
Development of a CNN-Based Household Waste Image Classification Model for Automatic Identification

Achmad Isra'M¹, Azmuri Wahyu Azinar¹, Suhendro Busono¹, Hamzah Setiawan¹

¹ Muhammadiyah University of Sidoarjo, Faculty of Science and Technology, Informatics, Jl. Raya Gelam No. 250, Pagerwaja, Gelam, Candi District, Sidoarjo Regency, East Java 61271, Indonesia

Article History

Submitted : 25-06-2026
Revised : 19-08-2026
Accepted : 20-08-2026
Published : 22-09-2026
DOI: 10.32664/smatika.v16i03.2436

Keyword

Automatic Waste Sorting; Convolutional Neural Network (CNN); Deep Learning; Image Processing; Waste Classification

*Corresponding Author:

mamatsatunya@gmail.com

Abstract

The problem of household waste in Indonesia remains a major challenge due to the low level of waste sorting and the lack of technology utilization in the waste identification process. This study aims to develop a Deep Learning-based inorganic waste identification system using the Convolutional Neural Network (CNN) method to improve the accuracy of automatic classification in the household environment. The research method used is an experiment with a quantitative approach. The dataset used consists of 92 images of inorganic waste covering the categories of plastic, paper, glass, and metal. The data went through a preprocessing stage in the form of image resizing to 224×224 pixels, normalization, and data augmentation, then divided into training and testing data with a ratio of 80:20. The CNN model was trained using variations in epoch parameters, batch size, and learning rate, then evaluated using metrics of accuracy, precision, recall, F1-score, and confusion matrix. The results showed that the training parameters had a significant effect on model performance. The optimal value was obtained at epoch 20 with 97% accuracy, a batch size of 8 with 97% accuracy, and a learning rate of 1×10^{-3} with 90% accuracy. The best parameter combination produced a model with 97% classification accuracy, along with balanced precision, recall, and F1-score values. These results demonstrate the effectiveness of the CNN method in automatically identifying inorganic waste types. The developed system has the potential to be applied in household environments to assist waste sorting, improve waste management efficiency, and support environmental conservation by raising public awareness of the importance of sustainable waste management. Although the model achieved 97% accuracy under a single 80:20 train-test split, its performance decreased substantially to $32.69\% \pm 6.93\%$ under stratified 5-fold cross-validation, indicating potential overfitting and limited generalization.

1. Introduction

Waste management remains one of the major environmental challenges faced by Indonesia due to the large volume of waste generated from household and other daily activities [1]. The national waste management data provided through the National Waste Management Information System (SIPSN) show that waste generation remains substantial, while a considerable proportion of waste is still not properly managed [1]. Waste

segregation at the source is an important component of household waste management because separating waste according to its type can facilitate subsequent recycling and treatment processes [2]. However, household waste segregation remains an important challenge in Indonesia, indicating the need for approaches that can support more efficient waste sorting at the source [2]. Consequently, automated waste classification has emerged as a potential approach to support more efficient waste sorting and sustainable waste management [3],[4]. Recent developments in Artificial Intelligence (AI), particularly Deep Learning, have provided opportunities for developing image-based waste classification systems [3], [4]. Deep Learning-based methods have been widely investigated for automatic waste classification because they can learn discriminative features directly from image data [3], [5],[6].

One of the most widely used architectures for image classification is the Convolutional Neural Network (CNN), which can automatically learn spatial features and visual patterns from images [4],[7]. The ability of CNN-based models to extract visual characteristics makes them suitable for identifying different categories of waste from digital images [7],[8]. Several previous studies have investigated CNN-based approaches for waste classification. Bahagia and Akbar reported that CNN could be used to distinguish organic and inorganic waste categories [8]. Similarly, Diki developed a CNN-based waste classification system and reported high classification accuracy on the evaluated dataset [7]. Sadida Aulia et al. investigated household waste classification using a CNN-based approach [9]. Furthermore, Agustiani et al. compared several deep learning architectures, including EfficientNetB3, MobileNetV2, and ResNet50, for waste classification and reported that EfficientNetB3 achieved the best performance among the evaluated architectures [10]. Other studies have also investigated CNN-based approaches for multi-class waste classification [11],[12],[13],[14]. Although these studies have demonstrated the potential of deep learning for waste classification, differences remain in dataset size, waste categories, model architectures, and experimental conditions [3], [10], [15], [5]. Several studies have used publicly available waste datasets or datasets collected under specific image acquisition conditions. For example, previous household waste classification studies have investigated datasets containing different waste categories and image characteristics [9]–[5].

Consequently, model performance obtained from a particular dataset may not necessarily represent performance under different household image conditions. Real household environments can introduce variations in lighting, background, object orientation, object position, and image quality. These factors can influence the visual characteristics available to an image classification model. Previous studies have therefore investigated household waste classification and real-time waste detection using deep learning approaches [15], [16]. This indicates that evaluating waste classification methods in household-oriented settings remains relevant. In addition to model architecture, training configuration can influence the performance of a CNN model. Parameters such as the number of epochs, batch size, and learning rate determine how the model learns from the training data and can affect the resulting classification performance. Previous waste classification studies have mainly focused on developing and evaluating classification models using different CNN or deep learning architectures [7],[5],[17],[6]. However, fewer studies have specifically emphasized a controlled empirical analysis of the influence of training hyperparameters on CNN performance under a small-data setting.

Luntungan developed an automatic waste classification system using CNN-based deep learning to support smart waste management [18]. Simbolon et al. investigated household waste image classification using a deep learning model and considered four categories of household inorganic waste, namely glass, plastic, metal, and paper [15]. These studies demonstrate the feasibility of applying deep learning to household waste classification. However, the present study focuses specifically on the empirical evaluation of selected CNN training hyperparameters rather than proposing a new CNN architecture. Therefore, this study does not claim novelty in applying CNN to household waste classification itself. Instead, its contribution lies in providing a controlled empirical evaluation of the sensitivity of CNN classification performance to selected training hyperparameters, namely the number of epochs, batch size, and learning rate, under a small-data setting. The analysis is followed by a confirmatory experiment to identify a reproducible and practical training configuration for a small household-waste image dataset.

This research aims to develop an automatic household waste identification system based on Deep Learning using the Convolutional Neural Network (CNN) method. The system is designed to classify four categories of inorganic waste, namely plastic, paper, glass, and metal, using image data obtained from a publicly available dataset. In addition, this study evaluates the impact of several training parameters, including the number of epochs, batch size, and learning rate, on classification performance. The results are expected to provide empirical insights into CNN training under small-data conditions and contribute to the development of practical image-based waste classification systems for supporting household waste sorting activities.

Table 1. Comparative Summary of Previous Studies on Waste Classification

Study	Method/Architecture	Dataset Size	Classes	Accuracy
Diqi (2022)	CNN	22,564	2	99.69%
Bahagia & Akbar (2024)	CNN	25,112	2	95%
Agustiani et al. (2025)	EfficientNetB3	4,650	6	93%
Simbolon et al. (2025)	MobileNetV2	NR	4	>80%
Luntungan (2025)	CNN	NR	2	94.86%*
This study	CNN + Hyperparameter Optimization	92	4	97%

Based on the comparison presented in Table 1, previous studies have demonstrated that deep learning methods, particularly Convolutional Neural Networks (CNNs) and transfer learning approaches, can achieve high performance in waste classification. One study employed a CNN using a dataset of 22,564 images covering two waste classes and achieved an accuracy of 99.69% [7]. Similarly, another study applied a CNN to 22,564 images from two waste categories and reported an accuracy of 95% [8]. A comparative study of several transfer learning architectures used a dataset of 4,650 images across six waste classes [10]. In that study, EfficientNetB3 achieved the highest accuracy of 93%, followed by ResNet50 at approximately 91%, while MobileNetV2 achieved approximately 70–73%. Another study employed MobileNetV2 for four-class household waste classification and reported an accuracy above 80% [9]. In addition, a CNN-based waste classification system achieved a training accuracy of 94.86% [15].

Compared with these previous studies, the present study uses a substantially smaller dataset consisting of only 92 images distributed across four waste classes. The best CNN configuration obtained through hyperparameter evaluation achieved an accuracy of 97% using 20 epochs, a batch size of 8, and a learning rate of 1×10^{-3} . However, this result should be interpreted with caution because the evaluation was conducted using a limited dataset and a single train–test split. Therefore, the 97% accuracy should not be interpreted as evidence that the proposed model outperforms previous studies using substantially larger datasets. Instead, the result demonstrates that the selected hyperparameter configuration can produce favorable classification performance under the experimental conditions of this study.

Overall, previous studies generally employ larger datasets and focus on the application or comparison of CNN and transfer learning architectures. In contrast, the present study emphasizes the empirical evaluation of CNN hyperparameters, particularly the number of epochs, batch size, and learning rate, for four-class household waste classification using a limited dataset. Therefore, the contribution of this study is not to claim superior accuracy over previous approaches, but to investigate how hyperparameter configurations influence CNN classification performance under limited-data conditions.

2. Research Method

This research employed an experimental method with a quantitative approach to develop and evaluate an automatic household waste classification system based on Deep Learning using the Convolutional Neural Network (CNN) method. The experimental approach was chosen to investigate the effect of several training parameters on the performance of the classification model and to determine the optimal parameter

configuration for achieving the highest classification accuracy. To determine the optimal CNN training configuration, three hyperparameter experiments were conducted: epoch variation, batch size variation, and learning rate variation. Each experiment was evaluated using accuracy, precision, recall, and F1-score metrics. Only one parameter was modified at a time while keeping the other training configurations fixed to ensure a fair comparison. Generally methods include the following

2.1 Research flow

The research was conducted through several stages, namely data collection, data preprocessing, dataset splitting, CNN model training, model evaluation, and system implementation. The overall research flow is illustrated in Figure 1.

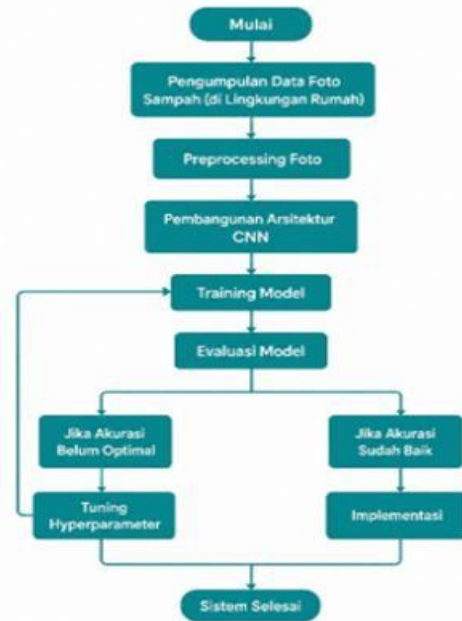


Figure 1. Research Flowchart

The process begins with collecting waste image datasets from online sources. The collected images are then preprocessed to improve data quality and prepare them for the training process. After preprocessing, the dataset is divided into training and testing data. The CNN model is trained using the training dataset with different combinations of epoch, batch size, and learning rate values. Subsequently, the trained model is evaluated using the testing dataset. Finally, the best-performing model is integrated into an automatic waste classification system. Hyperparameter tuning was performed sequentially using a one-factor-at-a-time (OFAT) strategy. The effects of the number of epochs, batch size, and learning rate were evaluated independently while the other parameters were kept constant. The best-performing value obtained for each parameter was subsequently used in the final configuration. A confirmation experiment was then conducted using the resulting combination of selected hyperparameter values.

2.2 Dataset Collection

The dataset consisted of 92 original images representing four household waste categories. Since data augmentation was performed online during the training process, augmented images were not permanently stored as additional files. Instead, image transformations were generated dynamically for each training batch. With 73 training images, a batch size of 8, and 25 training epochs, the training process involved 10 steps per epoch, corresponding to 80 image instances processed per epoch and 2,000 effective image exposures throughout the training process. Therefore, the physical dataset remained at 92 original images, while the effective number of training image instances increased through online augmentation [9],[15] Organic and

residual waste were excluded from this study because their identification may require non-visual characteristics, such as odor, moisture content, or material composition, which are beyond the scope of the proposed image-classification approach.

Table 2. Dataset and Training Configuration

Parameter	Value
Original images	92
Number of classes	4
Training images	73
Testing images	19
Image size	224 × 224 pixels
Batch size	8
Epochs	25
Augmentation	Online/dynamic
Effective training image exposures	2,000

2.3 Transfer Learning Implementation

Data preprocessing was performed to improve image quality and increase the robustness of the model. The preprocessing stage consisted of the following steps: Image Resizing All images were resized to 224 × 224 pixels to ensure uniform input dimensions for the CNN model. Normalization Pixel values were normalized into the range of 0–1 to accelerate the learning process and improve training stability. Data Augmentation Data augmentation techniques such as rotation, horizontal flipping, zooming, and brightness adjustment were applied to increase data diversity and reduce overfitting. After preprocessing, Due to the limited number of available images, Stratified 5-Fold Cross-Validation was employed instead of a single train-test split. The 92 images were divided into five stratified folds while maintaining the class distribution. In each iteration, four folds were used for training and one fold was used for validation. The process was repeated five times so that each image was used once as validation data. Online data augmentation was applied only to the training subset to prevent data leakage, where 80% of the data was used for model training and 20% for performance evaluation. To improve data diversity and reduce the risk of overfitting, online data augmentation was applied to the training images. The augmentation process was performed dynamically during model training rather than by permanently creating additional image files. The applied transformations included image rotation, horizontal flipping, zooming, and brightness variation. These transformations generated different image representations across training iterations while maintaining the original dataset size of 92 images. With 73 images in the training subset, a batch size of 8, and 25 epochs, the training procedure produced 2,000 effective image exposures.

2.4 CNN Model Architectur

The proposed CNN architecture accepts RGB images with an input size of 224 × 224 × 3. The first convolutional layer consists of 32 filters with a 3 × 3 kernel followed by the ReLU activation function and a 2 × 2 MaxPooling layer. The second convolutional layer uses 64 filters (3 × 3) followed by ReLU activation and MaxPooling. The third convolutional layer contains 128 filters (3 × 3) with ReLU activation and MaxPooling. The extracted feature maps are then flattened and connected to a fully connected layer containing 128 neurons with ReLU activation. A Dropout layer (0.5) is applied to reduce overfitting before the output layer. Finally, a Dense layer with 4 neurons and Softmax activation is used to classify the four waste categories (plastic, paper, glass, and metal). The architecture contains approximately 11,169,476 trainable parameters.

2.4.1 Input Layer

The input images were resized to 224 × 224 pixels with three RGB color channels. Therefore, the input tensor of the proposed CNN was defined as 224 × 224 × 3. The input images were normalized to a pixel-value range of 0 to 1 before being fed into the network.

2.4.2 Convolutional Layers

Table 3. CNN Architecture

Layer	Filters/Neurons	kernel	Activation
Conv2D-1	32	3 × 3	ReLU
Conv2D-2	64	3 × 3	ReLU
Conv2D-3	128	3 × 3	ReLU

The feature extraction component consists of three convolutional layers. The first convolutional layer uses 32 filters with a 3×3 kernel, followed by a ReLU activation function. The second convolutional layer increases the number of filters to 64, while the third convolutional layer uses 128 filters. All convolutional layers employ 3×3 kernels and ReLU activation to extract hierarchical visual features from the input images.

2.4.3 Pooling Layers

Each convolutional layer is followed by a 2×2 max-pooling operation. Max pooling is used to reduce the spatial dimensions of the feature maps while retaining the most prominent features. This operation also helps reduce the computational complexity of the subsequent layers.

2.4.4 Fully Connected and Output Layers

After the convolutional and pooling operations, the resulting feature maps are flattened into a one-dimensional vector using a flattening layer. The flattened features are then passed to a fully connected layer containing 128 neurons with ReLU activation. A dropout layer with a rate of 0.5 is applied to reduce the risk of overfitting. Finally, the output layer contains four neurons corresponding to the four waste categories and uses the softmax activation function to generate class probabilities.

2.4.5 Trainable Parameters

Total params: 11,169,476 (42.61 MB)

Trainable params: 11,169,476 (42.61 MB)

Non-trainable params: 0 (0.00 B)

The proposed CNN contains 11,169,476 total parameters, all of which are trainable. No non-trainable parameters are present because the proposed CNN is trained from scratch.

2.4.6 Tabel Arsitektur

Table 4. CNN Architecture

Layer	Configuration	Activation	Function
Input	$224 \times 224 \times 3$	–	Input image
Conv2D-1	32 filters, 3×3	ReLU	Feature extraction
MaxPooling-1	2×2	–	Downsampling
Conv2D-2	64 filters, 3×3	ReLU	Feature extraction
MaxPooling-2	2×2	–	Downsampling
Conv2D-3	128 filters, 3×3	ReLU	Feature extraction
MaxPooling-3	2×2	–	Downsampling
Flatten	–	–	Feature vector conversion
Dense	128 neurons	ReLU	Feature learning
Dropout	0.5	–	Regularization

Layer	Configuration	Activation	Function
Output	4 neurons	Softmax	Classification

The proposed architecture contains three convolutional layers, three max-pooling layers, one fully connected layer, one dropout layer, and one output layer. The final classification layer consists of four neurons corresponding to plastic, paper, glass, and metal waste.

2.5 Model Evaluation

The performance of the CNN model was evaluated using several classification metrics, namely accuracy, precision, recall, and F1-score. These metrics were calculated based on the confusion matrix generated during testing. Accuracy was used to measure the overall correctness of the classification model and is defined as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Precision measures the proportion of correctly predicted positive instances:

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

Recall measures the ability of the model to identify all relevant instances:

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

The F1-score represents the harmonic mean of precision and recall:

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

where TP denotes True Positive, TN denotes True Negative, FP denotes False Positive, and FN denotes False Negative. In addition to these metrics, a confusion matrix was used to analyze the distribution of classification results and identify potential misclassification among waste categories. The best-performing model was selected based on the highest accuracy value while maintaining balanced precision, recall, and F1-score values.

3. Result and Discussions

3.1 Effect of Epoch on Model Performance

The first experiment was conducted to analyze the effect of the number of epochs on the performance of the CNN model. The evaluation metrics used were accuracy, precision, recall, and F1-score. The results are presented in Table 5 and figure 2. The initial evaluation using a single 80:20 train-test split produced an accuracy of 97%. However, this result should be interpreted with caution because the test set consisted of only 19 images. With such a small test set, a small number of correctly or incorrectly classified samples can substantially affect the reported accuracy. Therefore, the 97% accuracy obtained from the single split may not fully represent the model's generalization capability.

Table 5. Effect of Epoch on CNN Performance

no	Jumlah gambar	epoch	Batch size	Learning rate	akurasi	precision	recall	F1-score
1	92	1	8	1e-4	0,40	0,41	0,40	0,34
2	92	3	8	1e-4	0,67	0,76	0,67	0,68
3	92	5	8	1e-4	0,83	0,84	0,83	0,83
4	92	7	8	1e-4	0,89	0,89	0,89	0,89
5	92	9	8	1e-4	0,91	0,92	0,91	0,91
6	92	11	8	1e-4	0,91	0,92	0,91	0,91
7	92	13	8	1e-4	0,91	0,92	0,91	0,91

no	Jumlah gambar	epoch	Batch size	Learning rate	akurasi	precision	recall	F1-score
8	92	15	8	1e-4	0,95	0,96	0,95	0,95
9	92	17	8	1e-4	0,95	0,96	0,95	0,95
10	92	20	8	1e-4	0,97	0,97	0,97	0,97

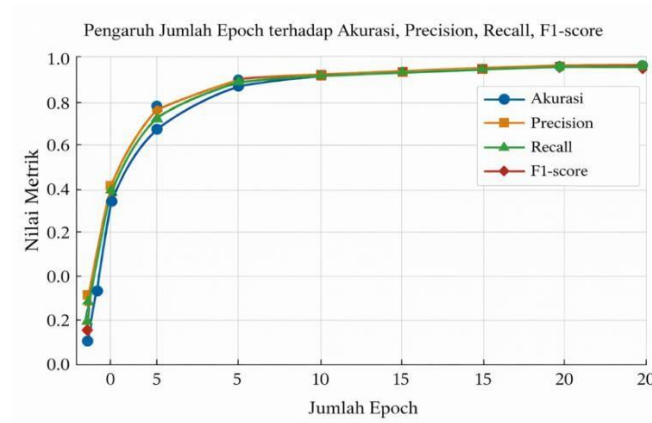


Figure 2. Graph of the Increase in Accuracy, Precision, Recall, and F1-Score with the Addition of Epochs

Table 2 shows the influence of epoch variation on CNN performance. The results indicate that increasing the number of epochs improves classification performance during the early training phase because the model gradually learns discriminative features from waste images. However, after reaching a certain number of epochs, the performance improvement becomes relatively stable, indicating that the model has reached convergence. The best performance was obtained at the optimal epoch configuration, where the model achieved the highest accuracy, precision, recall, and F1-score. Nevertheless, excessive training epochs may increase the possibility of overfitting, especially considering the limited dataset size used in this study.

3.2 Effect of Batch Size on Model Performance

The second experiment evaluated the impact of batch size on CNN classification performance. Various batch sizes ranging from 2 to 20 were tested while maintaining the other parameters constant. The initial evaluation using a single 80:20 train-test split produced an accuracy of 97%. However, this result should be interpreted with caution because the test set consisted of only 19 images. With such a small test set, a small number of correctly or incorrectly classified samples can substantially affect the reported accuracy. Therefore, the 97% accuracy obtained from the single split may not fully represent the model's generalization capability.

Table 6. Effect of Batch Size on CNN Performance

no	Jumlah gambar	epoch	Batch size	Learning rate	akurasi	precision	recall	F1-score
1	92	25	2	1e-4	0,91	0,91	0,91	0,91
2	92	25	4	1e-4	0,95	0,95	0,95	0,95
3	92	25	6	1e-4	0,95	0,96	0,95	0,95
4	92	25	8	1e-4	0,97	0,97	0,97	0,97
5	92	25	10	1e-4	0,95	0,95	0,95	0,95
6	92	25	12	1e-4	0,95	0,95	0,95	0,95
7	92	25	14	1e-4	0,96	0,97	0,96	0,96
8	92	25	16	1e-4	0,95	0,96	0,95	0,95
9	92	25	18	1e-4	0,91	0,92	0,91	0,91
10	92	25	20	1e-4	0,94	0,94	0,94	0,94

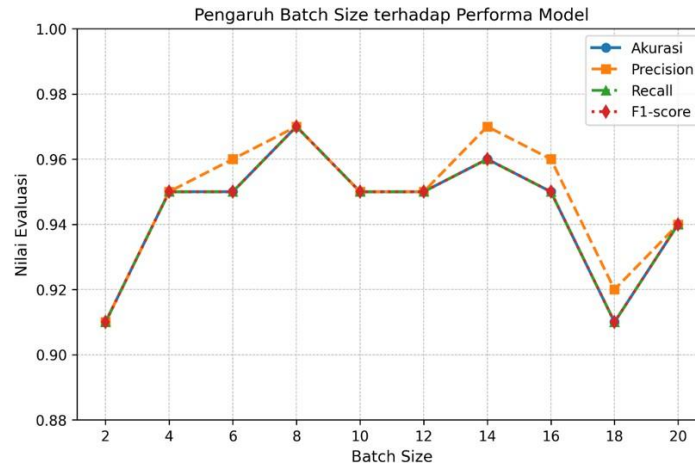


Figure 3. Graph of Accuracy, Precision, Recall, and F1-Score Evaluation Based on Batch Size Variations

The batch size experiment demonstrates that different batch configurations influence the stability and accuracy of the CNN training process. A smaller batch size allows the model to update weights more frequently, which may improve feature learning on a limited dataset. However, extremely small batch sizes can increase training instability. Based on the experimental results, the selected batch size achieved the highest evaluation performance. This configuration provided a suitable balance between learning stability and computational efficiency.

3.3 Effect of Learning Rate on Model Performance

The third experiment investigated the influence of learning rate on CNN training performance. Several learning rate values ranging from 1×10^{-1} to 1×10^{-20} were evaluated.

Table 7. Effect of Learning Rate on CNN Performance

no	Jumlah gambar	epoch	Batch size	Learning rate	akurasi	precision	recall	F1-score
1	92	25	20	1e-1	0,35	0,34	0,35	0,27
2	92	25	20	1e-3	0,90	0,92	0,90	0,90
3	92	25	20	1e-5	0,73	0,74	0,73	0,73
4	92	25	20	1e-7	0,32	0,32	0,32	0,28
5	92	25	20	1e-9	0,31	0,31	0,31	0,27

Preliminary experiments with learning rates below $1e-5$ ($1e-7$ to $1e-20$) consistently resulted in accuracy of approximately 0.31–0.32, which is close to the random-guessing baseline for the four-class classification task. Therefore, these values were excluded from the main analysis because they did not provide meaningful performance variation and indicated insufficient convergence within the given training budget.

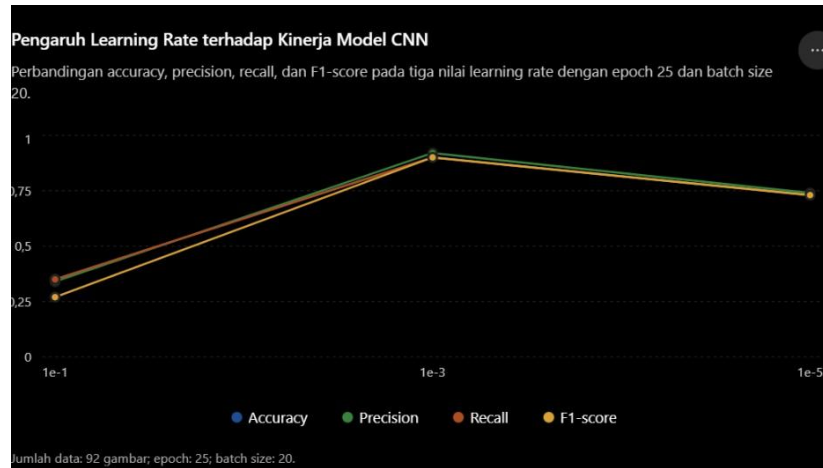


Figure 4. Evaluation Graph of Accuracy, Precision, Recall, and F1-Score at Various Learning Rate Values

The learning rate experiment shows that optimization performance is highly dependent on the selected learning rate value. A high learning rate may cause unstable updates and prevent the model from reaching optimal convergence, while an excessively small learning rate may slow the learning process and result in insufficient feature optimization. The selected learning rate provided the best balance between convergence speed and classification performance. These results indicate that proper hyperparameter selection is essential, particularly when training CNN models using limited image datasets. The learning rate search was restricted to the range of 1e-1 to 1e-5, consistent with commonly reported effective ranges for small CNNs trained with Adam-family optimizers. Values below 1e-5 were excluded after preliminary tests confirmed that they prevented meaningful convergence within the training budget.

3.4 Results of Stratified 5-Fold Cross-Validation

Table 8. Stratified 5-Fold Cross-Validation Results

Fold	Training	Validation	Accuracy	Precision	Recall	F1-Score
Fold 1	73	19	0.368421	0.502924	0.368421	0.345990
Fold 2	73	19	0.210526	0.257519	0.210526	0.210223
Fold 3	74	18	0.333333	0.273810	0.333333	0.290965
Fold 4	74	18	0.333333	0.266667	0.333333	0.293827
Fold 5	74	18	0.388889	0.375000	0.388889	0.367725
Mean	—	—	0.326901	0.335184	0.326901	0.301746
Std. Dev.	—	—	0.069268	0.105137	0.069268	0.060964

Table 8 presents the results of the Stratified 5-Fold Cross-Validation performed on the dataset consisting of 92 images. The dataset was divided into five folds while maintaining the class distribution in each fold. In each iteration, one fold was used as the validation set, while the remaining four folds were used for training. The number of training samples ranged from 73 to 74 images, while the validation samples ranged from 18 to 19 images. The results show that the model performance varied across the five folds.

Fold 1 achieved an accuracy of 36.84%, precision of 50.29%, recall of 36.84%, and F1-score of 34.60%. Fold 2 produced the lowest performance, with an accuracy of 21.05%, precision of 25.75%, recall of 21.05%, and F1-score of 21.02%. Fold 3 and Fold 4 obtained an accuracy of 33.33%, with F1-scores of 29.10% and 29.38%, respectively. Fold 5 achieved the highest accuracy of 38.89%, with a precision of 37.50%, recall of 38.89%, and F1-score of 36.77%.

Overall, the model obtained a mean accuracy of 32.69% $\pm$ 6.93%, precision of 33.52% $\pm$ 10.51%, recall of 32.69% $\pm$ 6.93%, and F1-score of 30.17% $\pm$ 6.10%. The variation in performance across folds indicates that the model's predictive performance is sensitive to the composition of the training and validation subsets. This variation is particularly relevant given the limited dataset size of only 92 images.

For a four-class classification problem, the theoretical accuracy of random guessing is approximately 25%. Therefore, the mean accuracy of 32.69% indicates that the model performs above the random baseline, but its generalization performance remains limited. These findings suggest that the small dataset size may restrict the model's ability to learn representative features and generalize consistently to unseen samples. Consequently, a larger and more diverse dataset, together with further evaluation using transfer learning architectures, is recommended for future research.

To obtain a more reliable estimate of model generalization, stratified 5-fold cross-validation was subsequently performed. In contrast to the single 80:20 split, cross-validation evaluates the model across multiple training and validation subsets. The proposed CNN achieved a mean accuracy of $32.69\% \pm 6.93\%$ across the five folds.

The substantially lower performance compared with the 97% accuracy obtained from the single split indicates that the earlier result may have been influenced by the limited size and composition of the test set, suggesting possible overfitting.

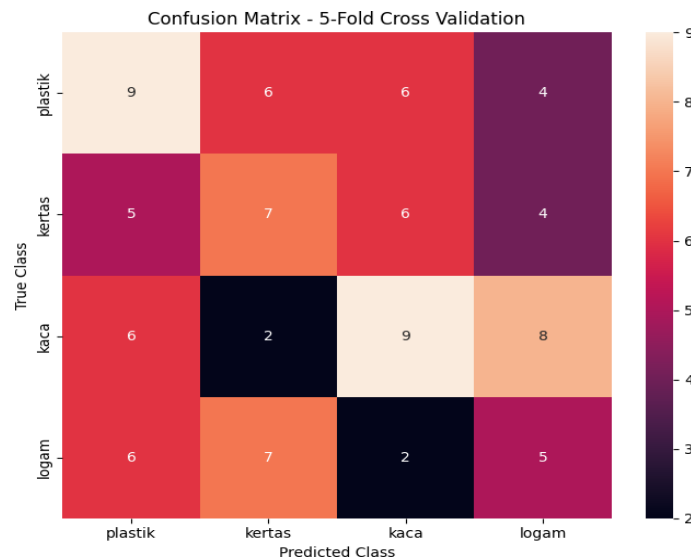


Figure 5. Aggregated Confusion Matrix from 5-Fold Cross-Validation

Figure 5 presents the aggregated confusion matrix obtained from the Stratified 5-Fold Cross-Validation for the four waste categories: plastic, paper, glass, and metal. The rows represent the actual classes, while the columns represent the predicted classes. The diagonal values indicate correctly classified samples, whereas the off-diagonal values represent misclassified samples. The model correctly classified 9 plastic images, 7 paper images, 9 glass images, and 5 metal images. Therefore, 30 out of 92 validation samples were correctly classified, corresponding to an overall accuracy of approximately 32.61%.

The relatively low number of correct predictions indicates that the model still has difficulty distinguishing among the four waste categories. For the plastic class, 9 samples were correctly classified, while 6 samples were incorrectly predicted as paper, 6 as glass, and 4 as metal. For the paper class, 7 samples were correctly classified, whereas 5 were predicted as plastic, 6 as glass, and 4 as metal. The glass class achieved 9 correct predictions; however, 6 samples were classified as plastic, 2 as paper, and 8 as metal. The metal class obtained 5 correct predictions, while 6 samples were predicted as plastic, 7 as paper, and 2 as glass.

The largest misclassification occurred between the glass and metal classes, where 8 glass samples were predicted as metal. In addition, plastic was frequently misclassified as paper and glass, with 6 samples for each category. These results indicate that the visual characteristics of several waste categories may overlap, making classification difficult for a CNN trained with a limited number of images.

The class-wise results also show relatively low recognition rates. Based on the confusion matrix, the recall values are approximately 36.00% for plastic, 31.82% for paper, 36.00% for glass, and 25.00% for metal. Metal obtained the lowest recall, indicating that this class was the most difficult for the model to identify correctly.

These findings are consistent with the cross-validation results presented in Table 8, where the mean accuracy was 32.69% and the mean F1-score was 30.17%.

Overall, the confusion matrix demonstrates that the model has limited generalization capability across the four waste categories. The results are likely influenced by the small dataset size of 92 images, which limits the diversity of visual characteristics available during training. Therefore, increasing the dataset size, improving data diversity, and evaluating transfer learning architectures may help improve the classification performance in future studies.

3.5 Confirmation Experiment Using the Selected Hyperparameter Configuration

Based on the sequential one-factor-at-a-time (OFAT) tuning experiments, the best-performing values for each hyperparameter were identified individually. The selected configuration consisted of 20 epochs, a batch size of 8, and a learning rate of 1×10^{-3} . A confirmation experiment was then conducted by combining these selected values into a single configuration and retraining the CNN model. The purpose of this experiment was to verify the performance of the final configuration obtained from the sequential OFAT tuning process. The results of the confirmation experiment are presented. The configuration of 20 epochs, batch size of 8, and learning rate of 1×10^{-3} achieved the best performance among the configurations evaluated during the sequential tuning process. However, because the hyperparameters were tuned sequentially rather than simultaneously, this configuration should not be interpreted as the globally optimal combination across all possible hyperparameter settings.

3.6 Discussion

This study employed a sequential one-factor-at-a-time (OFAT) tuning strategy rather than a joint grid search; therefore, the reported configuration should be interpreted as a locally optimal rather than globally optimal setting. A confirmation experiment was then conducted by combining these selected values into a single configuration and retraining the CNN model. The purpose of this experiment was to verify the performance of the final configuration obtained from the sequential OFAT tuning process. The results of the confirmation experiment. The configuration of 20 epochs, batch size of 8, and learning rate of 1×10^{-3} achieved the best performance among the configurations evaluated during the sequential tuning process. However, because the hyperparameters were tuned sequentially rather than simultaneously, this configuration should not be interpreted as the globally optimal combination across all possible hyperparameter settings.

The obtained accuracy of 97% indicates that the proposed CNN model achieved promising performance for household waste image classification. However, direct comparison with previous studies should be interpreted carefully because differences exist in dataset size, number of waste categories, image characteristics, and experimental settings. Although the proposed model achieved a high accuracy value, the evaluation dataset used in this study is relatively limited compared with several previous studies [7],[5]. The high performance may be influenced by the controlled image acquisition environment and the characteristics of the collected household waste images [15],[19]. Therefore, the obtained accuracy should be considered as an initial evaluation of model feasibility rather than evidence that the proposed model universally outperforms previous approaches.

Nevertheless, the results demonstrate that CNN can effectively extract visual features from household waste images and provide accurate classification results for the tested categories. Future studies should evaluate the model using larger and more diverse datasets containing variations in lighting conditions, backgrounds, camera devices, and real-world household environments.

The selection of a CNN architecture trained from scratch in this study was based on the objective of developing a lightweight and interpretable model architecture specifically designed for household waste classification. The proposed CNN was developed using a relatively simple architecture to evaluate the feasibility of extracting waste-related visual features from the collected household waste images. However, we acknowledge that transfer learning approaches, including MobileNet, ResNet, EfficientNet, and Vision Transformer (ViT), are

widely adopted for small-scale image classification tasks because pre-trained models can utilize previously learned feature representations from large-scale datasets such as ImageNet.

These approaches may provide better generalization, particularly when the available training dataset is limited. Therefore, the obtained accuracy of 97% should not be interpreted as evidence that the proposed CNN architecture outperforms transfer learning-based methods. Instead, the result demonstrates the feasibility of applying a custom CNN model for the evaluated household waste dataset. A comprehensive comparison with transfer learning architectures will be conducted in future research to determine the optimal model for household waste classification.

The Stratified 5-Fold Cross-Validation results indicate that the proposed CNN achieved an average accuracy with a standard deviation. The relatively small standard deviation indicates that the model performance remained relatively consistent across the five folds. However, the limited dataset size of 92 images remains a constraint and may affect the generalization capability of the model.

The substantial difference between the accuracy obtained from the single train-test split and the stratified 5-fold cross-validation requires careful consideration. The proposed CNN achieved 97% accuracy under the single 80:20 split, whereas its mean accuracy decreased to $32.69\% \pm 6.93\%$ under stratified 5-fold cross-validation. The model also obtained an accuracy of 33.33% in the transfer learning comparison. These differences are primarily associated with the limited dataset size, which consists of only 92 original images.

In the single-split experiment, only 19 images were used for testing, making the accuracy highly sensitive to the specific composition of the test set. Consequently, the 97% accuracy may represent an optimistic estimate of model performance and may be affected by overfitting to the training data. In contrast, stratified 5-fold cross-validation evaluates the model across multiple subsets and therefore provides a more robust indication of its generalization capability.

The substantially lower cross-validation performance suggests that the model has difficulty generalizing to unseen samples. Therefore, the 97% accuracy obtained from the single-split experiment should not be considered the primary indicator of the proposed model's performance. The cross-validation results are considered more appropriate for assessing generalization given the limited dataset size.

3.7 Comparison with Transfer Learning Models

Table 9. Performance Comparison of CNN and Transfer Learning Models

Model	Accuracy	Precision	Recall	F1-score
Proposed CNN	0.333333	0.189899	0.333333	0.237654
MobileNetV2	0.777778	0.777778	0.777778	0.773449
ResNet50	0.777778	0.782407	0.777778	0.767725
EfficientNetB0	0.888889	0.909259	0.888889	0.887446

The performance comparison between the proposed CNN and the transfer learning models is presented in Table 9. The proposed CNN achieved an accuracy of 33.33%, while MobileNetV2, ResNet50, and EfficientNetB0 achieved accuracies of 77.78%, 77.78%, and 88.89%, respectively. These results indicate that EfficientNetB0 achieved the best classification performance among the evaluated models, whereas the proposed CNN obtained the lowest accuracy. This finding suggests that pretrained transfer learning architectures can provide more effective feature representations and better generalization than a CNN trained from scratch when using a limited dataset. The results in Table 9 should also be interpreted in the context of the evaluation protocol. The proposed CNN achieved 33.33% accuracy in the transfer learning comparison, which is substantially lower than the 97% accuracy obtained using the initial single 80:20 train-test split. This difference indicates that the performance of the proposed CNN is highly sensitive to dataset partitioning and the evaluation procedure. Therefore, the results obtained from different evaluation protocols should not be directly compared unless the same dataset partitioning and evaluation procedure are applied. For a fair comparison, all evaluated

architectures were assessed using the same dataset split and evaluation procedure.

3.8 Overfitting Analysis

3.8.1 Training accuracy vs validation accuracy

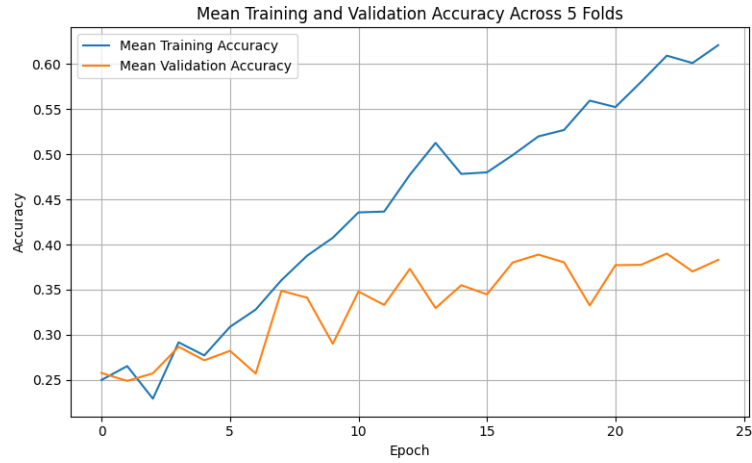


Figure 6. Mean Training and Validation Accuracy Across 5 Folds

The training and validation accuracy curves were analyzed to evaluate the learning behavior and generalization capability of the proposed CNN model. Training accuracy represents the classification performance of the model on the training data, whereas validation accuracy represents its performance on previously unseen validation data. As shown in Figure 6, the training accuracy generally increases as the number of epochs increases, indicating that the model progressively learns the visual features of the waste categories. The validation accuracy also shows an increasing trend, although fluctuations between epochs may occur due to the relatively small dataset used in this study. The difference between training and validation accuracy is used as an indication of the generalization gap. A relatively small gap between the two curves indicates that the model can generalize its learned features to validation data. Conversely, if the training accuracy continues to increase while the validation accuracy stagnates or decreases, this indicates a potential overfitting problem. Therefore, the training and validation accuracy curves provide an important basis for evaluating whether the model learns general patterns rather than memorizing the training images.

3.8.2 Training Loss VS Validation Loss

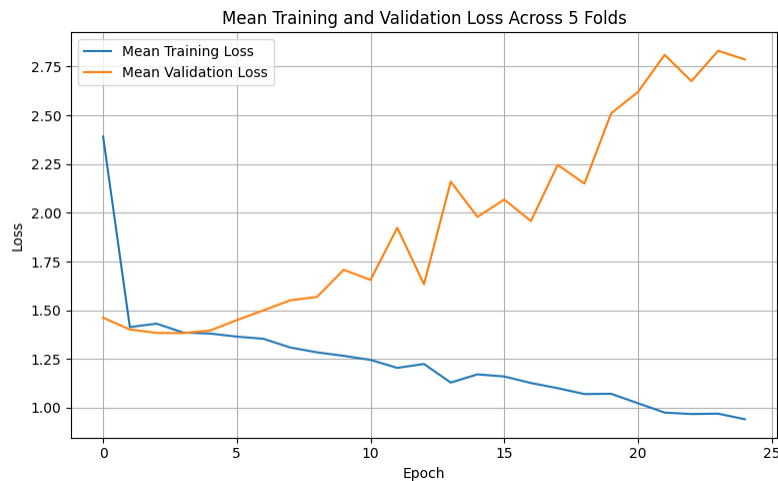


Figure 7. Mean Training and Validation Loss Across 5 Folds

The training and validation loss curves were further examined to identify potential overfitting during the model training process. Training loss measures the error produced by the model on the training data, while validation loss measures the error on the validation data. As illustrated in Figure 7, the training loss generally decreases throughout the training process, indicating that the model progressively minimizes the classification error on the training data. The validation loss is also evaluated to determine whether the reduction in training error is accompanied by improved generalization to unseen data. When both training and validation loss decrease simultaneously, the model demonstrates a stable learning process. However, an increasing validation loss while training loss continues to decrease indicates that the model may be learning specific characteristics of the training data rather than generalizable features. This condition is commonly interpreted as an indication of overfitting. The loss curves are therefore considered together with the accuracy curves to provide a more reliable assessment of the model's generalization capability.

3.8.3 Overfitting Analysis

The training and validation curves indicate a potential overfitting tendency in the later training epochs. The training accuracy continues to increase while the improvement in validation accuracy becomes limited. At the same time, the training loss continues to decrease, whereas the validation loss tends to increase or fluctuate at a higher level. This behavior suggests that the CNN model increasingly fits the training samples but does not obtain a proportional improvement in its ability to generalize to unseen validation samples. The observed generalization gap is particularly relevant because the dataset contains a relatively limited number of images. Therefore, the high training performance should not be interpreted solely as evidence of strong generalization. The use of five-fold cross-validation provides additional evaluation across different training and validation subsets. Nevertheless, the observed differences between training and validation performance indicate that the model should be interpreted with consideration of the limited dataset size.

3.9 Performance Metrics Comparison

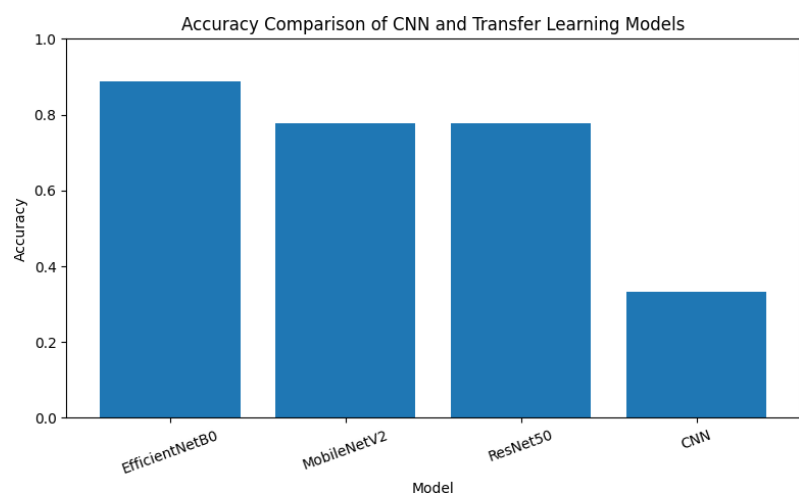


Figure 8. Accuracy Comparison of CNN and Transfer Learning Models

Figure 8 illustrates the accuracy performance comparison across all evaluated models. The results demonstrate that EfficientNetB0 achieved the highest accuracy at approximately 89% (0.89), outperforming all other architectures. Both MobileNetV2 and ResNet50 exhibited moderate and identical accuracy levels of around 78% (0.78). In contrast, the baseline CNN model registered the lowest accuracy, reaching only 33% (0.33). This significant performance gap highlights the strong capability of transfer learning models in correctly classifying overall instances compared to a standard, non-pretrained CNN.

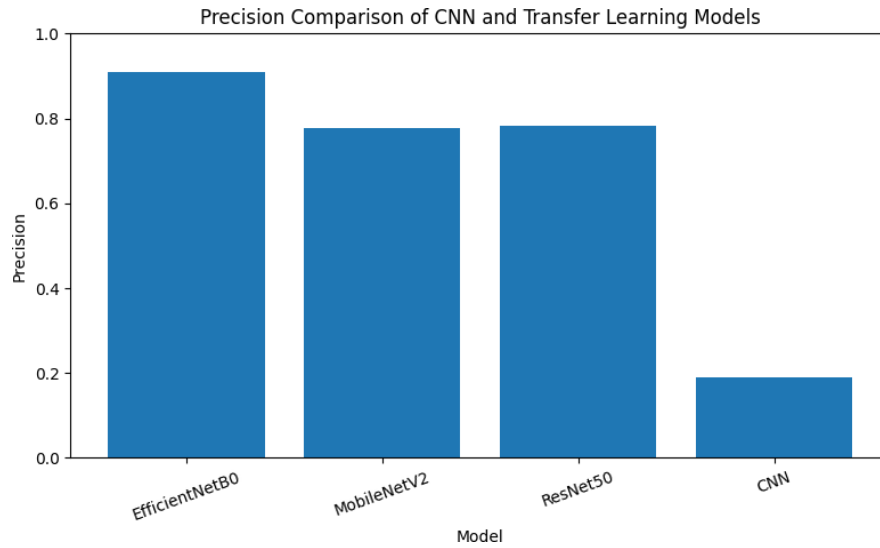


Figure 9. Precision Comparison of CNN and Transfer Learning Models

Figure 9 highlights the precision performance of the models, measuring the exactness of positive predictions. EfficientNetB0 leads with the highest precision score of approximately 91% (0.91), indicating a low rate of false-positive classifications. MobileNetV2 and ResNet50 showed comparable precision scores, reaching around 78% (0.78). Meanwhile, the standard CNN model performed poorly, obtaining a precision score of less than 20% (~0.19). These results confirm that transfer learning significantly minimizes misclassification errors in positive predictions.

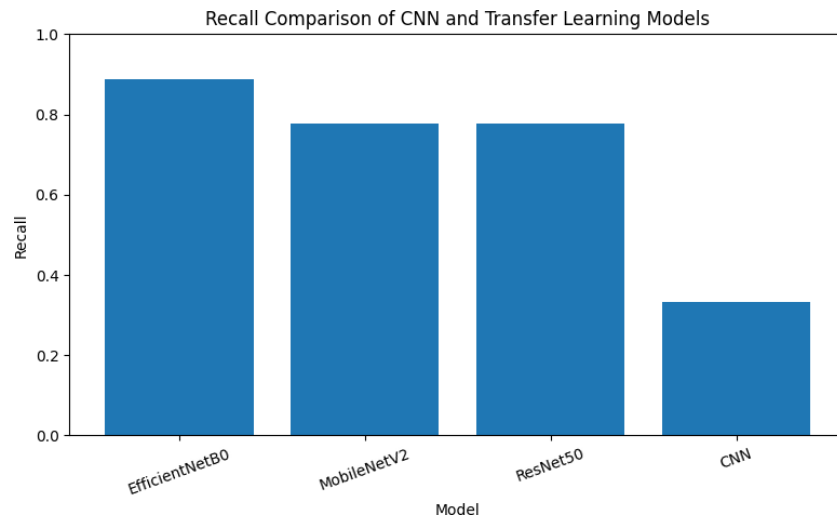


Figure 10. Recall Comparison of CNN and Transfer Learning Models

Figure 10 displays the recall (sensitivity) comparison, which measures each model's ability to identify all relevant positive instances. EfficientNetB0 achieved the top recall score at approximately 89% (0.89). Both MobileNetV2 and ResNet50 followed closely with identical recall values of 78% (0.78). Conversely, the standard CNN architecture struggled significantly, achieving a recall of only 33% (0.33). The high recall achieved by EfficientNetB0 confirms its effectiveness in minimizing false negatives during detection.

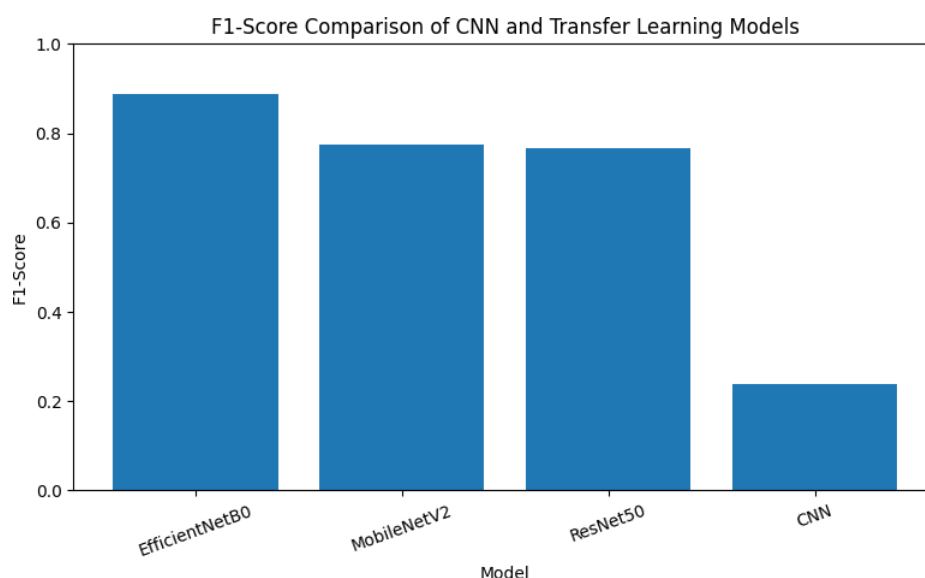


Figure 11. F1-Score Comparison of CNN and Transfer Learning Models

Figure 11 presents the F1-Score comparison, representing the harmonic mean of precision and recall to evaluate overall balanced performance. EfficientNetB0 attained the highest F1-Score at approximately 89% (0.89), demonstrating superior stability and balance between precision and sensitivity. MobileNetV2 and ResNet50 achieved nearly identical scores of 78% (0.78) and 77% (0.77), respectively. Finally, the baseline CNN model recorded a low F1-Score of 24% (0.24), further proving that transfer learning is far more robust for this task.

4. Conclusions and Future Works

This research successfully developed a CNN-based household waste image classification model capable of automatically identifying different waste categories from image data. The developed model achieved satisfactory classification performance based on accuracy, precision, recall, and F1-score evaluation. Furthermore, the trained model was integrated into a simple web-based application that enables users to upload waste images and obtain automatic classification results. However, the current implementation is still limited to image-based identification and has not yet supported real-time camera detection or deployment on edge devices. Future research can focus on developing real-time waste detection systems using mobile devices or embedded platforms.

The proposed system was designed to classify four categories of inorganic waste, namely plastic, paper, glass, and metal, using digital image data. Experimental results demonstrated that CNN is highly effective in recognizing and classifying waste objects based on their visual characteristics, including shape, texture, and color. The implementation of preprocessing techniques such as image resizing, normalization, and data augmentation contributed significantly to improving data quality and enhancing model performance during the training process.

The evaluation results revealed that training parameters play a crucial role in determining the effectiveness of the CNN model. Among the tested configurations, the optimal performance was achieved using 20 epochs, a batch size of 8, and a learning rate of 1×10^{-3} . Under these conditions, the model obtained a classification accuracy of 97%, accompanied by high and balanced precision, recall, and F1-score values. The experimental results demonstrate that the proposed CNN can perform waste image classification; however, its generalization capability is limited by the small dataset size. Although an accuracy of 97% was obtained using a single 80:20 train-test split, stratified 5-fold cross-validation resulted in a mean accuracy of $32.69\% \pm 6.93\%$. The substantial difference between these results suggests that the single-split accuracy may provide an optimistic estimate of model performance and may be affected by overfitting. Therefore, the cross-

validation result provides a more reliable indication of the model's generalization capability. These findings highlight the importance of using robust validation strategies when developing CNN-based classification models with limited datasets.

Future studies should increase the number and diversity of real-world waste images and investigate more robust architectures and regularization techniques to improve generalization performance. These findings indicate that proper hyperparameter optimization can significantly improve model convergence, reduce classification errors, and enhance the overall reliability of the waste identification system.

Furthermore, the developed system has the potential to support household waste management by assisting users in sorting waste automatically and accurately. The ability to identify waste categories in real time can contribute to increasing public awareness of waste segregation practices, promoting recycling activities, and reducing environmental pollution. Therefore, the proposed system can serve as an intelligent technological solution for supporting sustainable waste management in household environments.

Despite the promising results, this study has several limitations. The dataset used consisted of only 92 images and was limited to four categories of inorganic waste. Consequently, the model may not fully represent the diversity of waste objects encountered in real-world environments. Future research should consider utilizing larger and more diverse datasets, including organic waste and residual waste categories, to improve model generalization capabilities.

In addition, more advanced Deep Learning architectures such as MobileNet, ResNet, EfficientNet, or Vision Transformer (ViT) can be explored and compared to achieve higher classification performance. Future studies may also focus on integrating the model into web-based or mobile applications, enabling direct implementation by households and supporting the development of smart waste management systems in real-world environments.

Based on Stratified 5-Fold Cross-Validation, the proposed CNN achieved an average accuracy of $32.69\% \pm 6.93\%$ with a precision of $33.52\% \pm 10.51\%$, recall of $32.69\% \pm 6.93\%$, and F1-score of $30.17\% \pm 6.10\%$.

5. Declarations

Author Contributions

Author 1 Responsible for conceptualization of the research, formulation of the research methodology, dataset collection and preparation, image preprocessing, CNN model development, model training, hyperparameter configuration, model evaluation, analysis and interpretation of experimental results, and preparation of the initial manuscript. Author 2 Responsible for supporting the collection and organization of the household waste image dataset, data labeling and classification, dataset verification, preprocessing procedures, and assisting in the implementation and testing of the CNN model. Author 3: Responsible for assisting in the development and implementation of the CNN model, conducting model training and evaluation, preparing performance evaluation results, including accuracy, precision, recall, F1-score, and confusion matrix, and assisting in the analysis of experimental results. Author 4: Responsible for reviewing and validating the research methodology and experimental results, providing critical revisions to the manuscript, improving the presentation and interpretation of the findings, and supervising the overall preparation of the manuscript for publication.

Funding

The authors received no specific funding for this research.

Conflicts of Interest

The authors declare that there are no conflicts of interest related to the publication of this article.

Data Availability Statement

The dataset used in this study consists of household waste images collected from primary and secondary sources. The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Use of Artificial Intelligence (AI)

Artificial Intelligence (AI) was used as a supporting tool during the preparation of this manuscript, particularly to assist with grammar correction, sentence structure improvement, academic language refinement, translation, and manuscript organization. AI was also used to provide suggestions for improving the clarity, coherence, and readability of several sections of the manuscript. AI was not used to generate research data, perform image preprocessing, develop or train the Convolutional Neural Network (CNN) model, conduct model evaluation, or determine the research results and conclusions. All research activities, including data collection and preparation, image preprocessing, CNN model development, model training, evaluation, analysis of experimental results, and interpretation of findings, were conducted and verified by the authors. All AI-generated suggestions were carefully reviewed revised, and validated by the authors to ensure their accuracy and consistency with the research data, methodology, and findings. The authors remain fully responsible for the accuracy, originality, and final content of the manuscript.

Ethical Clearance

This study did not involve human participants, animals, or personal data requiring ethical approval. Therefore, ethics approval was not required.

6. References

- [1] SIPSN, "Regulasi SIPSN," *Sipsn*, vol. 6, no. 2, pp. 162–171, 2025, [Online]. Available: <https://sipsn.menlhk.go.id/sipsn>
- [2] Badan Pusat Statistik, "Metadata Statistik Kegiatan Sakernas," pp. 1–19, 2021, [Online]. Available: <https://sirusa.web.bps.go.id/metadata/>
- [3] A. Ikhlas and B. Hendrik, "Literature Review: A Comparative Study of Waste Classification using Deep Learning Algorithms," *Sistemasi*, vol. 14, no. 3, p. 1360, 2025, doi: 10.32520/stmsi.v14i3.5163.
- [4] C. J. Yi and C. F. Kim, "AI-Powered Waste Classification Using Convolutional Neural Networks (CNNs)," *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 10, pp. 67–75, 2024, doi: 10.14569/IJACSA.2024.0151009.
- [5] M. E. Purba, A. Z. Situmorang, G. L. Br Ginting, M. W. P. Lubis, and F. M. Sinaga, "Klasifikasi Sampah Organik dan Anorganik Menggunakan Algoritma CNN," *J. Sifo Mikroskil*, vol. 26, no. 1, pp. 37–54, 2025, doi: 10.55601/jsm.v26i1.1510.
- [6] A. Asmawati, Y. Kurniawan, and R. N. Fadilah, "EFISIENSI PEMILAHAN SAMPAH DAUR ULANG BERBASIS ANDROID MENGGUNAKAN MODEL CNN SSD-MOBILENET V2," *Journal*, vol. 11, no. 2, pp. 2654–8704, 2025.
- [7] M. Diqi, "Waste Classification using CNN Algorithm," *Int. Conf. Inf. Sci. Technol. Innov.*, vol. 1, no. 1, pp. 130–135, 2022, doi: 10.35842/icostec.v1i1.17.
- [8] G. A. Bahagia and M. Akbar, "Klasifikasi Sampah Organik Dan Anorganik Menggunakan Metode Convolutional Neural Network (Cnn)," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 5, pp. 10349–10355, 2024, doi: 10.36040/jati.v8i5.11038.
- [9] D. S. Aulia, H. Arwoko, and E. Asmawati, "Klasifikasi Sampah Rumah Tangga Menggunakan Metode Convolutional Neural Network," *Metik J.*, vol. 8, no. 2, pp. 114–120, 2024, doi: 10.47002/metik.v8i2.956.
- [10] S. Agustiani, H. Haryani, A. Junaidi, R. R. Putri, and M. G. Adam Z, "Comparative Optimization of EfficientNetB3, MobileNetV2, and ResNet50 for Waste Classification," *J. Inform.*, vol. 12, no. 2, pp. 131–137, 2025, doi: 10.31294/inf.v12i2.27533.
- [11] T. N. Sari, M. P. Sari, T. A. Putri, and R. Pashya, "Otomatisasi Klasifikasi Sampah Menggunakan Convolutional Neural Network (CNN) Sebagai Sistem Pendukung Keputusan," in *Pros. SEMNAS INOTEK (Seminar Nas. Inov. Teknol.)*, 2025, pp. 720–729. [Online]. Available: <https://proceeding.unpkediri.ac.id/index.php/inotek/article/view/7413/4935>

- [12] L. Rahmawati and S. Defianti, "Klasifikasi Jenis Sampah Menggunakan Convolutional Neural Network untuk Mendukung Pengelolaan Limbah Waste Type Classification Using Convolutional Neural Network for Supporting Waste Management," vol. 6, no. 1, pp. 61–68, 2026, doi: 10.59395/exaxyj37.
- [13] M. Haqqi, L. Rochmah, A. D. Safitri, R. A. Pratama, and Tarwoto, "Implementation Of Machine Learning To Identify Types Of Waste Using CNN Algorithm," *J. Fasilkom*, vol. 14, no. 3, pp. 761–765, 2024, doi: 10.37859/jf.v14i3.8116.
- [14] 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, 2024, doi: 10.1007/s11042-024-18939-w.
- [15] J. J. C. Simbolon, Robet, and Hendri, "Household Waste Image Classification Using Deep Learning Model," *J. Artif. Intell. Eng. Appl.*, vol. 5, no. 1, pp. 1681–1690, 2025, doi: 10.59934/jaiea.v5i1.1690.
- [16] A.-M. Al-Mamun, R. Hossain, M. M. M. Sharmin, E. Kabir, and M. A. Iqbal, "Garbage classification using convolutional neural networks (CNNs)," *Mater. Sci. Eng. Int. J.*, vol. 7, no. 3, pp. 140–144, 2023, doi: 10.15406/mseij.2023.07.00217.
- [17] M. F. Setiawan, A. H. Rismawyana, P. T. Informatika, C. Utara, and K. Cimahi, "Convolutional Neural Network (Cnn) Pada Klasifikasi Grade," vol. 13, no. 3, 2024, [Online]. Available: <https://doi.org/10.23960/jitet.v13i3S1.7805>
- [18] Kartiko, A. P. Yudha, N. D. Aryanto, and M. A. Farabi, "Klasifikasi Sampah Menggunakan Convolutional Neural Network," *Indones. J. Data Sci.*, vol. 3, no. 2, pp. 72–81, 2022, doi: 10.56705/ijodas.v3i2.33.
- [19] S. P. Luntungan, "Development of Automatic Waste Classification System using CNN-Based Deep Learning to Support Smart Waste Management," *INOVTEK Polbeng - Seri Inform.*, vol. 10, no. 1, pp. 214–224, 2025, doi: 10.35314/wst8mh87.