JASMINE
Journal of Intelligent Systems and Machine Learning
https://journals.telkomuniversity.ac.id/jasmine

Article in Press

Transfer Learning for Medical Waste Image Classification Using EfficientNet-B0

Reni Kartika Suwandi[1*], **Asep Saeppani**[1], **and Irfan Fadil**[1]
[1]Information Technology Faculty, Universitas Sebelas April, Sumedang, Indonesia

*Author to whom any correspondence should be addressed.
E-mail: 220660121130@student.unsap.ac.id

Received: December 9, 2025
Accepted for publication: February 12, 2026
Published:

ABSTRACT

Medical waste management requires careful handling due to its potential to cause infection and environmental hazards when not properly treated. Manual classification of medical waste is time-consuming, highly dependent on human accuracy, and prone to error. This study employs a transfer learning approach using the EfficientNet-B0 architecture to automatically classify medical waste images. The dataset consists of 23 waste categories and undergoes preprocessing steps including data cleaning, resizing, normalization, and data augmentation. The model is initialized with ImageNet-pretrained weights and refined through fine-tuning. Experimental results demonstrate stable classification performance, achieving an average accuracy of 92.2%, precision of 94.1%, recall of 92.2%, and an F1-score of 92.1%. The results indicate that EfficientNet-B0 provides a competitive balance between classification performance and computational efficiency for medical waste image classification, particularly under limited data conditions. However, this study is limited by the size and scope of the dataset and the absence of real-world deployment evaluation, which may affect the generalizability of the results. Despite these limitations, the proposed approach offers a feasible basis for further research toward automated medical waste sorting systems.

Keywords: digital image, EfficientNet-B0, image classification, medical waste, transfer learning

I. Introduction

Medical waste is classified as hazardous material that poses serious risks to public health and the environment if not properly managed [1]. The rapid growth of healthcare services, including those intensified during large-scale outbreaks such as COVID-19, has significantly increased the volume of medical waste generated worldwide [1], [2]. In many healthcare facilities, waste sorting is still performed manually, resulting in time inefficiency, dependence on human accuracy, and increased exposure to infectious materials. These challenges motivate the development of automated and intelligent systems for medical waste classification.

Recent advances in deep learning, particularly Convolutional Neural Networks (CNNs), have led to substantial improvements in image-based classification tasks. Transfer learning has become a widely adopted strategy, enabling pretrained models on large-scale datasets such as ImageNet to be adapted to specific target tasks through fine-tuning [3], [4]. This approach is especially beneficial in medical and environmental imaging, where large labeled datasets are often unavailable, as it improves performance while reducing training time and data requirements [5], [6].

Prior studies have demonstrated the feasibility of applying deep learning to medical waste image classification. Kunwar and Rai [7] evaluated transfer learning architectures aligned with bin color standards, while Akkajit and Sukkuea [8] applied CNN-based models for multiclass medical waste classification. Lightweight and efficient models have also been proposed, such as MedBin [9], alongside deeper architectures including ResNeXt-based frameworks [10]. In addition, computer vision–assisted sorting systems have been explored to support real-world medical waste workflows [11].

To provide a concise overview of existing research, Table 1 summarizes representative studies on medical waste image classification, highlighting the employed models, application contexts, and key results.

Table 1. Summary of Previous Studies on Medical Waste Image Classification.

No	Author (Year)	Method	Use Case	Key Results
1	Kunwar & Rai (2025)	Transfer learning (CNN)	Bin color–based medical waste	High classification accuracy
2	Akkajit & Sukkuea (2024)	CNN-based model	Medical waste images	Strong multiclass performance
3	Xu et al. (2025)	MedBin (lightweight CNN)	Medical waste management	Efficient end-to-end classification
4	Zhou et al. (2022)	ResNeXt-based model	Medical waste images	High accuracy across categories
5	van den Broek & Sharami (2025)	Hybrid capsule network	Medical waste images	Improved robustness
6	Mythili & Anbarasi (2022)	En-SegNet-DNN + TL	Medicinal trash images	Enhanced feature representation
7	Bruno et al. (2023)	Computer vision–assisted sorting	Real-world workflows	Practical sorting support
8	Nallapaneni et al. (2021)	AI-based sorting system	COVID-related medical waste	Data-driven waste management
9	Ardana & Kusrini (2025)	EfficientNet-B0, ResNet-50	General waste images	EfficientNet-B0 competitive
10	Kurniawan et al. (2025)	EfficientNet-B0	Waste image classification	Efficient and lightweight

As summarized in Table 1, existing studies report promising results but often rely on limited datasets and place less emphasis on balancing classification accuracy with computational efficiency and evaluation robustness. These limitations motivate the exploration of efficient transfer learning–based models that can achieve stable performance under constrained data conditions.

EfficientNet architectures have been specifically designed to balance accuracy and computational cost through compound scaling [17]. EfficientNet-B0, as the baseline variant, has demonstrated strong performance in various medical imaging tasks [18], [19] and general waste classification scenarios [20], [21], making it a suitable backbone for medical waste image classification.

Accordingly, this study applies transfer learning using EfficientNet-B0 to develop an image-based classification model for medical waste. The model is evaluated using a cross-validation strategy to improve the reliability of performance assessment [22], [23], with the aim of supporting safer and more efficient medical waste management practices.

II. Related Work

Recent research on medical waste management increasingly emphasizes the integration of data-driven and intelligent systems to address the growing volume and heterogeneity of waste streams. Healthcare waste poses significant risks to public health and the environment when handling and segregation processes remain manual and inconsistent [1]. To mitigate these risks, several studies have incorporated artificial intelligence into broader waste management frameworks, including AI-based decision support systems for COVID-19-related medical waste within circular economy contexts [18] and blockchain-assisted classification pipelines with integrated tracking mechanisms for solid waste streams [12]. These approaches highlight the importance of intelligent classification modules as core components of modern waste management systems.

Within the waste classification domain, deep convolutional neural networks (CNNs) have been widely adopted to distinguish multiple waste categories and support operational decision-making. Enhanced CNN-based frameworks for general waste classification aimed at sustainable waste handling have been reported in [13], while other studies have connected deep learning models with smart city infrastructures through sensor-based and networked systems [14]. Focusing specifically on medical waste, several studies have formulated the problem as an image recognition task. Deep learning models capable of distinguishing multiple medical waste categories have been proposed in [17], and classifiers aligned with country-specific bin color policies have been reported in [6]. Additional CNN-based approaches have demonstrated promising performance using datasets collected directly from healthcare environments [7], [10]. However, many of these datasets and evaluation protocols are not standardized or publicly accessible, limiting reproducibility and fair comparison across studies.

From a methodological perspective, transfer learning has emerged as a dominant strategy for medical image classification under limited labeled data conditions. Comprehensive reviews in [2] and [3] demonstrate that fine-tuning pretrained CNN backbones consistently improves performance across diverse medical imaging tasks when annotation resources are constrained. Further studies indicate that combining transfer learning with appropriate regularization and data augmentation strategies often outperforms training models from scratch [5]. More specialized techniques, such as modality-bridge transfer learning, have been introduced to reduce domain discrepancies between natural and medical images [4], while transfer learning–based CNN pipelines with rigorous validation protocols have been applied to automated detection tasks in ophthalmology and related medical imaging applications [11]. These findings establish transfer learning as a well-founded baseline for medical waste image classification.

EfficientNet-based architectures have attracted increasing attention due to their favorable trade-off between classification accuracy and computational efficiency. EfficientNet, introduced in [16], employs a compound scaling strategy that jointly adjusts network depth, width, and input resolution to achieve strong performance with a constrained number of parameters. EfficientNet-B0 and its variants have been validated in a range of medical imaging tasks, including eye disease classification [23] and retinal disease recognition from fundus images [24]. In the context of waste classification, EfficientNet-B0 has been evaluated both in comparison with deeper architectures such as ResNet-50 [20] and as the backbone of efficient multiclass waste classification pipelines [25]. These results suggest that EfficientNet-B0 is a suitable candidate for medical waste image classification scenarios where both predictive accuracy and inference efficiency are critical.

Data preprocessing, augmentation, and evaluation strategies also play essential roles in the effective application of transfer learning to limited datasets. Augmentation libraries such as Albumentations provide flexible and efficient image transformation techniques that are widely adopted in computer vision and medical imaging studies [15]. A systematic review in [21] highlights that well-designed augmentation policies are often crucial for improving generalization, particularly in the presence of class imbalance and label noise. For performance evaluation, cross-validation protocols are commonly recommended to obtain reliable and unbiased performance estimates. Detailed guidance on k-fold cross-validation for AI-based medical imaging is provided in [22], and earlier CNN-based studies demonstrate its use for assessing model stability under limited data conditions [11]. Despite these advances, many medical waste classification studies still rely on private or narrowly scoped datasets [17], [19], and EfficientNet-based architectures have not been systematically evaluated on multiclass medical waste datasets under rigorous cross-validation schemes.

In light of these observations, the present study applies transfer learning using EfficientNet-B0 to a publicly available 23-class medical waste image dataset, integrating a structured preprocessing and augmentation pipeline with a 5-fold cross-validation evaluation protocol. By explicitly documenting the training configuration and evaluation strategy, this work aims to provide a reproducible and comparable baseline for future research on medical waste image classification.

III. Material and Methods

A. *Research Design*

This study adopts an experimental design to develop a deep learning-based model for medical waste image classification. The research workflow is organized systematically, starting from the collection of a medical waste image dataset, image preprocessing, application of transfer learning on the EfficientNet-B0 architecture, model training, and finally performance evaluation using a 5-fold cross-validation scheme.

This design is intended to ensure that the resulting model is not only able to recognize visual patterns with high accuracy but also exhibits consistent behavior and good generalization capability in distinguishing different types of medical waste, as shown in Figure 1.

In general, the main stages of the study include: (1) collection and selection of medical waste images; (2) standardization of image quality and format through preprocessing; (3) utilization of transfer learning on EfficientNet-B0 as the basis for designing the classification model; (4) model training using the specified hyperparameter configuration; and (5) performance evaluation using several classification metrics under the 5-fold cross-validation scheme. A 5-fold cross-validation scheme is employed to obtain a more stable and reliable performance estimate, as commonly recommended in CNN-based medical imaging studies

[11], [22]. With this design, the study aims to produce a medical waste image classification model that is accurate, stable, and generalizable across different data variations.

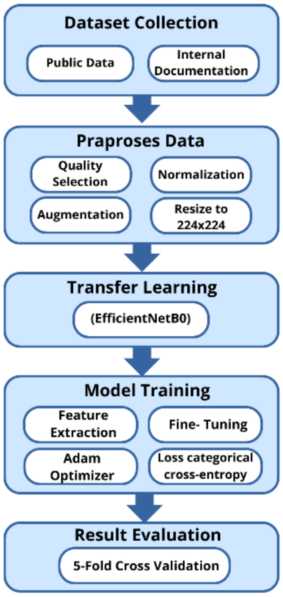

Figure 1. Research flowchart

B. *Dataset*

This study uses a digital medical waste image dataset obtained from the publicly available Medical Waste dataset on the Kaggle platform. The dataset consists of 9,634 images distributed across 23 categories of medical waste. Each category represents waste objects that are commonly found in healthcare facilities, including various types of medical gloves, syringes, infusion bottles, metal and plastic medical instruments, and tissue residues as shown in Figure 2.

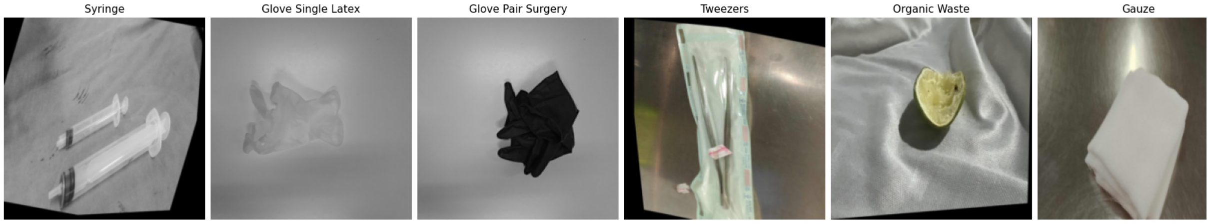

Figure 2. Representative samples from selected medical waste categories in the dataset. Only a subset of classes is shown for visualization clarity, while the complete dataset consists of 23 categories.

The number of images per class ranges from 300 to 845 samples. The gauze class contains the largest number of images (845), while classes such as urine_bag, glove_single_latex, and glove_pair_surgery are among the categories with the fewest samples (around 300 images each). The distribution of image counts for all classes is shown in Figure 3. This class imbalance is an important consideration when designing data augmentation techniques and data splitting strategies during model training.

Before being used for training, the dataset was filtered to ensure that only images with acceptable visual quality were included. Images that were blurred, had extreme lighting conditions, or contained excessive noise were removed to avoid disrupting the learning process of the model. Such filtering procedures are commonly applied in medical and waste image classification studies to maintain dataset consistency and to ensure that the visual representations learned by the model remain meaningful [2], [3].

C. *Data Preprocessing*

Data preprocessing was carried out to ensure that the images used in this study have adequate visual quality and are formatted according to the input requirements of the EfficientNet-B0 architecture. This

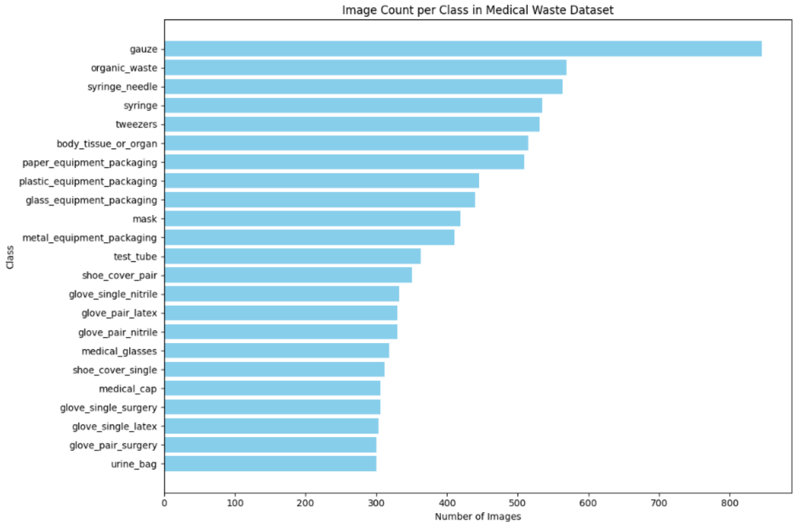

Figure 3. Distribution of the number of images for each class in the medical waste dataset

stage plays a critical role in the success of model training, as data quality and consistency strongly influence the ability of neural networks to learn discriminative visual patterns [2], [3]. Improper preprocessing can lead to slow convergence, overfitting, or an overall degradation in performance.

The preprocessing pipeline begins with data cleaning, in which duplicate, blurred, or irrelevant images are removed so that the dataset is free from noise and artifacts that could interfere with the learning process. Images with extremely dark or bright illumination, as well as images containing objects outside the defined categories, are also excluded. Such filtering procedures are consistent with medical image classification studies that emphasize the importance of high-quality and consistent data as the foundation for stable model performance [11].

After the initial screening, all images are resized to 224 × 224 pixels, which is the standard input size for the EfficientNet-B0 architecture [16]. This spatial standardization ensures that every image is processed within the same feature space and helps maintain training stability. A similar resizing strategy is applied in other modern models such as MedBin for medical waste classification, which reported that standardized input dimensions improve the stability of the training process [8].

The next step is pixel value normalization, where intensity values in the range 0-255 are scaled to the [0, 1] range. This transformation accelerates convergence and stabilizes the computations inside the network, particularly in transfer learning settings that are sensitive to the magnitude of the input [2], [3]. Such normalization has become a common practice in medical image processing and is a key component of modern training pipelines.

To increase data diversity and strengthen the generalization ability of the model, data augmentation is applied using several transformations, including random rotations up to ±30°, horizontal flipping, zooming, and brightness adjustments. These transformations enrich the visual representation of the dataset without explicitly increasing the number of original images, encouraging the model to recognize objects from different viewpoints and under varying illumination conditions. Priority work in medical imaging has highlighted the importance of augmentation in improving robustness and reducing overfitting, especially when the dataset is relatively small or imbalanced [21]. The use of such transformations is further supported by augmentation libraries such as Albumentations, which provide fast and flexible implementations widely adopted in recent image-based studies [15].

Finally, label encoding is performed by assigning a unique numerical label to each medical waste category. This step enables the model to associate each image with the corresponding target class during training, ensuring a consistent optimization process and allowing the predictions to be evaluated using standard classification metrics.

D. Model and EfficientNet-B0 Architecture

EfficientNet-B0 is used as the base model for the medical waste image classification task. This architecture, as shown in Figure 4, employs a compound scaling strategy that jointly scales network depth, width, and input resolution in a coordinated manner to achieve a balanced trade-off between accuracy and computational efficiency [3], [16]. In this study, EfficientNet-B0 acts as a feature extractor within a transfer learning framework, leveraging ImageNet-pretrained weights to support learning on a relatively limited medical waste dataset [2], [3].

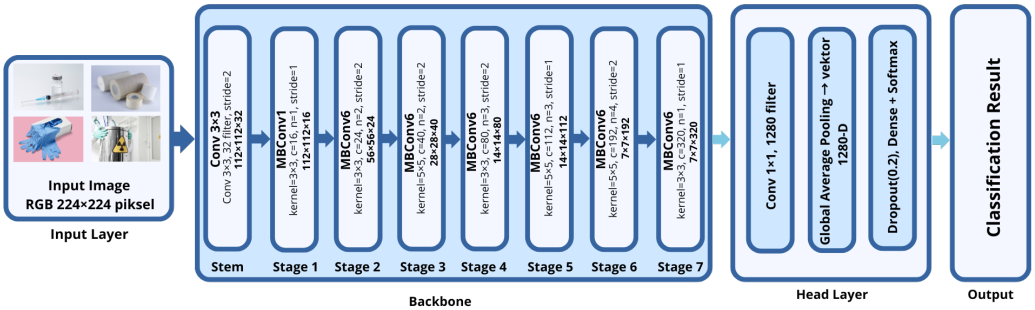

Figure 4. Architecture diagram of the EfficientNet-B0-based classification model

Structurally, EfficientNet-B0 can be divided into three main components: the stem, the backbone, and the head layer [16]. The stem processes input images of size 224 × 224 pixels using an initial convolutional layer to extract low-level features such as edges and simple textures. The backbone consists of a sequence of Mobile Inverted Bottleneck Convolution (MBConv) blocks arranged hierarchically. Each MBConv block uses depthwise separable convolution, expansion layers, squeeze-and-excitation (SE) modules, and the Swish activation function to improve computational efficiency and feature representation capacity [16]. The spatial resolution of the feature maps is gradually reduced, from 112 × 112 down to 7 × 7, as the network depth increases. In general, the two-dimensional convolution operation applied to a feature map can be expressed as in Equation 1:

$$Y(i\,,j)\ =\sum_{m}\sum_{n}X(i+m,j+n)\times K(m,n) \tag{1}$$

where X denotes the input image or feature map from the previous layer, K is the convolution kernel (filter), and Y is the resulting output feature map. The indices i and j represent the spatial coordinates of the output feature map, while m and n denote the kernel indices used during the convolution operation. In EfficientNet-B0, this computation is implemented using depthwise separable convolutions to reduce the number of parameters and floating-point operations while maintaining strong representational power [16].

Each convolutional block in EfficientNet-B0 employs the Swish activation function, which has been shown to provide smoother gradients and better performance than ReLU, particularly in large-scale models [16]. The Swish activation is defined as in Equation 2:

$$f(x) = x\ \times\ \sigma(x)\ =\frac{x}{(1+\mathrm{e}^{-x})} \tag{2}$$

where x represents the input activation value to the activation function and $\sigma(x)$ denotes the logistic sigmoid function. This activation helps maintain gradient stability during training and accelerates model convergence.

After feature extraction in the backbone, the head of the network applies a 1 × 1 convolutional layer with 1,280 filters to aggregate spatial information across channels. The resulting feature maps are then summarized using Global Average Pooling (GAP), producing a 1,280-dimensional feature vector. A fully connected layer with a Softmax activation is subsequently applied to generate class probability estimates. The Softmax function for class i is defined as in the following Equation 3:

$$S_{\mathrm{i}}(x) = \frac{e^{x_i}}{\left(1\,\sum_{j=1}^{N} e^{x_j}\right)} \tag{3}$$

where $Si(x)$ is the predicted probability for class i, x_i and x_j denote the logits associated with classes i and j, respectively, and N is the total number of classes. The class with the highest probability is selected as the final prediction.

Within the transfer learning paradigm, EfficientNet-B0 pretrained on ImageNet is treated as a source model $fs(x;\theta s)$. Adaptation to the medical waste domain is performed through fine-tuning, which updates the model parameters based on the target dataset. Conceptually, this process can be expressed as in the following Equation 4:

$$f_t(x) = f_s(x;\ \theta_s) + \Delta\theta_t \tag{4}$$

where $f_t(x)$ denotes the adapted target model and $\Delta\theta_t$ represents the parameter updates learned from the medical waste dataset. This approach enables EfficientNet-B0 to extract features that are more relevant to the visual characteristics of medical waste, even when the training dataset is limited in size [2], [3], [20], [25].

Overall, the use of EfficientNet-B0 provides an efficient feature extraction mechanism with stable performance, making this architecture well suited for medical waste image classification tasks that require a balance between accuracy and computational cost [16], [20].

E. *Evaluation Protocol*

Model evaluation was conducted to assess the performance and stability of the proposed medical waste image classification model. A 5-fold cross-validation scheme was employed to ensure that the model not only achieves high performance on a single subset of the data but also generalizes well across the entire dataset. This evaluation strategy is widely adopted in CNN-based medical imaging studies, as it provides more reliable performance estimates and reduces dependence on a particular train–test split [11], [22].

In the 5-fold cross-validation procedure, the dataset is partitioned into five subsets of approximately equal size. In each iteration, four folds are used for training, while the remaining fold is reserved for testing. This process is repeated five times, ensuring that each fold serves as the test set exactly once. During each iteration, several evaluation metrics are computed, including accuracy, precision, recall, and F1-score. The final performance of the model is reported as the average of these metrics across all five folds.

Mathematically, accuracy (A), precision (P), recall (R), and F1-score are defined in Equation 5, Equation 6, Equation 7, and Equation 8:

$$A = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

$$P = \frac{TP}{TP + FP} \tag{6}$$

$$R = \frac{TP}{TP + FN} \tag{7}$$

$$F1 = 2 \text{ x } \frac{P\ X\ R}{P + R} \tag{8}$$

where TP (true positives) are samples that belong to the positive class and are correctly predicted as positive by the model, TN (true negatives) are samples that do not belong to the positive class and are correctly predicted as negative, FP (false positives) are samples that do not belong to the positive class but are incorrectly predicted as positive, and FN (false negatives) are samples that belong to the positive class but are incorrectly predicted as negative. For the multiclass setting, precision, recall, and F1-score are computed using macro-averaging across all classes.

The mean value of a given evaluation metric M over the k=5 folds is computed as shown in Equation 9 below, where Mi denotes the value of metric M obtained in the i-th fold and k represents the total number of folds.

$$\bar{M} = \frac{1}{k}\sum_{i=1}^{k} M_i \tag{9}$$

All evaluations are performed after the model has been fine-tuned on the EfficientNet-B0 architecture. The averaged results across the five folds are used to assess the consistency and robustness of the model's performance, as well as to identify potential overfitting or bias toward specific subsets of the data. Consequently, the reported performance is expected to be more reliable and generalizable to unseen medical waste images.

IV. Results and Discussion

The EfficientNet-B0–based classification model was implemented using the architecture described in Section III.D, where the pretrained EfficientNet-B0 backbone was employed as a feature extractor within a transfer learning framework. Model training was conducted using the Adam optimizer with a learning rate of 0.0001, a batch size of 32, and 30 training epochs.

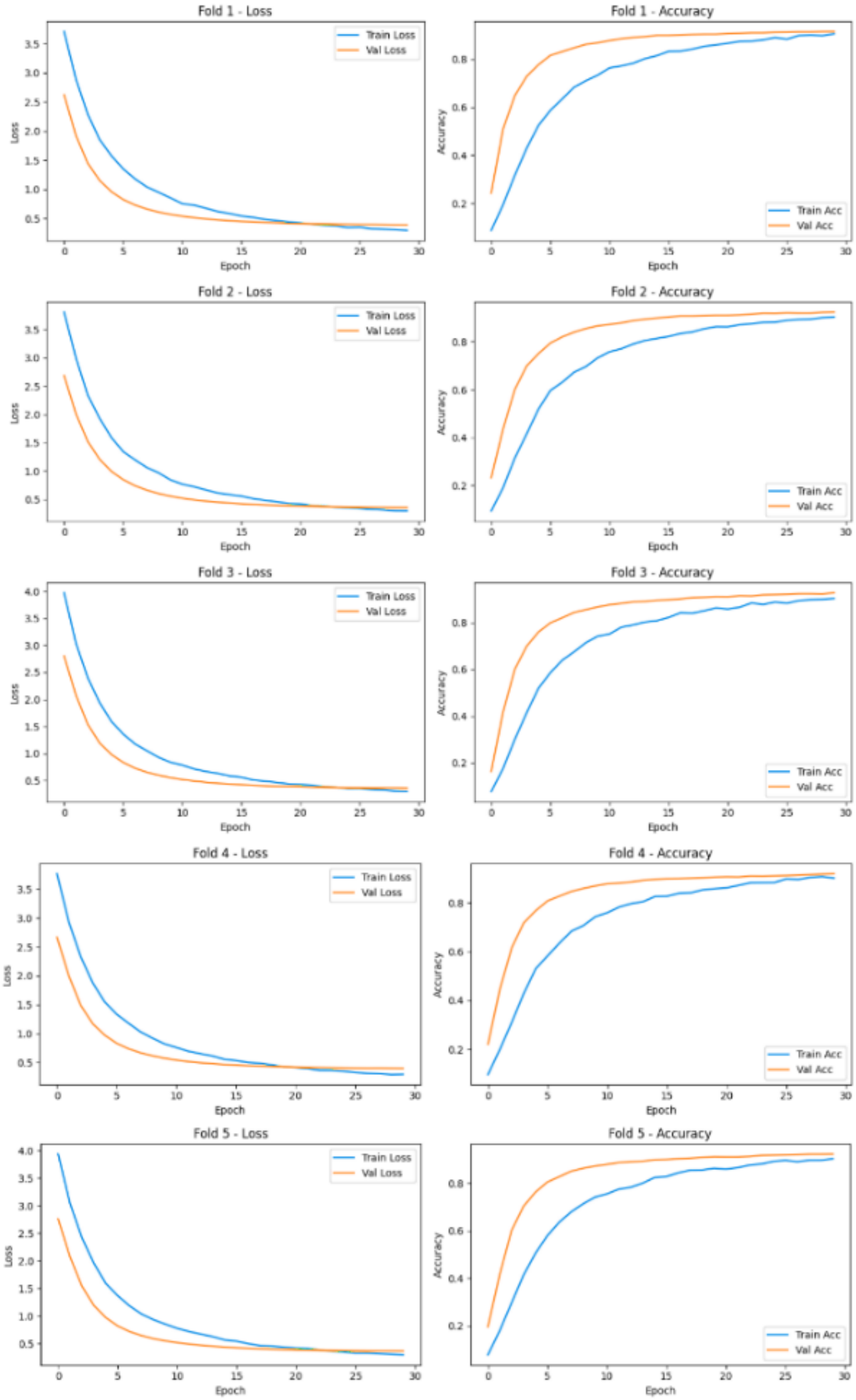

Figure 5. Training and validation loss and accuracy curves for the EfficientNet-B0 model.

The training and validation curves presented in Figure 5 show a stable convergence behavior. Both training and validation loss decrease consistently as the number of epochs increases, while validation accuracy improves steadily and remains close to the training accuracy throughout the training process. This pattern suggests that the model learns discriminative features effectively without exhibiting pronounced overfitting, indicating that the combination of fine-tuning and data augmentation contributes positively to generalization. Similar effects have been reported in recent medical imaging studies employing transfer learning and augmentation strategies [2], [21].

To obtain a robust estimate of classification performance, the model was evaluated using the 5-fold cross-validation scheme described in Section III.E. The accuracy, precision, recall, and F1-score obtained for each fold are summarized in Table 2. Accuracy values range from 0.9159 to 0.9284, resulting in an average accuracy of 0.9222 across the five folds. The mean precision reaches 0.9411, while the average recall and F1-score are 0.9222 and 0.9221, respectively. The relatively narrow variation of all four metrics across folds indicates that the model exhibits consistent behavior under different train–test partitions.

Table 2. Performance of the EfficientNet-B0 model using 5-fold cross-validation.

Fold	**Accuracy**	**Precision**	**Recall**	**F1-Score**
1	0.9159	0.9381	0.9159	0.9158
2	0.9241	0.9457	0.9241	0.9247
3	0.9284	0.9433	0.9284	0.9279
4	0.9192	0.9413	0.9192	0.9186
5	0.9225	0.9391	0.9225	0.9224
Average	0.9222	0.9411	0.9222	0.9221

The distribution of evaluation metrics across the five folds is illustrated in Figure 6. The limited spread of accuracy, precision, recall, and F1-score confirms the stability of the classifier across different data partitions. Precision exhibits the highest median value, indicating that the model effectively limits false positive predictions. The close alignment between recall and F1-score further suggests a balanced trade-off between sensitivity and precision, which is particularly important in safety-critical applications such as medical waste handling.

Figure 6. Distribution of accuracy, precision, recall, and F1-score across the 5 folds.

Further insight into class-wise performance is provided by the confusion matrix shown in Figure 7. The EfficientNet-B0 classifier correctly identifies most medical waste categories, as reflected by strong diagonal values. Classes such as syringe, tweezers, and organic_waste demonstrate particularly high recognition rates. A limited number of misclassifications are observed among glove-related classes, especially between single and pair variants, which share similar visual characteristics in terms of shape, material, and color. However, the misclassification rate within these classes remains below approximately 7% and does not significantly affect overall model performance.

Overall, the experimental results demonstrate that the proposed transfer learning approach using EfficientNet-B0 yields an accurate and computationally efficient model for multiclass medical waste image classification on a limited dataset. The achieved average accuracy of 92.2% and precision of 94.1% are consistent with prior studies on medical and general waste classification employing CNN-based and lightweight architectures [8], [17], [20]. In addition, previous work on computer vision–assisted medical waste sorting has reported comparable performance levels in practical settings [19].

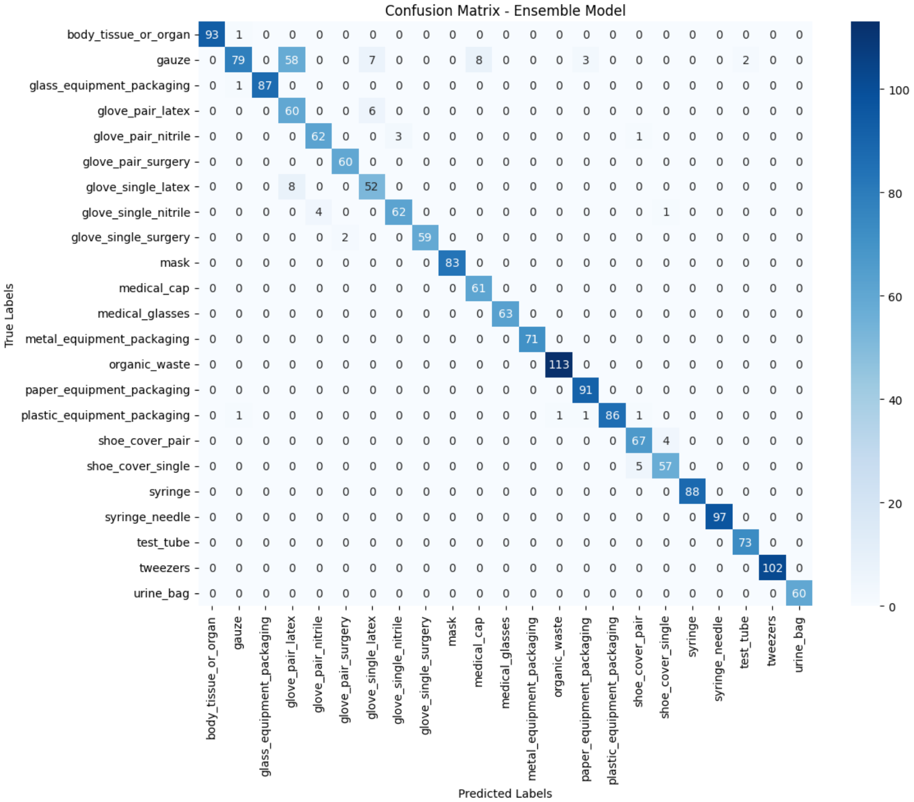

Figure 7. Confusion matrix of the medical waste image classification results using EfficientNet-B0.

The use of 5-fold cross-validation enhances the reliability of the reported results by reducing the influence of a particular train–test split, in line with recommended evaluation practices in AI-based medical imaging [11], [22]. Considering the stability across folds, the favorable per-class performance, and consistency with existing literature, the EfficientNet-B0–based model represents a viable foundation for further development of automated medical waste classification systems. A direct comparison with alternative lightweight architectures is left for future work, as the focus of this study is on evaluating the stability and effectiveness of EfficientNet-B0 under limited data conditions.

V. Conclusion

This study presents a transfer learning–based approach using the EfficientNet-B0 architecture for medical waste image classification. Experimental results obtained through a 5-fold cross-validation scheme demonstrate that the proposed model achieves stable and consistent performance, with an average accuracy of 92.2%, precision of 94.1%, recall of 92.2%, and an F1-score of 92.1%. These results indicate that EfficientNet-B0 is capable of learning discriminative visual patterns across 23 medical waste categories, even under limited data conditions. The application of fine-tuning and data augmentation contributes to stable convergence and balanced class-wise performance, as reflected in the confusion matrix analysis, where most categories are recognized with high accuracy and only minor confusion occurs among visually similar classes. Despite these promising results, this study has several limitations. The experiments are conducted on a single publicly available dataset with a limited scale and do not include real-world deployment or validation in operational healthcare environments. In addition, the proposed model is not experimentally compared with other lightweight architectures such as MobileNet or YOLO-based models, which may exhibit similar efficiency characteristics. Future work should therefore focus on expanding the dataset with images collected from diverse healthcare facilities, conducting systematic comparisons with alternative efficient architectures, and integrating the model into real-world medical waste sorting systems to further assess its robustness, scalability, and practical applicability.

Conflicts of Interest

The authors declare no conflicts of interest.

Author Contributions Statement

Reni Kartika Suwandi: conceptualization, methodology design, data collection, data curation, implementation, model training, experimental analysis, and writing original draft preparation.
Asep Saeppani: supervision, refinement of methodology, guidance on experimental design, review and editing. Irfan Fadil: supervision, validation of results, technical feedback on model development, and review and editing. All authors have read and approved the final version of the manuscript.

Acknowledgment

This research received no external funding. The authors would like to thank their academic supervisors and colleagues for the constructive feedback, as well as the supporting staff from their institutions for the technical and administrative assistance provided during this work.

References

[1] E. Janik-Karpinska, R. Brancaleoni, M. Niemcewicz, W. Wojtas, M. Foco, M. Podogrocki, and M. Bijak, "Healthcare Waste—A Serious Problem for Global Health," Healthcare, vol. 11, no. 2, p. 242, Jan. 2023, doi: 10.3390/healthcare11020242H.

[2] H. E. Kim, A. Cosa-Linan, N. Santhanam, M. Jannesari, M. E. Maros, and T. Ganslandt, "Transfer learning for medical image classification: a literature review," BMC Medical Imaging, vol. 22, no. 1, p. 69, 2022, doi: 10.1186/s12880-022-00793-7.

[3] P. Kora, C. P. Ooi, O. Faust, U. Raghavendra, A. Gudigar, W. Y. Chan, K. Meenakshi, K. Swaraja, P. Plawiak, and U. R. Acharya, "Transfer learning techniques for medical image analysis: A review," Biocybernetics and Biomedical Engineering, vol. 42, no. 1, pp. 79–107, 2022, doi: 10.1016/j.bbe.2021.11.004.

[4] H. G. Kim, Y. Choi, and Y. M. Ro, "Modality-bridge transfer learning for medical image classification," in Proc. 10th Int. Congr. Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Shanghai, China, Oct. 2017, pp. 1–6, doi: 10.1109/CISP-BMEI.2017.8302286.

[5] J. Chae and J. Kim, "An investigation of transfer learning approaches to overcome limited labeled data in medical image analysis," Applied Sciences, vol. 13, no. 15, p. 8671, 2023, doi: 10.3390/app13158671.

[6] S. Kunwar and P. Rai, "Health care waste classification using deep learning aligned with Nepal's bin color guidelines," arXiv:2508.07450 [cs.CV], 2025, doi: 10.48550/arXiv.2508.07450.

[7] P. Akkajit and A. Sukkuea, "Medical waste classification using convolutional neural network," in E3S Web of Conferences, vol. 530, Art. no. 04001, 2024, doi: 10.1051/e3sconf/202453004001.

[8] X. Xu, C. Wang, Q. Yi, J. Ye, X. Kong, S. Q. Ashraf, K. D. Dearn, and A. M. Hajiyavand, "MedBin: A lightweight end-to-end model-based method for medical waste management," Computers in Industry, vol. 163, 2025.

[9] B. van den Broek and J. P. R. S. Sharami, "Improving medical waste classification with hybrid capsule networks," arXiv pre-print arXiv:2503.10426, 2025.

[10] T. Mythili and A. Anbarasi,"A concatenation of deep and texture features for medicinal trash image classification using En-SegNet-DNN-based transfer learning," Materials Today: Proceedings, vol. 66, pp. 3420–3428, 2022, doi: 10.1016/j.matpr.2022.03.129.

[11] A. Akram and R. Debnath, "An efficient automated corneal ulcer detection method using convolutional neural network," in Proc. 22nd Int. Conf. Computer and Information Technology (ICCIT), Dhaka, Bangladesh, Dec. 2019, pp. 1–6. doi: 10.1109/CISP-BMEI.2017.8302286.

[12] A. M. Alabdali, "Blockchain-based solid waste classification with AI-powered tracking and IoT integration," Scientific Reports, vol. 15, no. 1, Art. no. 15197, 2025, doi: 10.1038/s41598-025-97030-2.

[13] M. M. Islam, S. M. M. Hasan, M. R. Hossain, M. P. Uddin, and M. A. Mamun, "Towards sustainable solutions: Effective waste classification framework via enhanced deep convolutional neural networks," PLOS ONE, vol. 20, no. 6, p. e0324294, 2025, doi: 10.1371/journal.pone.0324294.

[14] M. K. Hasan, A. Atta, M. A. Khan, A. S. Akram, G. F. Issa, and M. Hassan, "Smart waste management and classification system for smart cities using deep learning," in Proc. 2022 Int. Conf. Business Analytics for Technology and Security (ICBATS), Dubai, United Arab Emirates, 2022, pp. 1–6, doi: 10.1109/ICBATS54253.2022.9759087.

[15] A. Buslaev, A. Parinov, E. Khvedchenya, V. I. Iglovikov, and A. A. Kalinin, "Albumentations: Fast and flexible image aug-mentations," arXiv preprint arXiv:1809.06839, 2018.

[16] M. Tan and Q. V. Le, “EfficientNet: Rethinking model scaling for convolutional neural networks,” arXiv preprint arXiv:1905.11946, 2020.

[17] H. Zhou, X. Yu, A. Alhaskawi, Y. Dong, Z. Wang, Q. Jin, X. Hu, Z. Liu, V. G. Kota, M. H. A. Hasan Abdulla, S. H A. Ezzi, B. Qi, J. Li, B. Wang, J. Fang, and H. Lu, “A deep learning approach for medical waste classification,” Scientific Reports, vol. 12, p. 2159, 2022, doi: 10.1038/s41598-022-06146-2.

[18] M. K. Nallapaneni, M. A. Mohammed, K. H. Abdulkareem, R. Damaševičius, S. A. Mostafa, M. S. Maashi, and S. S. Chopra, "Artificial intelligence-based solution for sorting COVID-related medical waste streams and supporting data-driven decisions for smart circular economy practice,” Process Safety and Environmental Protection, vol. 152, pp. 482–494, 2021, doi: 10.1016/j.psep.2021.06.026.

[19] A. Bruno, C. Caudai, G. R. Leone, M. Martinelli, D. Moroni, and F. Crotti, “Medical waste sorting: a computer vision approach for assisted primary sorting,” in Proc. ICASSP 2023 IEEE Int. Conf. Acoustics, Speech and Signal Processing Workshops (ICASSPW), 2023, pp. 1–5, doi: 10.1109/ICASSPW59220.2023.10193520.

[20] W. M. Ardana and K. Kusrini, “Optimasi algoritma convolutional neural network dengan arsitektur EfficientNet-B0 dan ResNet-50 untuk klasifikasi jenis sampah,” MALCOM: Indonesian Journal of Machine Learning and Computer Science, vol. 5, no. 4, pp. 1274–1286, 2025, doi: 10.57152/malcom.v5i4.2030.

[21] T. Islam, M. S. Hafiz, J. R. Jim, M. M. Kabir, and M. F. Mridha, “A systematic review of deep learning data augmentation in medical imaging: recent advances and future research directions,” Healthcare Analytics, vol. 5, p. 100340, 2024, doi: 10.1016/j.health.2024.100340.

[22] T. J. Bradshaw, Z. Huemann, J. Hu, and A. Rahmim, “A guide to cross-validation for artificial intelligence in medical imaging,” Radiology: Artificial Intelligence, vol. 5, no. 4, e220232, 2023, doi: 10.1148/ryai.220232.

[23] I. D. A. P. P. Tentriajaya and I. G. N. L. Wijayakusuma, “Eye disease classification using EfficientNet-B0 based on transfer learning,” Journal of Applied Informatics and Computing, vol. 9, no. 4, pp. 1415–1422, 2025, doi: 10.30871/jaic.v9i4.9743.

[24] A. A. Rakib, M. M. Billah, A. S. Ahamed, H. M. Imamul, and M. S. A. Masum, “EfficientNet-based model for automated classification of retinal diseases using fundus images,” European Journal of Computer Science and Information Technology, vol. 12, no. 8, pp. 48–61, 2024, doi: 10.37745/ejcsit.2013/vol12n84861.

[25] T. Kurniawan, Khadijah, and R. Kusumaningrum, “An efficient model for waste image classification using EfficientNet-B0,” Jurnal Teknik Informatika (JUTIF), vol. 6, no. 3, pp. 1147–1158, 2025, doi: 10.52436/1.jutif.2025.6.3.4417.