Comparative Analysis of CNN and MobileNetV2 With Data Augmentation and Transfer Learning for Waste Classification

Deas Aghellar[1], Aswan Supriyadi Sunge[2]
[1,2]Faculty of Computer Engineering and Design, Nusaputra University, Sukabumi, Indonesia
[1]deas.aghellar@nusaputra.ac.id, [2]aswan.sunge@pelitabangsa.ac.id

Accepted 21 June 2026
Approved 28 June 2026

***Abstract*— Waste management has become an increasingly important environmental challenge that requires efficient and intelligent classification systems. Deep learning-based approaches have shown significant potential in automating waste classification and improving recycling processes. This study compares the performance of Convolutional Neural Network (CNN) and MobileNetV2 models for image-based waste classification and evaluates the impact of data augmentation and optimization techniques on classification performance.**

The experiments were conducted using the publicly available TrashNet dataset, which contains 2,527 images categorized into six waste classes: cardboard, glass, metal, paper, plastic, and trash. Data augmentation techniques, including rotation, translation, zooming, horizontal flipping, and brightness adjustment, were applied to improve data diversity. MobileNetV2 was implemented using transfer learning with pretrained ImageNet weights, while optimization techniques such as dropout, learning rate adjustment, and early stopping were employed to enhance model performance and generalization capability.

Experimental results show that MobileNetV2 outperformed the conventional CNN model, achieving a validation accuracy of 69.8%, whereas CNN achieved 48.5%. Further evaluation indicated that MobileNetV2 obtained an overall accuracy of 71.97%, precision of 73%, recall of 72%, and F1-score of 71%. Confusion matrix analysis revealed strong classification performance for paper and cardboard categories, while the trash category remained the most challenging due to higher visual diversity and limited training samples. These findings demonstrate that MobileNetV2 combined with data augmentation and optimization techniques provides an effective and computationally efficient approach for waste classification tasks.

***Index Terms*— Waste Classification; Deep Learning; CNN; MobileNetV2; Transfer Learning; Data Augmentation.**

I. INTRODUCTION

Waste management has become an increasingly complex environmental issue due to population growth and intensified human activities. Ineffective waste management can result in environmental pollution, public health problems, and a decline in quality of life. One of the primary challenges lies in the waste sorting process, which is still largely performed manually, making it inefficient and prone to human error. Therefore, technology-based solutions are essential to improve the accuracy and efficiency of waste classification [1], [2].

In recent years, deep learning approaches, particularly Convolutional Neural Networks (CNNs), have been widely applied in image classification due to their ability to automatically extract features [3]. However, conventional CNN models often suffer from limitations in computational efficiency and generalization capability, especially when trained on limited datasets. To address these issues, lightweight architectures such as MobileNetV2 have been developed to improve efficiency without sacrificing accuracy, particularly through transfer learning techniques [4], [5].

Previous studies have compared CNN and MobileNet models in various classification tasks. Zhang et al. [6] demonstrated that MobileNet provides better computational efficiency compared to conventional CNN. Other studies highlight the importance of data augmentation in improving model performance, especially for limited datasets [7]. Although previous studies have explored CNN and MobileNetV2 for waste classification, most research primarily focuses on model comparison and often evaluates optimization techniques separately. Limited studies comprehensively investigate the combined impact of data augmentation, transfer learning, transfer learning with frozen weights, dropout, and learning rate adjustment on waste classification performance using lightweight architectures. Recent studies have shown that transfer learning and lightweight architectures such as MobileNetV2 can significantly improve waste classification performance while reducing computational complexity [18], [20], [23].

Therefore, this study addresses the identified research gap by proposing an integrated framework that combines CNN and MobileNetV2 comparison with data augmentation and optimization techniques. The objective is to evaluate how these techniques contribute to improving classification accuracy, robustness, and generalization capability in waste classification tasks.

II. METHOD

Research Design

This study adopts an experimental approach to compare the performance of CNN and MobileNetV2 models for image-based waste classification. Data augmentation and optimization techniques are integrated to enhance model performance and generalization.

The research workflow includes data collection, preprocessing, model training, evaluation, and result analysis. This approach enables an objective comparison of model performance.

2.2 Related Works

Waste classification has attracted increasing attention in recent years due to the growing environmental impact of improper waste management. Deep learning approaches, particularly Convolutional Neural Networks (CNNs), have been widely adopted because of their ability to automatically extract discriminative visual features from images [2], [3].

Several studies have investigated lightweight architectures to improve computational efficiency while maintaining classification performance. MobileNet and MobileNetV2 have become popular choices because of their depthwise separable convolution mechanism, which significantly reduces the number of parameters and computational requirements [4], [5]. Previous studies reported that MobileNet-based architectures provide competitive performance for image classification tasks while being suitable for resource-constrained devices [6]. More recent studies have demonstrated that MobileNetV2 combined with transfer learning can achieve promising results for waste classification tasks while maintaining computational efficiency [17], [18], [23].

Data augmentation has also been recognized as an effective strategy for improving model robustness and reducing overfitting, particularly when training data are limited [7]. Recent studies further reported that augmentation techniques such as rotation, zooming, and brightness adjustment contribute significantly to improving classification accuracy and model generalization [21], [24]. Furthermore, transfer learning enables models to leverage knowledge learned from large-scale datasets, resulting in improved classification performance and faster convergence [8], [18], [22].

Despite these advances, most previous studies focused either on model comparison or on individual optimization techniques. Limited research has comprehensively investigated the combined effects of transfer learning, data augmentation, dropout, learning rate adjustment, and early stopping within a unified waste classification framework. Therefore, this study proposes an integrated approach that combines CNN and MobileNetV2 comparison with multiple optimization strategies to improve classification accuracy and generalization capability.

Table 1. Summary of Related Works

No	Study	Method	Main Contribution	Limitation
1	Rahman et al. [2]	CNN	Automated waste detection using convolutional neural networks	Limited generalization capability
2	LeCun et al. [3]	Deep Learning	Established the foundation of deep learning for image classification	Not specific to waste classification
3	Sandler et al. [4]	MobileNetV2	Introduced a lightweight architecture with improved efficiency	Requires transfer learning for optimal performance
4	Howard et al. [5]	MobileNet	Developed efficient CNN architectures for mobile devices	Lower representational capacity than larger networks
5	Zhang et al. [6]	Lightweight CNN	Improved computational efficiency for image classification	Limited feature extraction capability
6	Shorten and Khoshgoftaar [7]	Data Augmentatio n	Demonstrated the effectiveness of image augmentation techniques	Does not address model architecture optimization
7	Li et al. [8]	CNN + Data Augmentatio n	Improved classification accuracy using augmentation techniques	Sensitive to augmentation settings
8	Ahmed et al. [9]	Transfer Learning	Enhanced waste classification performance using transfer learning	Dependent on pretrained models
9	Brown et al. [10]	Deep Learning Optimization	Investigated optimization techniques for deep learning models	Did not focus on waste classification
10	This Study	CNN vs MobileNetV2 + Optimization	Combines transfer learning, data augmentation, dropout, learning rate adjustment, and early stopping	Limited dataset size

Although previous studies have demonstrated promising performance in waste classification using CNN and MobileNetV2, several limitations remain. Most studies primarily focused on model comparison without comprehensively integrating optimization strategies such as data augmentation, dropout, learning rate adjustment, and transfer learning. In addition, several studies reported performance degradation when dealing with visually similar waste categories or limited datasets. Therefore, this study proposes an integrated approach combining CNN and MobileNetV2 with multiple optimization techniques to improve classification performance and generalization capability.

2.3 Literature Review Flowchart

As shown in Figure 1, the literature review flowchart illustrates the systematic process of conducting the Systematic Literature Review (SLR), starting from problem identification, followed by literature search, screening, eligibility assessment, data extraction, analysis and synthesis, research gap identification, and concluding with the literature review summary.

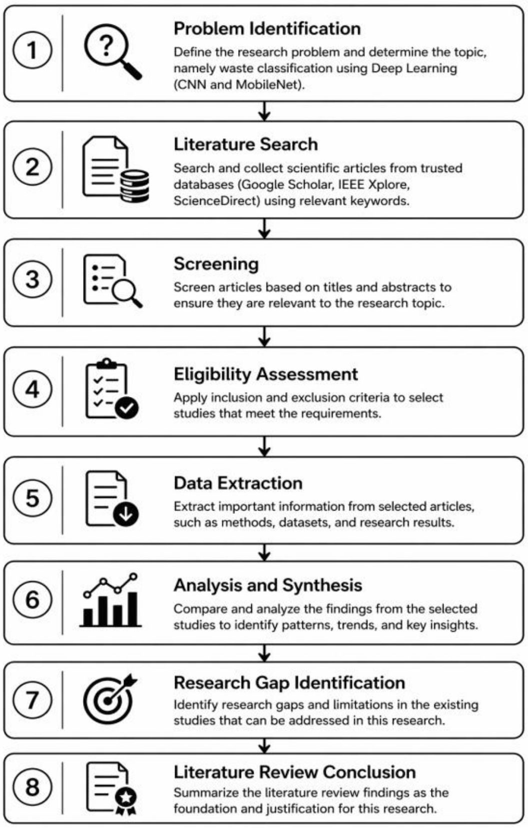

Figure 1. Flowchart

2.4 Dataset

The dataset used in this study was obtained from the publicly available TrashNet dataset available on Kaggle [11].. The dataset consists of six waste categories, namely cardboard, glass, metal, paper, plastic, and trash, with a total of approximately 2,527 images.

Figure 2. Representative samples from the TrashNet waste classification dataset, consisting of six categories:cardboard, glass, metal, paper, plastic, and trash

As illustrated in Figure 2, the dataset consists of six waste categories, namely cardboard, glass, metal, paper, plastic, and trash. Each category contains images with varying lighting conditions, object orientations, and background complexity to represent real-world waste classification scenarios. Each image is in either ***.jpg*** or ***.png*** format with varying resolutions. To ensure uniform input for the model, all images are resized to 224×224 pixels, which is compatible with the requirements of the Deep Learning architecture used, particularly MobileNetV2.

The dataset is divided into two main parts:

Training data: 80% of the total dataset (approximately 2,024 images), used to train the model.

Validation data**:** 20% of the total dataset (approximately 503 images), used to evaluate model performance during training.

The data splitting process is performed automatically using a validation split parameter in the

ImageDataGenerator, ensuring a balanced distribution across classes.
Furthermore, to increase data variability and reduce the risk of overfitting, data augmentation techniques are applied to the training data. These include rotation, translation, scaling (zoom), horizontal flipping, and brightness adjustment. These techniques aim to simulate various real-world image conditions.

With diverse dataset characteristics and appropriate preprocessing, the resulting model is expected to have good generalization capability in classifying different types of waste.

2.4.1 Data Augmentation Strategy

To improve model robustness and reduce overfitting, data augmentation techniques were applied to the training dataset using the *ImageDataGenerator* framework. The augmentation process was designed to simulate various real-world waste image conditions, including variations in object orientation, lighting, and scale.
The augmentation techniques applied in this study are summarized in Table 2.

Table 2. Data Augmentation Configuration

Technique	Parameter
Rotation Range	20°
Zoom Range	0.2
Width Shift Range	0.2
Height Shift Range	0.2
Horizontal Flip	True
Brightness Range	0.8–1.2
Rescaling	1/255

These augmentation techniques help improve the diversity of the training data and enable the model to better generalize to unseen data. Rotation and shifting were applied to simulate changes in object positioning, while brightness adjustment was used to represent varying illumination conditions commonly found in real-world environments. Furthermore, horizontal flipping and zooming were introduced to increase variability in object appearance and scale.

2.5 Model Architecture

2.5.1 Convolutional Neural Network (CNN)

The Convolutional Neural Network (CNN) model used in this study is a basic architecture designed to automatically extract features from input images. The CNN architecture consists of several main layers arranged sequentially as follows:

Input Layer. The input images have a size of 224×224×3 (RGB), which conforms to the standard requirements of Deep Learning models.

Convolutional Layer. This layer functions to extract features from images using filters or kernels. In this study, several convolutional layers are used with a kernel size of 3×3 and ReLU (Rectified Linear Unit) activation to enhance the model's non-linear capability.

Pooling Layer. After the convolution process, MaxPooling with a size of 2×2 is applied to reduce the feature dimensions (downsampling) and decrease computational complexity. This process also helps reduce the risk of overfitting.

Feature Extraction (Stacked Layers). A combination of convolutional and pooling layers is applied in a stacked manner to produce more complex and abstract feature representations from the input images.

Flatten Layer. The extracted features are transformed into a one-dimensional vector so that they can be processed by the fully connected layers.

Fully Connected Layer (Dense Layer). This layer performs classification based on the extracted features. A dense layer with ReLU activation is used.

Dropout Layer. To reduce overfitting, dropout is applied with a certain rate (e.g., 0.5–0.7), which randomly deactivates a portion of neurons during training.

Output Layer. The final layer uses the Softmax activation function to produce classification probabilities across six waste categories. The CNN architecture consists of several main stages, namely the input layer, convolutional layers for feature extraction, pooling layers for dimensionality reduction, a flatten layer for data transformation, and fully connected layers for classification. This model is used as a baseline in this study, as illustrated in Figure 3.

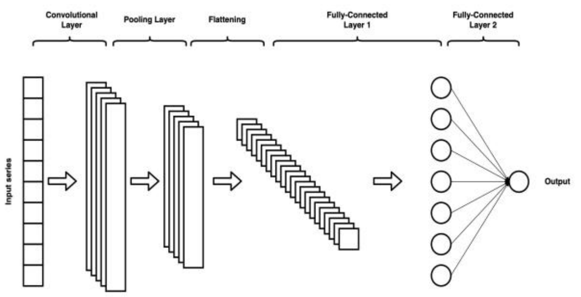

Figure 3. CNN Architecture

2.5.2 MobileNetV2

The MobileNetV2 model is used in this study as the main model due to its advantages in computational efficiency and lightweight architecture. MobileNetV2 is designed using the concept of depthwise separable convolution, which reduces the number of parameters and computational cost compared to conventional CNNs.

In this study, MobileNetV2 is implemented using a transfer learning approach by utilizing pretrained weights from the ImageNet dataset. This approach enables the model to leverage previously learned basic features such as edges, textures, and other visual patterns.

The MobileNetV2 architecture in this study consists of several main components:

Base Model (Feature Extractor). The main part of MobileNetV2 is used as a feature extractor with pretrained weights. In the initial stage, most layers are frozen to preserve the model's prior knowledge.

Global Average Pooling Layer. This layer is used to reduce feature dimensions without significantly increasing the number of parameters, while still retaining important information.

Fully Connected Layer (Dense Layer). Added to adapt the model to the number of classes in the dataset.

Output Layer. Uses the Softmax activation function to produce classification probabilities across six classes.

In this study, transfer learning was implemented using pretrained ImageNet weights. The pretrained layers were kept frozen during training to preserve previously learned visual representations while reducing computational complexity.

MobileNetV2 utilizes the concepts of depthwise separable convolution and inverted residual blocks, allowing the model to be more lightweight and efficient. This model is implemented using transfer learning and with frozen pretrained weights to improve classification performance, as illustrated in Figure 4.

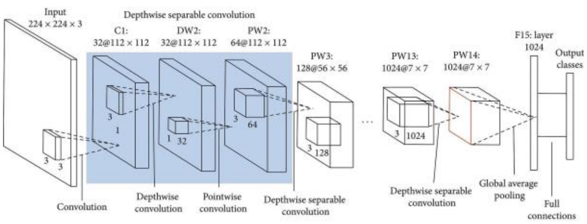

Figure 4. MobileNetV2 architecture used as the pretrained feature extractor in this study.

As illustrated in Figure 4, MobileNetV2 employs depthwise separable convolution and inverted residual blocks to reduce computational complexity while maintaining feature extraction capability. In this study, MobileNetV2 was implemented using transfer learning with pretrained ImageNet weights, while the pretrained layers remained frozen during training.

2.6 Experimental Configuration

The experimental setup was designed to evaluate and compare the performance of CNN and MobileNetV2 models for waste classification. Both models used input images resized to 224×224 pixels and were trained using the same training and validation datasets generated through the ImageDataGenerator framework with an 80:20 split ratio.

The CNN model consisted of three convolutional layers with 32, 64, and 128 filters, respectively, followed by MaxPooling layers for dimensionality reduction. A fully connected layer containing 128 neurons and a dropout rate of 0.5 were applied to reduce overfitting. The CNN model was trained using the Adam optimizer with a batch size of 32 for 10 epochs.

For MobileNetV2, transfer learning was implemented using pretrained ImageNet weights. The pretrained layers were kept frozen during training to preserve previously learned visual features. A Global Average Pooling layer, a dense layer with 128 neurons, and a dropout rate of 0.7 were added to adapt the model to the waste classification task. The MobileNetV2 model was trained using the Adam optimizer with a learning rate of 0.0001 and a batch size of 32 for 15 epochs.

To improve training stability and reduce overfitting, an early stopping mechanism was applied with a patience value of 3 epochs. The training process automatically restored the best-performing model weights based on validation performance.

The experimental configuration used in this study is summarized in Table 3. Both CNN and MobileNetV2 models were trained using the same dataset partition and preprocessing pipeline to ensure a fair comparison.

Table 3. Experimental Configuration

Parameter	CNN	MobileNetV2
Input Size	224×224	224×224
Batch Size	32	32
Epoch	10	15
Optimizer	Adam	Adam
Learning Rate	Default Adam	0.0001
Dropout	0.5	0.7

Parameter	CNN	MobileNetV2
Transfer Learning	No	Yes
Pretrained Weights	No	ImageNet
Early Stopping	Yes	Yes
Patience	3	3

Overall, the selected configuration was designed to balance classification performance, computational efficiency, and model generalization capability. The same experimental settings were maintained throughout the study to ensure a consistent comparison between CNN and MobileNetV2.

III. RESULT AND DISCUSSIONS

Experimental Results

This study evaluated the performance of two deep learning models, namely CNN and MobileNetV2, for image-based waste classification. Both models were trained using the same dataset, preprocessing procedures, and data augmentation strategies to ensure a fair comparison.

Table 4. Model training and validation Accuracy of CNN and MobileNetV2

Model	Training Accuracy	Validation Accuracy
CNN	54%	48.5%
MobileNetV2	70%	69.8%

Table 4 presents the training and validation accuracy achieved by each model. The results indicate that MobileNetV2 outperformed CNN, achieving a validation accuracy of 69.8%, whereas CNN obtained 48.5%. This finding demonstrates the effectiveness of transfer learning in improving classification performance.

To provide a more comprehensive evaluation, additional performance metrics were calculated for the MobileNetV2 model, including precision, recall, and F1-score. The detailed results are presented in Table 5 and discussed in Section 3.3.. The MobileNetV2 model achieved an overall accuracy of 71.97%, with a precision of 73%, recall of 72%, and F1-score of 71%. These results indicate that MobileNetV2 is capable of maintaining balanced classification performance across different waste categories while achieving reliable prediction accuracy.

The difference between the validation accuracy reported in Table 4 and the overall accuracy presented in Table 5 is attributed to the final evaluation procedure performed on the validation dataset after model training

The training and validation accuracy curves of both models are illustrated in Figure 5 , which shows the learning behavior of the models throughout the training process.

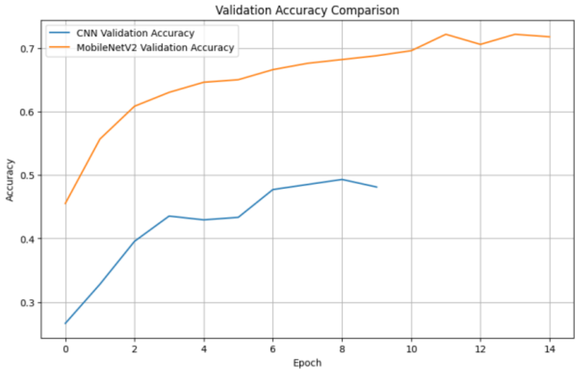

Figure 5. Comparison of validation accuracy between the CNN and MobileNetV2 models during the training process.

As shown in Figure 5, MobileNetV2 consistently achieved higher validation accuracy than CNN throughout the training process. The validation accuracy of MobileNetV2 increased steadily and stabilized at approximately 70%, whereas CNN showed slower improvement and reached a maximum validation accuracy of approximately 49%. This result further confirms the effectiveness of transfer learning in improving feature extraction and classification performance for waste classification tasks.

3.2 Model Performance Analysis

Based on the experimental results, MobileNetV2 demonstrated superior performance compared to the conventional CNN model in waste classification tasks. As shown in Table 4, MobileNetV2 achieved a validation accuracy of 69.8%, significantly outperforming CNN, which obtained 48.5%.

The lower performance of CNN may be attributed to its limited capability in extracting complex visual features from waste images, particularly when trained on a relatively small dataset. Furthermore, the CNN model exhibited indications of underfitting, as reflected by the relatively low accuracy observed in both training and validation phases. This suggests that the CNN architecture employed in this study may not have been sufficiently deep or complex to effectively capture discriminative patterns among waste categories.

In contrast, MobileNetV2 achieved better performance due to the use of transfer learning with pretrained ImageNet weights. The pretrained feature extraction mechanism enables MobileNetV2 to leverage previously learned low-level and high-level visual features, such as texture, shape, and edge information, thereby improving classification capability.

Figure 5 further illustrates that MobileNetV2 consistently achieved higher validation accuracy than CNN throughout the training process. The validation accuracy of MobileNetV2 increased steadily and stabilized at approximately 70%, whereas CNN showed slower improvement and reached a maximum validation accuracy of approximately 49%. This result confirms the effectiveness of transfer learning in improving feature extraction and classification performance for waste classification tasks.

3.3 Model Evaluation Metrics

To evaluate the classification performance more comprehensively, additional evaluation metrics including accuracy, precision, recall, and F1-score were calculated for the MobileNetV2 model. These metrics provide a more detailed assessment of model effectiveness beyond validation accuracy alone.

Table 5. Performance Evaluation Metrics of MobileNetV2

Metric	Score
Accuracy	71.97%
Precision	73%
Recall	72%
F1-score	71%

Table 5 presents the evaluation results obtained from the classification report. The MobileNetV2 model achieved an overall accuracy of 71.97%, indicating that the model correctly classified the majority of waste images in the validation dataset. Furthermore, the model obtained a weighted precision of 73%, recall of 72%, and F1-score of 71%.

The relatively balanced values of precision, recall, and F1-score indicate that the model maintained consistent classification performance across different waste categories. The weighted precision of 73% suggests that most predicted classes were correctly identified, while the recall value of 72% indicates that the model successfully detected a large proportion of actual waste categories. In addition, the F1-score of 71% demonstrates a balanced trade-off between precision and recall.

Overall, these results confirm that MobileNetV2 provides reliable classification performance and exhibits good generalization capability for waste classification tasks.

3.4 The Effect of Data Augmentation and Optimization

Data augmentation and optimization techniques played an important role in improving the performance of the waste classification models. Data augmentation increased the diversity of the training dataset by generating modified versions of existing images through transformations such as rotation, translation, zooming, horizontal flipping, and brightness adjustment. As a result, the models were exposed to a wider range of visual variations, enabling them to generalize better to unseen data.

The use of data augmentation helped reduce the risk of overfitting, particularly because the dataset contained a relatively limited number of images. By simulating different real-world conditions, the model became more robust to variations in object orientation, scale, illumination, and background complexity.

In addition to data augmentation, several optimization techniques were applied during training. Dropout layers were incorporated to reduce excessive reliance on specific neurons and improve model generalization. Furthermore, a lower learning rate was used in MobileNetV2 to ensure more stable parameter updates during training. Early stopping was also employed to prevent overfitting by terminating the training process when no significant improvement in validation performance was observed.

The combination of data augmentation and optimization techniques contributed to the improved performance of MobileNetV2, resulting in more stable training behavior and better validation accuracy compared to the CNN model.

3.5 Confusion Matrix Analysis

The confusion matrix was used to evaluate the distribution of model predictions across the six waste categories. Figure 6 presents the confusion matrix of the MobileNetV2 model on the validation dataset.

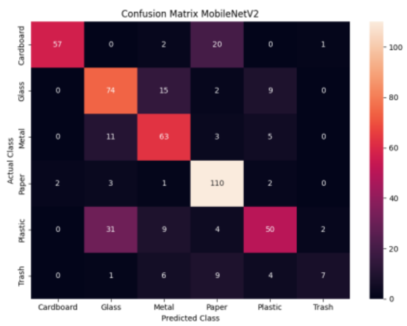

Figure 6. Confusion matrix of the MobileNetV2 model showing the distribution of predicted classes against actual classes in the validation dataset.

As shown in Figure 6, the model achieved strong classification performance for several categories, particularly paper and cardboard, which obtained F1-scores of 0.85 and 0.83, respectively. The high number of correctly classified samples in these categories indicates that the model was able to learn distinctive visual features effectively.

The metal category also demonstrated good performance, achieving a recall of 0.83 and an F1-score of 0.72. However, some misclassifications were observed between visually similar categories such as plastic, glass, and metal, suggesting overlapping visual characteristics among these waste types.

The trash category achieved the lowest performance, with an F1-score of 0.33 and a recall of 0.22. This result may be attributed to the limited number of training samples and the higher visual diversity within the trash category. Consequently, the model encountered greater difficulty in learning representative features for this class.

Overall, the confusion matrix analysis confirms that MobileNetV2 achieved relatively balanced classification performance, with an overall accuracy of 71.97%, weighted precision of 73%, recall of 72%, and F1-score of 71%.

3.6 Discussion

The experimental results demonstrate that MobileNetV2 significantly outperformed the conventional CNN model in waste classification tasks. MobileNetV2 achieved a validation accuracy of 69.8% and an overall evaluation accuracy of 71.97%, whereas CNN achieved a validation accuracy of only 48.5%. These findings indicate that transfer learning is effective in improving classification performance, particularly when working with relatively limited datasets.

The superior performance of MobileNetV2 can be attributed to its pretrained feature extraction capability. By utilizing weights learned from the large-scale ImageNet dataset, MobileNetV2 is able to capture low-level and high-level visual features such as edges, textures, shapes, and object patterns more effectively than a CNN trained from scratch. Consequently, the model exhibits better generalization capability and classification accuracy.

The confusion matrix analysis further revealed that the model performed particularly well in classifying the paper, cardboard, and metal categories. These classes achieved relatively high recall and F1-score values, indicating that their visual characteristics were sufficiently distinctive for the model to learn. In contrast, the trash category achieved the lowest performance, which may be attributed to the limited number of training samples and the greater visual diversity within the class.

In addition, some misclassifications were observed between plastic, glass, and metal categories. These errors may be caused by similarities in texture, color, reflective surfaces, and object appearance, which make it difficult for the model to distinguish between classes. Such challenges are commonly reported in waste classification studies and highlight the need for larger and more diverse datasets.

The implementation of data augmentation and optimization techniques also contributed positively to model performance. Data augmentation increased the diversity of training samples, while dropout and early stopping helped reduce overfitting and improve training stability. As a result, MobileNetV2 achieved a more balanced performance between training and validation data.

The findings obtained in this study are consistent with recent studies that reported the effectiveness of transfer learning and lightweight deep learning architectures for waste classification tasks [18], [20], [23]. Although the accuracy achieved in this study is lower than some state-of-the-art approaches, the proposed framework demonstrates competitive performance while maintaining computational efficiency.

3.7 Limitations

This study has several limitations. First, the dataset size is relatively limited, which may affect the model's ability to learn highly discriminative features. Second, certain classes, particularly trash, contain fewer samples and exhibit greater visual diversity, resulting in lower classification performance. Third, the pretrained MobileNetV2 layers were kept frozen during training, which may limit adaptation to the waste classification dataset.

IV. Conclusion

Based on the experimental results, it can be concluded that MobileNetV2 outperformed the conventional CNN model in image-based waste classification. The MobileNetV2 model achieved a validation accuracy of 69.8% and an overall evaluation accuracy of 71.97%, while the CNN model obtained a validation accuracy of 48.5%. These results indicate that transfer learning using pretrained ImageNet weights can significantly improve classification performance compared to a CNN model trained from scratch.

Additional evaluation metrics further confirmed the effectiveness of MobileNetV2, achieving a precision of 73%, recall of 72%, and F1-score of 71%. The confusion matrix analysis showed that the model performed particularly well in classifying paper, cardboard, and metal categories, while the trash category remained challenging due to limited training samples and greater visual diversity.

The implementation of data augmentation techniques, including rotation, translation, zooming, horizontal flipping, and brightness adjustment, contributed to improving model robustness and generalization capability. Furthermore, optimization strategies such as dropout, learning rate adjustment, and early stopping helped improve training stability and reduce the risk of overfitting.

Overall, the results demonstrate that MobileNetV2 is an effective and computationally efficient approach for waste classification tasks. Future research may focus on expanding the dataset size, addressing class imbalance issues, and exploring advanced transfer learning and fine-tuning strategies to further improve classification performance.

REFERENCES

[1] A. Kumar et al., "Waste Classification Using Deep Learning," Journal of Environmental Informatics, 2022.

[2] S. Rahman et al., "Automated Waste Detection Using CNN," IEEE Access, 2020.

[3] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[4] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4510–4520, 2018.

[5] A. Howard et al., "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," arXiv:1704.04861, 2017.

[6] Z. Zhang et al., "Lightweight CNN Models for Image Classification," 2021.

[7] S. Shorten and T. M. Khoshgoftaar, "A Survey on Image Data Augmentation for Deep Learning," Journal of Big Data, vol. 6, no. 1, pp. 1–48, 2019.

[8] H. Li et al., "Improving Image Classification Using Data Augmentation," 2023.

[9] M. Ahmed et al., "Transfer Learning for Waste Classification," 2024.

[10] J. Brown et al., "Deep Learning Optimization Techniques," 2022.

[11] G. Thung and M. Yang, "Classification of Trash for Recyclability Status," Stanford CS229 Project Report, 2016.

[12] M. Yang and G. Thung, "TrashNet: A Dataset for Image Classification of Waste," Stanford University Technical Report, 2021.

[13] S. R. Dubey, S. K. Singh, and R. K. Chaudhuri, "Activation Functions in Deep Learning: A Comprehensive Survey and Benchmark," Neurocomputing, vol. 503, pp. 92–108, 2022.

[14] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, "A Survey of the Recent Architectures of Deep Convolutional Neural Networks," Artificial Intelligence Review, vol. 53, no. 8, pp. 5455–5516, 2021.

[15] M. Tan and Q. Le, "EfficientNetV2: Smaller Models and Faster Training," in Proceedings of the International Conference on Machine Learning (ICML), 2021, pp. 10096–10106.

[16] S. V. Militante and N. V. Dionisio, "Deep Learning Implementation for Waste Segregation and Classification System," International Journal of Advanced Computer Science and Applications, vol. 12, no. 2, pp. 232–240, 2021.

[17] R. M. S. Verma and P. Singh, "Transfer Learning-Based Waste Classification Using MobileNetV2," Proceedings of the International Conference on Smart Computing and Electronic Enterprise, 2022.

[18] A. Alqahtani, M. Alsubaie, and M. Alqahtani, "Automatic Waste Classification Using Deep Transfer Learning Models," Sustainability, vol. 15, no. 4, 2023.

[19] J. H. Park and S. Lee, "Lightweight Deep Learning Models for Smart Waste Management Systems," Sensors, vol. 23, no. 10, 2023.

[20] M. A. Islam, M. R. Hossain, and M. S. Rahman, "Comparative Analysis of CNN and Transfer Learning Models for Waste Classification," IEEE Access, vol. 12, pp. 45612–45625, 2024.

[21] P. Sharma and R. Gupta, "Data Augmentation Strategies for Improving Waste Image Classification Performance," Journal of Imaging, vol. 10, no. 3, 2024.

[22] Y. Chen, X. Li, and H. Wang, "Transfer Learning for Resource-Efficient Waste Classification Systems," Expert Systems with Applications, vol. 245, 2024.

[23] S. Ahmed, M. Khan, and F. Ali, "MobileNetV2-Based Waste Classification for Smart Recycling Applications," Applied Sciences, vol. 14, no. 5, 2024.

[24] T. Nguyen, H. Tran, and K. Le, "Optimization Techniques for Deep Learning-Based Waste Classification," Engineering Applications of Artificial Intelligence, vol. 131, 2025.

[25] R. Kumar and V. Patel, "Advanced Deep Learning Approaches for Automated Waste Classification: A Review," IEEE Access, vol. 13, pp. 11245–11267, 2025.