0.98	0.95	0.97
Shoes	0.99	1.00	0.99
Trash	0.93	0.96	0.94

Table 8. Performance Metrics of MobileNetV2 Model

Class	Precision	Recall	F1-Score
Cardboard	0.96	0.96	0.96
Clothes	0.98	0.98	0.98
Construction	1.00	1.00	1.00
Ewaste	0.94	0.95	0.94
Glass	0.96	0.91	0.94
Medical	0.91	0.91	0.91
Metal	0.89	0.92	0.91
Organic	0.98	0.90	0.94
Paper	0.93	0.93	0.93
Plastic	0.87	0.92	0.89
Shoes	0.95	0.98	0.96
Trash	0.86	0.83	0.84

Table 9. Performance Metrics of InceptionV3 Model

Class	Precision	Recall	F1-Score
Cardboard	0.97	0.96	0.96
Clothes	0.98	0.99	0.98
Construction	1.00	1.00	1.00
Ewaste	0.95	0.95	0.95
Glass	0.96	0.95	0.95
Medical	0.94	0.90	0.92

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

Metal	0.87	0.93	0.90
Organic	0.97	0.93	0.95
Paper	0.92	0.94	0.93
Plastic	0.93	0.91	0.92
Shoes	0.97	0.99	0.98
Trash	0.86	0.88	0.87

Table 10. Performance Metrics of Proposed Hybrid Model

Class	Precision	Recall	F1-Score
Cardboard	0.99	0.96	0.98
Clothes	1.00	0.99	1.00
Construction	1.00	1.00	1.00
Ewaste	0.99	1.00	0.99
Glass	0.99	0.98	0.99
Medical	0.99	0.98	0.98
Metal	0.98	0.99	0.98
Organic	0.99	0.97	0.98
Paper	0.96	0.99	0.97
Plastic	0.96	0.97	0.97
Shoes	1.00	1.00	1.00
Trash	0.95	0.93	0.94

Tables 4-10 present performance metrics for DenseNet201, ResNet101, EfficientNet121, ConvNextBase, MobileNetV2, InceptionV3, and the proposed Hybrid model, including precision, recall, and F1 Score. DenseNet201 model achieved a weighted average accuracy, precision, recall, and F1-score of 0.96 on the test dataset, as illustrated in Table 4. It performed exceptionally in categories such as construction (F1-score: 1.00), clothes (0.99), and shoes (0.98). High F1-scores were also observed for cardboard, glass, metal, and e-waste (0.94–0.98), reflecting strong feature discrimination. Plastic, paper, and organic waste were classified reliably, despite some visual overlap. Lower performance was seen in medical waste (0.94) and trash (F1: 0.90), with trash showing the lowest recall (0.87), likely due to its heterogeneous nature. Overall, the model shows robust classification with minor misclassification in ambiguous classes.

According to the metrics in Table 5, ResNet101 model performed exceptionally well in classifying construction (F1-score: 1.00), clothes (0.99), and cardboard (0.98), indicating effective learning of distinct features. High F1-scores were also observed for e-waste (0.96) and glass (0.96), despite minor confusion with similar classes. Metal, plastic, and medical categories showed slightly reduced performance, likely due to visual overlap with e-waste. Trash achieved an F1-score of 0.90, reflecting moderate difficulty due to its mixed composition. The EfficientNetB7 model achieved perfect scores (F1-score: 1.00) in classifying construction waste and performed very well for shoes (0.98), clothes (0.98), and e-waste (0.97), as reported in Table 6, reflecting its ability to capture complex and diverse features. Categories such as glass, paper, organic, and cardboard also showed high F1-scores, indicating reliable classification of visually distinctive materials. Slightly lower scores for medical, metal, plastic, and trash suggest occasional misclassifications due to visual similarities.

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

Table 7 shows the performance metrics of ConvNeXtBase model , which attained a weighted average precision, recall, and F1-score of 0.98 on the test dataset. It achieves perfect or near-perfect results in categories such as construction (1.00), clothes (0.99), shoes (0.99), and e-waste (0.98), showcasing its robust feature extraction capabilities. Glass, cardboard, and paper also show very high F1 scores (0.98), indicating effective handling of materials with reflective or textural patterns. Medical, metal, organic, and plastic categories yield slightly lower yet strong F1 scores. Trash, despite its visual diversity, is well classified, reflecting ConvNeXt's strength even in challenging classes.

MobileNetV2 model achieved weighted precision, recall, and F1-score of 0.94 on the test dataset , as shown in Table 8. The model performed exceptionally in the construction class with a perfect score of 1.00 and shows strong results in clothes (0.98), shoes (0.96), and cardboard (0.96). Moderate F1-scores are observed in e-waste, organic, glass, and paper, though recall is slightly lower in categories such as glass and organic. Performance dips further in medical, metal, and especially in plastic and trash, highlighting the model's challenges with visually overlapping or heterogeneous materials. As reported in Table 9, the InceptionV3 model achieved weighted precision, recall, and F1-score of 0.95 on the test dataset. It obtained a perfect classification for construction (1.00) and high F1-scores in clothes (0.98), shoes (0.98), and cardboard (0.96), showing effective feature learning for visually distinct categories. E-waste, glass, and organic waste also performed well with F1 scores of 0.95, though slight confusion is evident in overlapping textures. Medical and plastic categories show moderate performance, while metal and trash had the lowest F1-scores, reflecting difficulty in distinguishing complex or visually similar materials. Overall, the model generalizes well.

The hybrid model achieved a weighted precision, recall, and F1-score of 0.98 on the test dataset, as reported in Table 10, reflecting excellent overall performance. Construction and shoes were classified perfectly (F1-score: 1.00), while categories such as clothes, ewaste, glass, metal, and medical also achieved near-perfect F1-scores (0.98–1.00), indicating strong discriminative capability across diverse materials. While paper and plastic maintained an F1-score of 0.97, showing minimal misclassification. Overall, the hybrid approach demonstrates superior generalization and robust classification across all 12 waste categories.

The classification results demonstrate that all models perform well across the 12 solid waste categories, with varying degrees of accuracy. The Hybrid and ConvNextBase models deliver the best results, each achieving a weighted F1-score of 0.98. They show excellent generalization and robust performance in classes such as cardboard, clothes, construction, ewaste, and shoes, with the hybrid model achieving perfect F1 scores for clothes, construction, organic and shoes. Among the other models, DenseNet201, ResNet101, and EfficientNetB7 each reach an F1-score of 0.96, performing strongly in distinct classes such as construction, clothes, and shoes, but showing slight weaknesses in more ambiguous categories like trash and medical waste. InceptionV3 follows closely with an F1 score of 0.95, maintaining good balance but struggling with metal and trash. MobileNetV2, though efficient, shows the lowest F1-score (0.94), performing well in simpler categories but facing challenges in classifying plastic and trash, likely due to visual overlap. Overall, the hybrid model stands out as the most reliable choice for solid waste classification due to its superior accuracy and consistency across all waste categories.

4.2. Analysis of Inference Time

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

The models were trained and evaluated using Kaggle's dual NVIDIA T4 GPU environment, enabling efficient handling of image data. As observed during inference, the proposed hybrid model achieved a total inference time of 69.29 seconds on the test dataset. This corresponds to an inference time of approximately 0.0157 seconds per image, translating to an inference speed of 64 images per second. These results indicate that the proposed hybrid model is not only accurate but also computationally efficient, making it well-suited for real-time or near real-time deployment in intelligent waste management systems.

4.3. Comparative Analysis of State-of-the-art models

The proposed hybrid model showed excellent classification ability, achieving a test accuracy of 98.45% on a dataset of 22,000 images, including 12 waste categories. This performance surpasses several existing state-of-the-art models: ResNet50 [7] reported an accuracy of 92.4%, ResNet18 [8] achieved 87.8%, DenseNet169 [10] attained 94.9%, and ResNeXt-101 [17] recorded 89.62%, each classifying waste across 4 to 7 categories. In comparison, other models such as LADS [14] (94.53%), DCNN [15] (93.28%), and Hybrid CNN-LSTM [16] (95.45%) primarily focus on binary classification of organic and recyclable waste, limiting their adaptability in more complex, real-world environments. The proposed hybrid model's superior accuracy is driven by the attention-based fusion of DenseNet201 and ConvNeXtBase, effectively combining dense feature reuse and modern convolutional efficiency to enhance the classification of heterogeneous and visually similar waste types.

5. CONCLUSION AND FUTURE WORK

This research presents a comprehensive approach to solid waste classification using deep learning, intending to enhance urban sustainability and promote effective environmental conservation. A curated and diverse dataset comprising 12 categories of solid waste and a total of 22000 was compiled from multiple public waste image datasets. The dataset includes single-object images with clean backgrounds and real-world images captured under varied environmental conditions, ensuring model generalizability.

Multiple state-of-the-art pretrained convolutional neural networks—DenseNet201, ResNet101, EfficientNetB7, MobileNetV2, InceptionV3, and ConvNeXtBase were fine-tuned and evaluated on the curated dataset. ConvNeXtBase and DenseNet201 achieved the highest test accuracies of 98.13% and 96.45%, respectively, indicating their strong capacity to extract meaningful features from complex waste imagery. Building upon these findings, a novel hybrid CNN model was proposed, integrating DenseNet201 and ConvNeXtBase via an attention-based fusion mechanism. This model effectively combined the complementary strengths of both architectures—DenseNet's feature reuse and gradient flow with ConvNeXt's modern convolutional efficiency and scalability. The hybrid model achieved a test accuracy of 98.45%, outperforming all individual models and demonstrating superior classification performance.

These results highlight the potential of deep learning-based hybrid techniques in automating solid waste classification with high accuracy, which can significantly aid smart waste management systems in real-world scenarios. The proposed technology not only enhances waste sorting precision but also promotes more sustainable urban practices by allowing for effective recycling and disposal.

Future work may involve real-time deployment of the hybrid model in edge devices, exploring more lightweight architectures for resource-constrained environments, and expanding the dataset to include additional subcategories of waste and more geographically diverse samples for broader generalization.

REFERENCES

1. U. N. Environment, “Global Waste Management Outlook 2024 | UNEP - UN Environment Programme.” Accessed: Jun. 17, 2025. [Online]. Available: https://www.unep.org/resources/global-waste-management-outlook-2024
2. 274, “India Solid Waste Management.” Accessed: Jun. 17, 2025. [Online]. Available: https://www.trade.gov/market-intelligence/india-solid-waste-management
3. J. Gutberlet and S. M. N. Uddin, “Household waste and health risks affecting waste pickers and the environment in low- and middle-income countries,” *Int. J. Occup. Environ. Health*, vol. 23, no. 4, pp. 299–310, Oct. 2017, doi: 10.1080/10773525.2018.1484996.
4. A. Lakhouit, “Revolutionizing urban solid waste management with AI and IoT: A review of smart solutions for waste collection, sorting, and recycling,” *Results Eng.*, vol. 25, p. 104018, Mar. 2025, doi: 10.1016/j.rineng.2025.104018.
5. P. Sharma and U. Vaid, “Emerging role of artificial intelligence in waste management practices,” *IOP Conf. Ser. Earth Environ. Sci.*, vol. 889, no. 1, p. 012047, Nov. 2021, doi: 10.1088/1755-1315/889/1/012047.
6. O. Adedeji and Z. Wang, “Intelligent Waste Classification System Using Deep Learning Convolutional Neural Network,” *Procedia Manuf.*, vol. 35, pp. 607–612, Jan. 2019, doi: 10.1016/j.promfg.2019.05.086.
7. Y. Wang, W. J. Zhao, J. Xu, and R. Hong, “Recyclable Waste Identification Using CNN Image Recognition and Gaussian Clustering,” Nov. 02, 2020, *arXiv*: arXiv:2011.01353. doi: 10.48550/arXiv.2011.01353.
8. D. Gyawali, A. Regmi, A. Shakya, A. Gautam, and S. Shrestha, “Comparative Analysis of Multiple Deep CNN Models for Waste Classification,” Aug. 14, 2020, *arXiv*: arXiv:2004.02168. doi: 10.48550/arXiv.2004.02168.
9. G.-L. Huang, J. He, Z. Xu, and G. Huang, “A combination model based on transfer learning for waste classification,” *Concurr. Comput. Pract. Exp.*, vol. 32, no. 19, p. e5751, 2020, doi: 10.1002/cpe.5751.
10. G. Sai Susanth, L. M. Jenila Livingston, and L. G. X. Agnel Livingston, “Garbage Waste Segregation Using Deep Learning Techniques,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1012, no. 1, p. 012040, Jan. 2021, doi: 10.1088/1757-899X/1012/1/012040.
11. Y. Yu and R. Grammenos, “Towards artificially intelligent recycling Improving image processing for waste classification,” Aug. 09, 2021, *arXiv*: arXiv:2108.06274. doi: 10.48550/arXiv.2108.06274.
12. M. Malik *et al.*, “Waste Classification for Sustainable Development Using Image Recognition with Deep Learning Neural Network Models,” *Sustainability*, vol. 14, no. 12, Art. no. 12, Jan. 2022, doi: 10.3390/su14127222.
13. A. S. Ouedraogo, A. Kumar, and N. Wang, “Landfill Waste Segregation Using Transfer and Ensemble Machine Learning: A Convolutional Neural Network

Approach," *Energies*, vol. 16, no. 16, Art. no. 16, Jan. 2023, doi: 10.3390/en16165980.

14. R. Chauhan, S. Shighra, H. Madkhali, L. Nguyen, and M. Prasad, "Efficient Future Waste Management: A Learning-Based Approach with Deep Neural Networks for Smart System (LADS)," *Appl. Sci.*, vol. 13, no. 7, Art. no. 7, Jan. 2023, doi: 10.3390/app13074140.
15. M. Chhabra, B. Sharan, M. Elbarachi, and M. Kumar, "Intelligent waste classification approach based on improved multi-layered convolutional neural network," *Multimed. Tools Appl.*, vol. 83, no. 36, pp. 84095–84120, Nov. 2024, doi: 10.1007/s11042-024-18939-w.
16. U. K. Lilhore, S. Simaiya, S. Dalal, and R. Damaševičius, "A smart waste classification model using hybrid CNN-LSTM with transfer learning for sustainable environment," *Multimed. Tools Appl.*, vol. 83, no. 10, pp. 29505–29529, Mar. 2024, doi: 10.1007/s11042-023-16677-z.
17. H. Ahmed Khan, S. S. Naqvi, A. A. K. Alharbi, S. Alotaibi, and M. Alkhathami, "Enhancing trash classification in smart cities using federated deep learning," *Sci. Rep.*, vol. 14, no. 1, p. 11816, May 2024, doi: 10.1038/s41598-024-62003-4.
18. "garythung/trashnet · Datasets at Hugging Face." Accessed: Jun. 14, 2025. [Online]. Available: https://huggingface.co/datasets/garythung/trashnet
19. S. I. Sam Single, "RealWaste." UCI Machine Learning Repository, 2023. doi: 10.24432/C5SS4G.
20. D. Demetriou, P. Mavromatidis, R. Mwombeki, H. Papadopoulos, M. Petrou, and D. Nicolaides, "Construction and Demolition Waste Object Detection Dataset," vol. 2, Sep. 2023, doi: 10.17632/24d45pf8wm.2.
21. "E Waste Image Dataset." Accessed: Jun. 21, 2025. [Online]. Available: https://www.kaggle.com/datasets/akshat103/e-waste-image-dataset
22. "Garbage Classification (12 classes)." Accessed: Jun. 21, 2025. [Online]. Available: https://www.kaggle.com/datasets/mostafaabla/garbage-classification
23. "Clase Congestion Dataset > Overview," Roboflow. Accessed: Jun. 22, 2025. [Online]. Available: https://universe.roboflow.com/clasificacion-despuntes/clase-congestion
24. "Metal 4 Dataset > Overview," Roboflow. Accessed: Jun. 22, 2025. [Online]. Available: https://universe.roboflow.com/trash-classification-mjez0/metal-4
25. G. Huang, Z. Liu, L. V. D. Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," presented at the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE Computer Society, Jul. 2017, pp. 2261–2269. doi: 10.1109/CVPR.2017.243.
26. K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," presented at the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE Computer Society, Jun. 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90.

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

27. M. Tan and Q. V. Le, “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,” *ArXiv*, May 2019, Accessed: Jun. 29, 2025. [Online]. Available: https://www.semanticscholar.org/paper/EfficientNet%3A-Rethinking-Model-Scaling-for-Neural-Tan-Le/4f2eda8077dc7a69bb2b4e0a1a086cf054adb3f9
28. M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, “MobileNetV2: Inverted Residuals and Linear Bottlenecks,” presented at the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE Computer Society, Jun. 2018, pp. 4510–4520. doi: 10.1109/CVPR.2018.00474.
29. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the Inception Architecture for Computer Vision,” presented at the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE Computer Society, Jun. 2016, pp. 2818–2826. doi: 10.1109/CVPR.2016.308.
30. Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, “A ConvNet for the 2020s,” presented at the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE Computer Society, Jun. 2022, pp. 11966–11976. doi: 10.1109/CVPR52688.2022.01167.