Journal of Computer Science

Original Research Paper

Performance Analysis of Garbage Classification Using Balanced and Unbalanced Dataset with EfficientNetV2B1 Architecture

[1]Retinderdeep Singh, [1]Gurjot Kaur, [1]Chander Prabha, [2]Ayman Noor, [3]Mohammad Zubair Khan and [4]Prakash Srivastava

[1]Department of Computer Science and Engineering, Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India
[2]Department of Computer Science, College of Computer Science and Engineering, Taibah University, Madinah, Saudi Arabia
[3]Department of Computer Science and Information, Taibah University, Madinah, Saudi Arabia
[4]Department of Computer Science and Engineering, Graphic Era University, Dehradun, India

Article history
Received: 03-05-2024
Revised: 31-05-2024
Accepted: 14-06-2024

Corresponding Author:
Chander Prabha
Department of Computer Science and Engineering, Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India
Email: prabhanice@gmail.com

Abstract: This research concentrates on EfficientNetV2B1 deep learning model to classify garbage collection tasks, with both balanced and unbalanced dataset configurations. The dataset includes 7,260 images for the balanced dataset and 15,515 for the unbalanced one, both of the datasets are used to train and evaluate the model. Training and evaluation of the deep learning model with the standard performance variables: Accuracy, precision, recall, F1-score, and AUC score. The results indicate the unbalanced dataset performs excellently, with an accuracy of 95.22%, precision at 95.28%, recall of 95.22%, and F1-score of 95.21%. In contrast, the fully balanced data set yields slightly less but still excellent results: Accuracy of 91.46%, precision of 91.60%, recall of 91.46%, and F1 score of 91.43%. The test scores for loss and accuracy in both datasets are 0.4296 and 0.9522 for the unbalanced dataset and 0.60189 and 0.9146 for the balanced dataset respectively. A study with a dataset containing balanced classes is beneficial for assessing EfficientNetV2B1 deep learning performance across different classes evenly, providing a fair evaluation of the model's ability to generalize. On the other hand, a study with an unbalanced class distribution can be useful for evaluating how well EfficientNetV2B1 deep learning handles class imbalance and its performance in minority classes. Both types of studies offer valuable insights into model behavior under different data scenarios.

Keywords: Garbage Classification, EfficientNetV2B1, Deep Learning, Image Classification, Performance Analysis

Introduction

In today's world, waste and the management of garbage has become an increasingly urgent issue. With exploding populations and ever-increasing urban collapses, efficient waste management systems are essential to maintain environmental sustainability and public health. Among the various aspects of waste management, garbage collection is essential. With traditional methods, not only will a lot of manpower be needed but the cost could also be high. With all sorts of technological advances, there is a growing awareness about using artificial intelligence and deep learning to tackle society's many problems. A challenging problem that has received a great deal of attention is how waste and garbage collection can be effectively managed. In the garbage collection process, the proper waste materials should be classified and separated it assists with recycling efforts, reduces environmental pollution, and makes use of resources more efficient. With traditional waste classification approaches, manual sorting is the standard. This is labor-intensive and time-consuming yet disposed to faults. In contrast, recent AI and deep learning advances provide a promising way to automate the garbage collection process (Ihsanullah *et al.*, 2022). The classification of garbage collection techniques provides a systematic way of understanding the divergent methods of memory management in use today. By categorizing garbage collection methods based on their underlying principles, developers can make decisions on memory management strategies that take into account factors like performance and resource consumption that also meet the demands of their applications. Garbage collection techniques can be classified according to several criteria.

DOI: 10.3844/jcssp.2024.1310.1321

Based on the events that trigger garbage collection, people use this standard. The main triggers are man-initiated and computer-scheduled memory pressure. Different algorithms for identifying and reclaiming memory-free GC techniques can be classified based on collection algorithms. These algorithms include mark and sweep, reference counting, types of copying, and generational collection. Whether the garbage collection shall be done concurrently with program execution. That is the concept of concurrency of GC i.e., the process happens simultaneously while live programs are running. Garbage encompasses virtually all disposable materials of various kinds. How these materials should be handled and disposed of depends upon what kind of material they are. Organic waste includes those materials able to degrade in the natural environment, such as food scraps and garden waste which can be put into compost heaps or subject to anaerobic digestion. Recycling materials like paper, cardboard, glass, metal cans, and some plastics are turned into new products, saving virgin materials and energy. Dangerous waste (certain chocolate only, batteries, electronic waste, and medical waste) can pose safety threats to human health and environmental pollution dangers. Therefore, it needs special treatment and processing to remove these symptoms at the source. E-waste is a treasure trove of useful parts but also a source of deadly toxins such as lead and mercury. It needs to be disposed of properly to protect the environment and human health. Construction and demolition waste such as concrete, brick, and soil can be heavy and difficult to manage in bulk but is chemically stable. The mixed nature of the trash is potentially harmful. Plastic waste, including disposable packaging, bottles, plastics, and other items made with plastics remains a major environmental problem that calls for reduction, reuse, and recording as well as policies rooted in circular economics. Comprehensive separation, storage, and treatment can help minimize the environmental impact and promote sustainable development. To ensure that the garbage system provides a full set of garbage processing and recycling facilities. The exchange rate may vary across different regional markets and applications in which the exchanger is present (Farooq *et al*., 2022; Abdallah *et al*., 2020). This study explores garbage collection classification employing deep learning algorithms, focusing in particular on the EfficientNetV2B1 architecture. The EfficientNetV2 series by Google scientists is a family of Convolutional Neural Network (CNN) models that have excelled in various image classification tasks.

Problem Statement

The problem statement of this study is a wide range of types and a rough distribution in data sets. Traditional methods are often done by hand, which has disadvantages for both manpower and error rates. Deep learning architectures have been successful in garbage sorting automation, but the outcome may be different based on how balanced or imbalanced the dataset is. So, to have a look at just how powerful deep learning models like EfficientNetV2B1 could be for garbage collection classification tasks while taking into account the distribution of such datasets and their effect on model performance. The EfficientNetV2B1 architecture has been chosen as the main deep learning model for this research on the strength of its established commitment to top image classification performance. Since developed as a sequel to EfficientNet, which introduced a pioneering scaling method that allows network depth, width and resolution to be balanced, the EfficientNetV2 series carries on these ideas but is suitable for achieving even more efficient models and higher accuracy. By utilizing the architecture innovations and optimization techniques encapsulated within the EfficientNetV2B1 model, researchers hope thereby to fully exploit the power of deep learning in garbage collection classification.

Contribution

This study makes significant contributions to the fusion of waste management and deep learning through rewording key elements of many issues. This study presents a combined dataset in which there are both balanced and unbalanced subsets, making it possible to compare the performance of various models under different distributions of datasets. Next, by making use of the latest EfficientNetV2B1 architecture, this study demonstrates that in garbage collection classification it tops all others in performance measures with less computing resource allocated. Thirdly, the study assesses the model's classification capabilities on both balanced and unbalanced data sets through the use of comprehensive performance evaluation metrics such as accuracy, precision, recall, F1-score, and ROC score. Both in detail and in general, it shows that with different dataset distributions, the performance of the model will vary. This not only sheds light on what dataset distribution means to classification accuracy and effectiveness but also will offer real-world insights that could become valuable for them from an academic standpoint. Finally, this research aims to further comprehension of deep learning-based methods in garbage collection classification and contribute to the construction of waste management systems that are more efficient and safer.

Literature Review

In recent years, varying methods of garbage classification have sprung up as a strategy for dealing with waste. The purpose of these systems is to boost recycling efficiency, reduce environmental pollution, and make maximum use of limited resources. Deep learning techniques particularly Convolutional Neural Networks (CNNs) are bringing groundbreaking innovation to the field of garbage classification. This technological revolution has delivered more accurate systems that are also faster and easier to scale up than ever before. This literature review sets out to examine the main studies, the technology's progress, and the future development of garbage classification constructed with deep learning methods.

Chen *et al*. (2022) proposed a garbage classification system based on an upgraded version of the ShuffleNet v2 architecture. The significance of their study was in optimizing network architectures for efficient garbage classification, emphasizing that model design can bring both high accuracy and robustness. They put forward innovative methods of feature extraction and classification which, compared against traditional methods led to a big leap in classification performance. Meng and Chu (2020) applied a deep-learning algorithm to classify all sorts of garbage with CNNs. Their study demonstrated the potential of deep learning algorithms in discriminating accurately between various kinds of waste. They pointed out that data preprocessing, data augmentation, and model fine-tuning played a crucial role in improving classification results. They also talked about the scalability of CNNs for big garbage classification tasks and tackled problems like the size of the data set you need or how many different classes there are in different portions of it.

Yang *et al*. (2021) proposed GarbageNet, an integrated learning framework designed to improve the robustness and generalization capacity of a garbage classification model. The new model includes advanced regularization techniques, ensemble learning strategies, and data augmentation methods to improve its performance under different conditions. Their research helped produce garbage classification systems that were more capable of standing up to the real world's complexity. Zhang *et al*. (2022) studied the dioxin emissions resulting from different kinds of garbage biodegradable. For that reason their construct links machine learning algorithms with environmental investigations of the latter. They used machine learning algorithms to process environmental impact assessment results. Their work also featured the importance of thinking about sustainable measures in conjunction with garbage classification frameworks, so to promote eco-friendly waste management solutions. According to the Song *et al*., (2024) report, an example of a household DEEPBIN deep learning-based garbage sorting system has been created. Their system took advantage of sustainable natural technologies, such as biodegradable materials and energy-efficient processing, thus promoting a greener waste management practice. This study underscored the integration of sustainability principles and technology in the research and development of intelligent garbage classification solutions.

Wang and Wen (2024) developed and deployed a real-time garbage classification detection system, based on YOLO v8's architecture. Their system showed the use of deep learning tech even in places with few resources, such as sanitation, stations, and smart cities. Their paper made an important contribution by providing a proven method for sorting waste on a large scale across many countries. Their work contributed directly to the building of industrial solutions, which are scalable and deployable for garbage sorting. In a paper, Wang (2024), conducted a thorough investigation of the classification technique for identifying garbage images based on deep learning approach methods. The author emphasized the role that features representation, model choice, and Hyperparameter tuning play in determining whether one gets high accuracy in classifying images of garbage as rubbish. His research was very useful indeed. While not every word is recalled verbatim, it is evident that the main characteristics of utilizing deep learning architectures for classifying various types of garbage were outlined. Jin *et al*. (2023) introduced a new machine vision system that enables the monitoring of waste in real time and can even separate waste from recyclables. Their system also showed that when artificial intelligence is used in garbage classification, the efficiency of separating and finishing off garbage is increased while the workload for waste management staff is reduced.

Li *et al*. (2023) focused on fusing feature representations for poor image classification in human-robot interactions. This study enhanced the interpretability and robustness of the garbage classification model under dynamic conditions. It makes an important contribution to the development of adaptive garbage classification systems that can handle complex interactions between people and robots. Yuan *et al*. (2020) research used CNNs to perform real-time multiple single garbage classification. It tackles the problems of real-time processing and classification accuracy in dynamic environments. The study stressed how model optimization is very beneficial for the real-world identification of trash in aerial or overhead imagery.

Rismiyati *et al*. (2020) utilized transfer learning with Xception architecture for garbage classification, gave powerful verification of the transference of the pre-trained model, and that this cross-domain transfer can lead to better results from your point of view. Their research has helped lay the groundwork for efficient and scalable garbage classification systems based on transfer learning. Zhao *et al*. (2022) proposed an intelligent refuse sorting scheme, which is founded on the new MobileNetV3-large architecture. In the system, they exploited higher-level feature extraction methods along with model optimization

DOI: 10.3844/jcssp.2024.1310.1321

techniques to earn state-of-the-art performance in garbage classification. However, pivotal to their research was the focal point on the continuous expansion of model architectures tailored for refuse sorting applications worldwide.

These studies show the various applications, methods, and challenges involved in applying deep learning technology to garbage classification. Researchers are using deep learning algorithms to develop intelligent, sustainable scalable garbage classification systems capable of solving modern society's complex garbage management problems.

Materials and Methods

The dataset used in this study is Garbage classification (Mohamed, 2021) which contains images of garbage demonstrating distinct categories of garbage collection. In those classes, each corresponding to the image is labeled with a special class of the garbage collection it represents. The dataset used in this study is indeed very diverse, containing items from various geographic locations and different waste management facilities. This diversity ensures the model's robustness and applicability to a wide range of real-world scenarios. There are 12 categories in this distribution: Battery, biological, brown-glass, cardboard, clothes, green-glass, metal, paper, plastic, shoes, trash, and white-glass (Fig. 1). In the unbalanced dataset, there are 15,515 examples. For training and testing purposes this dataset is divided into training and testing sets, with the training set having 13,932 images and the validation set 1,549 images.

In the balanced dataset, there are 7,260 images. similar to the unbalanced dataset, this dataset was divided into training and test sets for training and evaluation purposes. the training dataset of the balanced dataset contains 6,534 images, while the test dataset is composed of 726 images. Both datasets are preprocessed and augmented before training to introduce enough variations and resilience. These preprocessing and augmentation methods can improve the model's ability to generalize from previously unknown data and overall performance at the same time (Nnamoko *et al*., 2022; Malik *et al*., 2023).

In this analysis, the EfficientNetV2B1 architecture is used for garbage classification, as shown in Fig. 2. The dataset consists of 12 different types of garbage classes. Deep learning tasks are performed in the combination of this dataset and additional data about the promotion of garbage collection, acquired from Kaggle. This added information could enable broader critique studies, including testing and comparison of recognition results. The method encompasses loading and preparing image data and building the model on EffectiveNetV2 B1 hardware with custom classification levels.

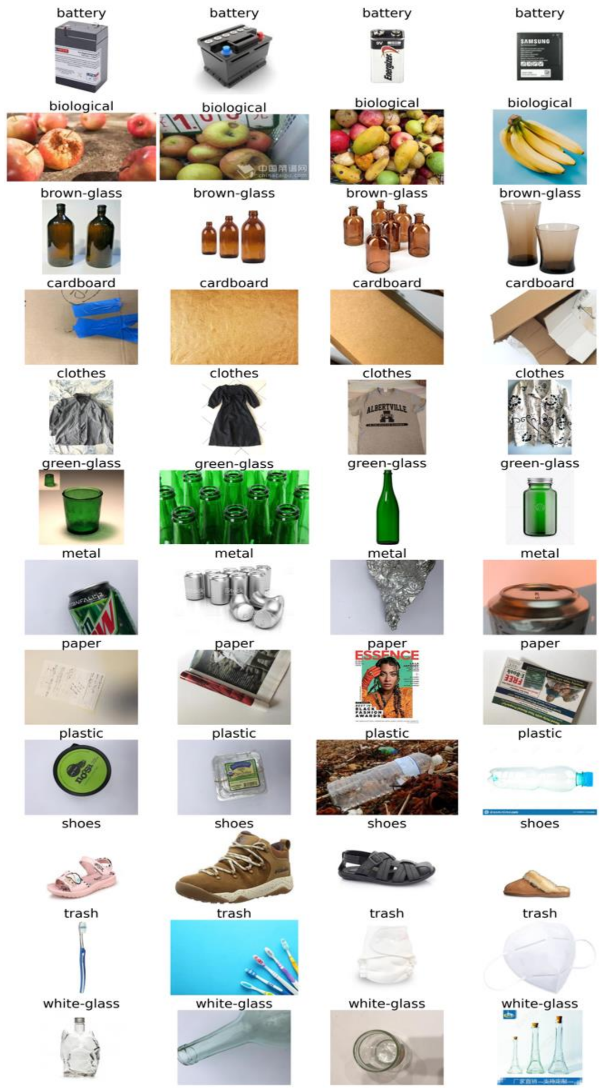


Fig. 1: Input dataset

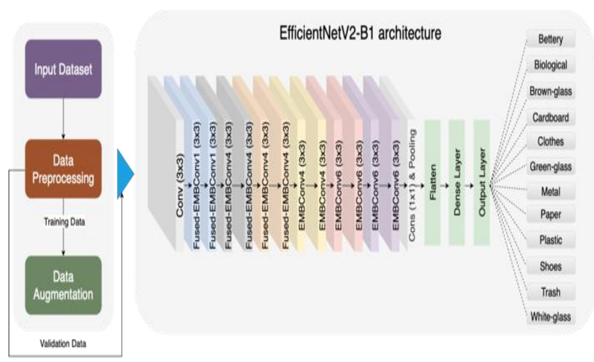


Fig. 2: Proposed methodology

DOI: 10.3844/jcssp.2024.1310.1321

Data Preprocessing and Augmentation

In the preprocessing and augmentation stages of the experiment, a variety of methods were used to improve the quality and variability of the dataset before training the model. Initially, images were resized to a uniform 224×224 pixel dimension to give all images an equal footing and also keep important elements. This resizing step can also reduce the computation complexity of training. Next, the values of image pixels were rescaled so that all fell between 0 and 1. Doing this makes the normalized data, allowing each feature to contribute equally to model learning. Also, data augmentation techniques were used to augment the dataset artificially. Other techniques such as random rotation, horizontal flipping and vertical flipping, changing image brightness and randomly zooming into images were used to alter the appearance of the dataset without changing its general class nature. These types of techniques serve to strengthen the model's ability to generalize to new data by exposing it to a wide range of scenarios and variations that may occur during inferencing. Additionally, to prevent the model from overfitting and to enhance its robustness in general, a dropout layer with a rate of 0.5 was inserted after the feature extraction stage. During training, this dropout layer randomly drops away some of the neurons unnecessarily forcing the model to learn more robust features by reducing this reliance on particular input patterns. In short, combining preprocessing techniques including resizing and normalization with data augmentation strategies and dropout regularization leads to a more comprehensive dataset, able to support the model as it learns and assists it in generalizing well on unfamiliar garbage images (Lou and Gou, 2023; Verma *et al*., 2019).

Model Architecture

EfficientNetV2B1 pre-trained model, with some extra layers to fine-tune and classify data. Here is a detailed breakdown of those layers (Fig. 3):

a. Base model (EfficientNetV2B1): Input shape: (224, 224, 3) takes an image 224 pixels high by 224 pixels wide and 3 channels of color (RGB). Include top: False the top classification layers of the pre-trained model are used. Weights: 'Imagenet'. The pre-trained weights are initialized by the ImageNet dataset. Trainable: False the weights of the trained model are frozen so that the learned characteristics from a previous or another training regime can be retained and no further training takes place. Output Shape: (None, 4, 4, 1280). The feature maps have spatial resolutions of 4×4 and 1280 channels
b. Sequential model (keras model): Initializes a sequential model to stack layers sequentially
c. Flatten layer: Converts the 4×4×1280 feature maps into a flattened vector with 20480 elements. Output Shape: (None, 20480)
d. Dropout layer: Applies dropout regularization with a dropout rate of 0.5 to avoid overfitting. Randomly drops 50% of input units during training. Output Shape: (None, 20480)
e. Dense layer (output layer): Adds a dense layer with 12 units, corresponding to the number of output classes. Apply softmax activation so that for each of the 12 classes, obtain its class probability
f. Summary: Model architecture: Outline of the model architecture, including layer types, output shapes, and trainable parameters in the various layers. The total, trainable, and non-trainable parameters are 7,176,896, 245,772 and 6,931,124

Overall, this architecture is based on EfficientNetV2B1 with other modifications like adding a flattening layer, dropping some of its nodes to create an element to the neural network, and introducing the final dense output classification layer. In simple words, the stem of the pre-trained model provides powerful feature extraction capability, while the additional layers provide fine-tuning and adaptability to specific classification tasks (Zhang *et al*., 2021b).

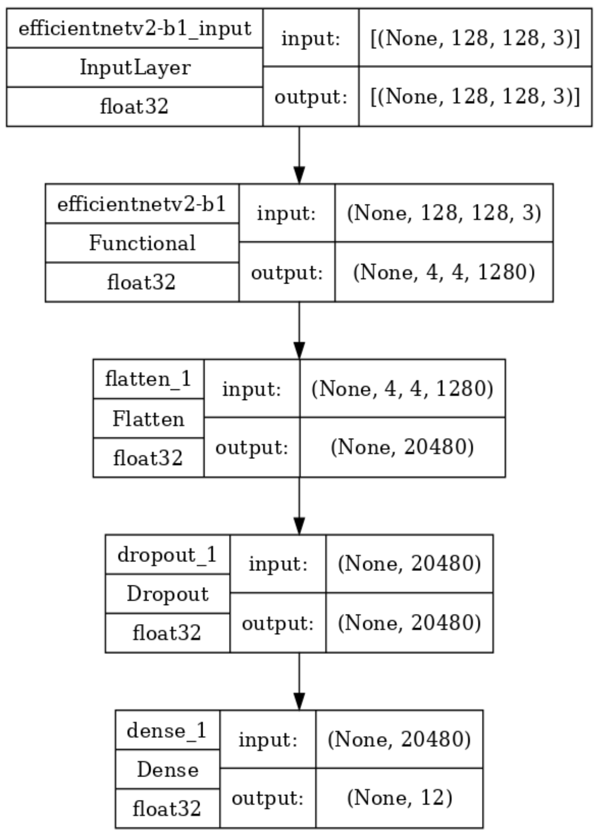


Fig. 3: Model architecture

DOI: 10.3844/jcssp.2024.1310.1321

Overfitting Analysis and Other Techniques

The training process was observed closely and losses accuracy tracked over training and validation. Our training curves were consistent both with training and validation so the signs of overfitting did not appear to be significant. In addition, we evaluated the model on a separate test set, providing an unbiased estimate of its generalizability.

Regularization: L2 regularization (weight decay) was introduced into the model with a view to penalizing large weights and thereby preventing overfitting.

Early stopping: Early stopping was performed based on the validation loss. Training was halted if the validation loss did not improve for a given number of epochs, putting the model in danger of overfitting instead of training data.

Results

The deep-learning model achieved promising results in the classification of all categories of waste. The EfficientNetV2B1 model has been trained on 10 epochs. The EfficientNetV2B1 model was trained for 10 epochs. An assessment of the performance of the EfficientNetV2B1 model is based on its ability to accurately categorize data from the input image in 12 specific varieties. In evaluating the validity of the proposed model, a range of performance indicators are used, as explained in the following sections A, B, and C.

Accuracy and Loss Analysis

In Fig. 4, the unbalanced dataset's training accuracy and loss and validation accuracy and loss are shown. Iterating over the training dataset once each epoch, the neural network model's training process involves 10 epochs. During training, for the training set and validation set, both the error and accuracy metrics were tracked.

During its first epoch, the training loss was 0.5180 with an accuracy of 87.54%, while the validation loss stood at 0.3576 with a precision of 92.82%. Both training and validation metrics improved significantly as training went on. By the final epoch, the training loss had reduced to 0.0992 and the accuracy was 98.14%, while the validation loss decreased even further to 0.5965 with a validation accuracy of 94.26%. This trend suggests that the model learned effectively from the training data and generalized well to unseen validation data. Analysis of the loss and accuracy trends over epochs shows that the model was not overfitting, as both training and validation statistics steadily improved. The final overall test performance also indicated high performance, with a test loss of 0.4296 and a test accuracy: Of 95.22%. The included data frame summarizes the loss and accuracy numbers for each epoch. Loss number drops while accuracy goes up, in a consistent and highly effective pattern over training. In general, the model was able to maintain high accuracy and effectively generalized to unseen data indicating its usefulness in classification tasks.

In Fig. 5, the training accuracy and loss and validation accuracy and loss of the balanced dataset are shown. At the beginning of the training, a TensorFlow caused layout optimization errors but the model kept on learning, updating its parameters, and eventually needed to be completed.

(a)

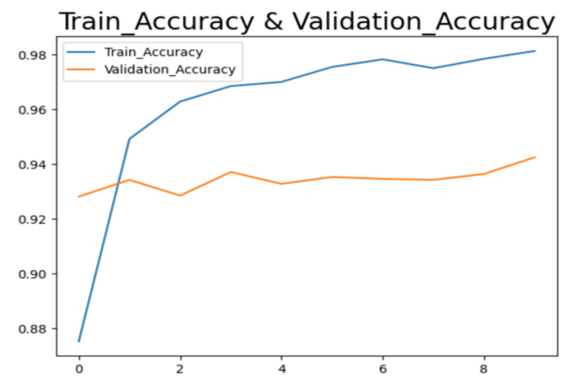
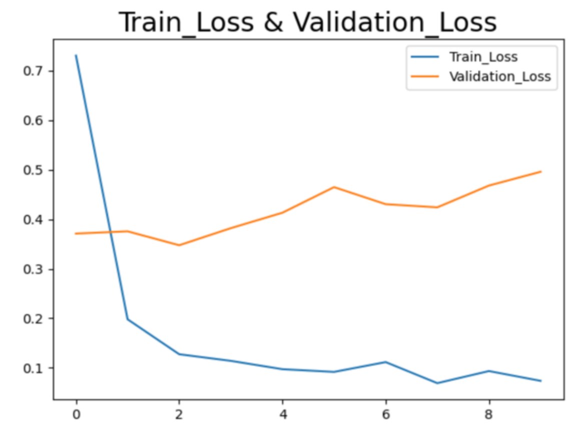
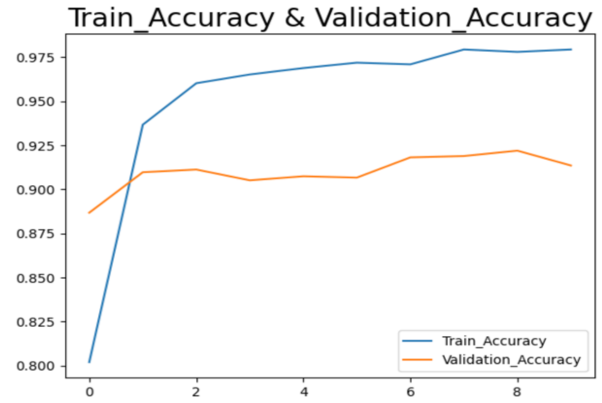


Fig. 4: (a) Training and validation curve of unbalanced dataset for loss; (b) Accuracy in training and validation curve of unbalanced dataset



(a)

DOI: 10.3844/jcssp.2024.1310.1321



(b)

Fig. 5: Training and validation curve of balanced dataset; (a) Loss and (b) Accuracy

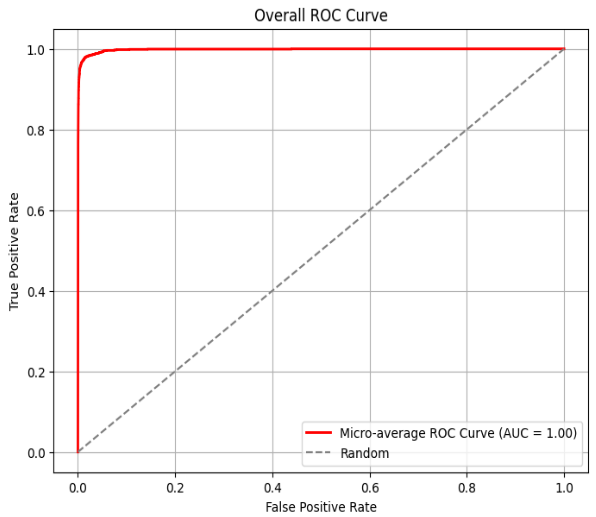
In the first epoch, the denoising units provide significantly reduced error with an average loss of 0.9142 and accuracy of 70.04% on the training set while only 0.2253 is lost per bit on validation. Although the loss continues to decrease and accuracy increases as training proceeds. By the last epoch, training loss was down to 0.0595 with an accuracy of 98.57% while validation losses increased to 0.4869 and validation accuracy grew slightly from 90.44%. Having tested this model on the test set, it found a test loss of 0.6019 and a test accuracy of 91.46%. This model transferred reasons rather well onto previously unobserved data, although at just a little lower than the validation set's performance level.

The data frame given just shows the loss and accuracy of every epoch, illustrating these trends that are downward on loss but upward for accuracy. Though it hit an error at first that was particular to TensorFlow, the performance of the model gradually improved, indicating it can learn and change in adversity.

ROC Curve Analysis

The Receiver Operation Characteristic curve (ROC) is a very useful tool for assessing the performance of the classification model, especially in situations such as garbage classification. A high Area Under Curve (AUC) value such as the exceptional AUC = 0.9998 observed for the overall ROC curve in Fig. 6(a) indicates good model performance in distinguishing between different categories of garbage items. That means it has a high true positive rate and keeps the false positive at low levels across all decision thresholds.

Through Fig. 6(b), the ROC curve performance for each class is further evaluated. The AUC values close to 1.0 (e.g., AUC = 1.0000) show that the near-perfect classification for Accuracy classes is something that the model separates effectively from others. Thus, those particular labels achieve almost near-to-perfect discrimination by a learner given their almost unique position within the graphs. Classes with slightly lower AUC values still demonstrate strong discrimination capability, if to a lesser extent. For example, AUC = 0.9999 and AUC = 0.9998 show this very clearly indeed. Even the classes with relatively lower AUC values, such as AUC = 0.9927, still mean quite good discrimination power. This also shows that overall classifier performance is sufficiently high for most garbage data items. However, there is still room to improve the model's capacity to distinguish these particular classes better. In total, ROC curve analysis allows the classifier to assess data types of different garbage items efficiently and accurately, promoting fine-tuning and optimal performance in model design (Zhang *et al*., 2021a).



(a)

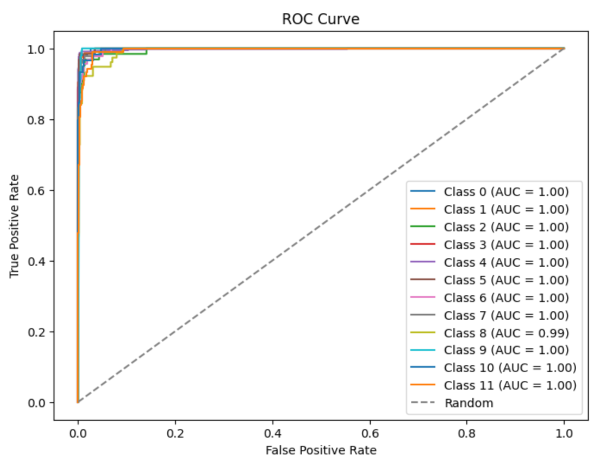


Fig. 6: (a) Overall ROC curve unbalanced dataset; (b) ROC curve analysis (class-wise) of unbalanced dataset

Figure 7 depicts the ROC curve analysis of the balanced dataset, which shows outstanding performance indicators. In Fig. 7(a), at its lowest point, this ROC curve performance achieves an AUC value of 1.00. This means that discrimination sums up life estimates. The balanced set of data in Fig. 7(b) is a more detailed explanation of ROC curve performance for individual classes. Even so, many classes showed admirable AUC values of 1.00. It seems that this model is extremely good at differentiating these classes from the others as demonstrated by these test results and the extensiveness of its training set. Ever Those classes with slightly lower AUC values, such as AUC = 0.97 and AUC = 0.99, still have good discriminability. They are signs of strong classification strength, with minimum overlap or cross-confusing between classes. In total, the ROC curve analysis for the balanced dataset shows the talent of the model to classify the multitude of garbage items with highly accurate results. It also demonstrates the technique's applicability and dependability outside real-world situations.

Confusion Matrix

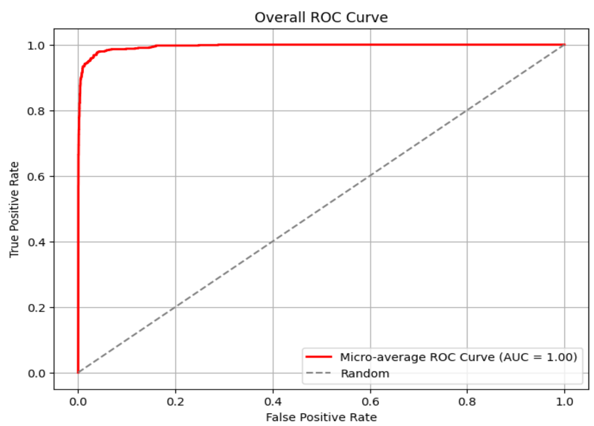
The confusion matrix is a tabular representation in Fig. 8 used to assess the effect of the classification model on the unbalanced data set. Each matrix is organized with rows standing for the actual category labels and columns for the expected category labels. Values within the matrix cell tell how many cases come from a specific combination of categories both in reality and forecast (Malik *et al*., 2023).

Observations made in that way are very useful. For Battery, the model correctly predicted 93 instances of battery. However, it misclassified 3 instances as brown-glass and 1 instance each as metal, paper, and white-glass. In turn biological received out of 84 instances of battery, the model accurately classified 80 instances. Nonetheless, it misclassified 1 instance each as clothes, plastic, and trash. Brown-glass correctly predicted 55 instances of Brown glass, but misclassified 7 instances as shoes, trash, and white-glass. In cardboard out of 90 instances, the model correctly classified 87 instances as cardboard. It misclassified 2 instances as cloths and 1 instance as metal. Cloths accurately predicted 560 instances of cardboard, with 6 instances misclassified as shoes.

In green-glass out of 63 instances, the model correctly classified 60 instances as clothes. It misclassified 1 instance each as biological and white-glass. Metal correctly predicted 83 instances of metal, but misclassified 7 instances as brown-glass, cloths, and white-glass. In paper out of 95 instances, the model accurately classified 84 instances as paper. However, it misclassified 5 instances as cardboard, cloths, plastic, and white-glass. Plastic correctly predicted 69 instances of plastic, but misclassified 8 instances as cardboard, paper, shoes, and class 11. In shoes out of 178 instances, the model accurately classified 174 instances as shoes. It misclassified 3 instances as brown-glass and 1 instance each as metal and paper. Trash correctly predicted 55 instances of trash. However, it misclassified 2 instances as battery and 1 instance each as cloths, plastic, and white-glass. In white glass out of 78 instances, the model accurately classified 75 instances as white glass. Nonetheless, it misclassified 1 instance each as brown glass, cardboard, cloths, metal, and plastic.

Overall, the confusion matrix provides a comprehensive overview of the model's performance for each class in the unbalanced dataset, highlighting areas of accurate classification and instances of misclassification that require further attention.

Figure 9 shows the confusion matrix used to evaluate the performance of a classification model on balanced datasets.



(a)

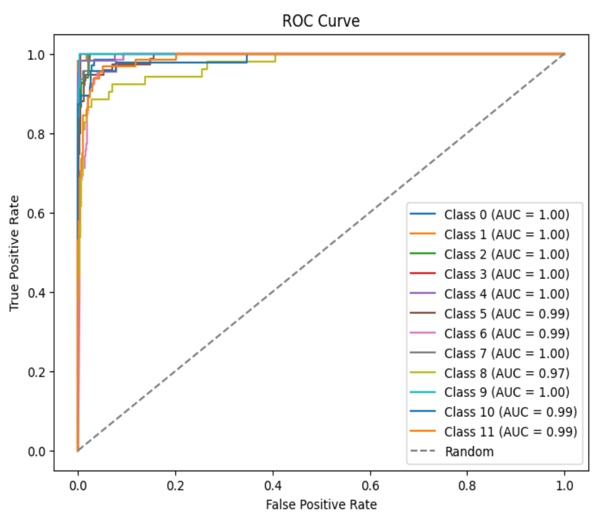


Fig. 7: (a) Overall ROC curve balanced dataset; (b) ROC curve analysis (class-wise) balanced dataset

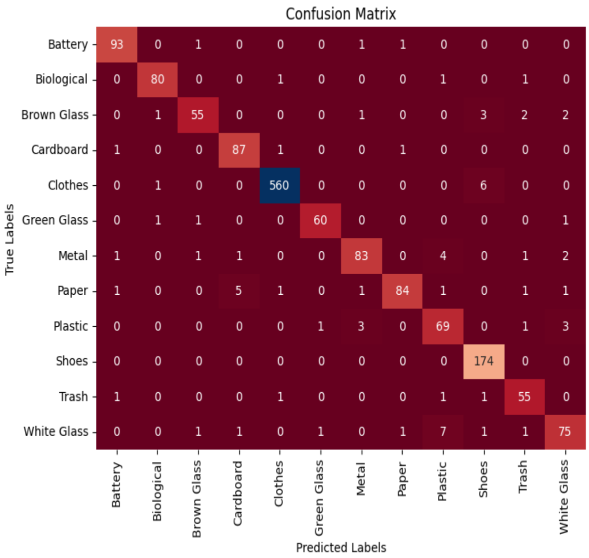


Fig. 8: Confusion matrix for unbalanced dataset

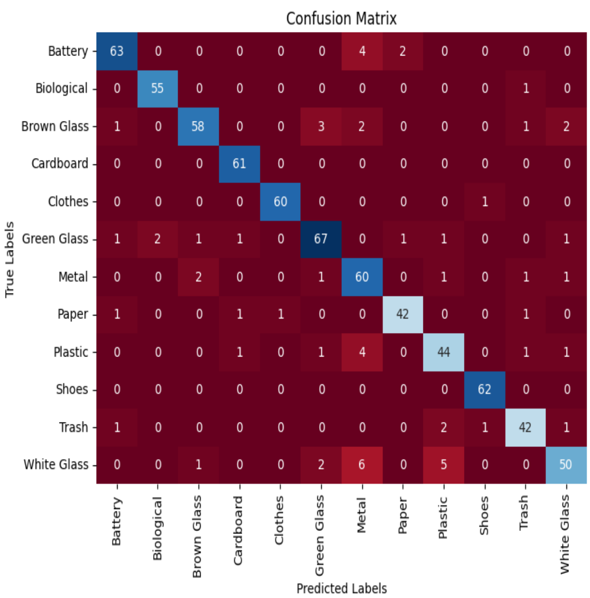


Fig. 9: Confusion matrix for balanced dataset

In battery, the model accurately predicted 63 cases out of 64, so there was just one mistake. Notably, its accuracy in distinguishing battery images is higher than with biological. Nevertheless, for biological the high accuracy persisted at 55 correct predictions and merely two misclassifications. Furthermore, brown glass had a slightly lower rate of accuracy, with 58 correct predictions and six misclassifications. This indicates that while doing well overall in this category of image recognition problems, brown-glass is one area where this model does have some difficulty distinguishing between classes. Moving on to cardboard, the model achieved perfect accuracy, correctly identifying all 61 instances of cardboard images. However, in cloths, which is also a classification of solids, the performance was excellent; with 60 correct predictions and only one misclassification. Instead, green-glass showed some variability, with 67 correct predictions but five misclassifications, indicating a slightly higher error rate compared to other classes.

Metal demonstrated similar performance 60 correct predictions and 5 misclassifications. Paper had a lower misclassification rate, with only two instances out of 45 being misclassified. Plastic and shoes were relatively good performers as well 44 and 62 correct predictions, respectively, and just three misclassifications each. Last trash had 42 of 44 right, whereas white glass was correct in 50 out of 51 instances. Overall, despite its good performance on all the lightly balanced categories in the dataset, this computer program made itself felt in a few areas, with errors and those areas where mistakes can only be avoided by adjustments. In the confusion matrix metal, plastic, and white glass are highly misclassified but it's because both balanced and unbalanced datasets suggest that the model effectively learned the underlying patterns rather than memorizing specific instances.

Performance Parameters

The performance of the model trained on the unbalanced dataset is defined in Table 1. This system achieved an accuracy of 0.9522, which indicates that it will correctly classify garbage images at a rate of more than 95%. Similarly, precision, which measures the accuracy of predictive cancer classifications, came in at 0.9528. This indicates the system's ability to effectively identify spam mail. Anyway, recall ratings are several positive affirmative classifications divided by total true positives in a vast majority sense quite similar to accuracy it would exist without a huge leap between voted access and transparent analysis of information. And F1-scores, which balance recall precision, testified again at 0.9521 that the first half of this process is nearly harmless to our interests whatsoever. In summary, all these indicators of model performance on the unbalanced data set underscore faithful patterns and effectiveness in generating benchmark images for repeated testing.

On the balanced dataset represented in Table 2, the model achieved excellent accuracy. Its score was 0.9146, meaning that the model is capable of accurately classifying garbage images yet maintains a balance across all classes of things precisely it achieved an index of 0.9161, which indicates how good it is at not leaving a single duplicate in the result dataset. While looking across our entire dataset while generating these classifications for free from false disputed information based on actual. This is the value of recall essentially checking out all positive instances in our data set before counting them and if it came close to generating accuracy values as 0.9146, then that demonstrates its effectiveness at identifying most of the real positives within a comprehensive picture that covers all potential negatives. In addition, the F1-score, a harmonized mix of recall and precision, came in at 0.9144 showing us its well-timbered response between these two aspects. This set of performance measures as a whole reflects the model's strong performance on the balanced dataset, showing its capabilities to accurately classify garbage images in many different classes and reliably predict future classifications whilst retaining balance.

Table 1: Performance parameter for unbalanced dataset

Name of class	Performance parameter of the unbalanced dataset			
	Precision	Recall	F1-score	Accuracy
Battery	0.96	0.97	0.96	0.95
Biological	0.96	0.96	0.96	0.95
Brown glass	0.93	0.86	0.89	0.95
Cardboard	0.93	0.97	0.95	0.95
Cloths	0.99	0.99	0.99	0.95
Green glass	0.97	0.95	0.96	0.95
Metal	0.93	0.89	0.91	0.95
Paper	0.97	0.88	0.92	0.95
Plastic	0.83	0.90	0.86	0.95
Shoes	0.94	1.00	0.97	0.95
Trash	0.89	0.93	0.91	0.95
White glass	0.89	0.85	0.87	0.95

Table 2: Performance parameter for balanced dataset

Name of class	Performance parameter of the balanced dataset			
	Precision	Recall	F1-Score	Accuracy
Battery	0.94	0.91	0.93	0.91
Biological	0.96	0.98	0.97	0.91
Brown glass	0.94	0.87	0.90	0.91
Cardboard	0.95	1.00	0.98	0.91
Cloths	0.98	0.98	0.98	0.91
Green glass	0.91	0.89	0.90	0.91
Metal	0.79	0.91	0.85	0.91
Paper	0.93	0.91	0.92	0.91
Plastic	0.83	0.85	0.84	0.91
Shoes	0.97	1.00	0.98	0.91
Trash	0.89	0.89	0.89	0.91
White glass	0.89	0.78	0.83	0.91

State-of-the-Art Comparison

In studies of garbage sorting launched recently by Majchrowska *et al*. (2022), the highest model accuracy is 75%, a result that shows how data variability and complexity make formidable challenges. Using the EfficientNet-B1 model, the research has surpassed these existing systems by a wide margin. For its part, our model has an accuracy of 95% on a non-unbalanced dataset and 91% for balanced ones. Clearly, we see that it works well across distribution patterns of the data. Moreover, the model's average ROC value of 0.99 between all classes demonstrates its excellent ability to differentiate between categories. This significant improvement underscores how well EfficientNet-B1's architecture captures the intricate features that need to be found and optimal classification performance techniques. Set against the background of existing approaches, our findings suggest that EfficientNet-B1 supplies a promising solution to the problem of garbage classification, with markedly better precision and reliability in all cases.

Conclusion

The research shows how EfficientNetV2B1 deep learning architecture is used for the classification of garbage images. Utilizing both unbalanced and balanced datasets that contained all 12 types of garbage. The study provided a comprehensive assessment of the model's performance and the significance of balancing the dataset. On the other hand, results show that the model trained in this study has high accuracies above 90%. Furthermore, when evaluating the test sets on the model they all performed soundly, which means they are effective in real-world scenarios.

Comprehensive evaluations, including confusion matrices, classification reports, and ROC curves, offer a detailed assessment of the model's performance in differentiating between different classes of garbage. The research suggests that the EfficientNetV2B1 architecture is a promising way to solve problems such as the classification of waste materials with both balanced and unbalanced datasets. Deep learning mechanisms in this approach not only outperform traditional methods by a large margin but also have a substantial impact on environmental conservation and waste management projects. As a result, efficient and sustainable solutions can be worked out. By further study and improvement of deep learning, models applied to garbage classification tasks in the future, this may be able to help make the environment healthier and create a cleaner land for the children.

Future Scope

In the future, the incorporation of cross-validation can be used to further validate the robustness of our model.

Cross-validation would provide a more reliable estimate of the model's performance and help in identifying any potential overfitting issues. Future studies can also further improve the accuracy in classifying the less accurately predicted classes by analyzing the features or feature maps. Evaluating the model's performance over a longer period in a real operational environment is crucial for understanding the stability of the model's predictions and its adaptability to changes in waste types and disposal practices. Future work can focus on deploying the model in a live operational setting to monitor its long-term performance and robustness. This will provide valuable insights into the model's practical applicability and help identify any necessary adjustments to enhance its effectiveness in real-world scenarios.

Acknowledgment

The authors acknowledge organizations for the support provided for carrying out the research work in the stipulated time.

Funding Information

The authors should acknowledge the funders of this manuscript and provide all necessary funding information.

Author's Contributions

All authors equally contributed to this study.

Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved.

References

Abdallah, M., Abu Talib, M., Feroz, S., Nasir, Q., Abdalla, H., & Mahfood, B. (2020). Artificial intelligence applications in solid waste management: A systematic research review. *Waste Management*, *109*, 231–246. https://doi.org/10.1016/j.wasman.2020.04.057

Chen, Z., Yang, J., Chen, L., & Jiao, H. (2022). Garbage classification system based on improved ShuffleNet v2. *Resources, Conservation and Recycling*, *178*, 106090. https://doi.org/10.1016/j.resconrec.2021.106090

Farooq, M., Cheng, J., Khan, N. U., Saufi, R. A., Kanwal, N., & Bazkiaei, H. A. (2022). Sustainable Waste Management Companies with Innovative Smart Solutions: A Systematic Review and Conceptual Model. *Sustainability*, *14*(20), 13146. https://doi.org/10.3390/su142013146

Ihsanullah, I., Alam, G., Jamal, A., & Shaik, F. (2022). Recent advances in applications of artificial intelligence in solid waste management: A review. *Chemosphere*, *309*, 136631. https://doi.org/10.1016/j.chemosphere.2022.136631

Jin, S., Yang, Z., Królczykg, G., Liu, X., Gardoni, P., & Li, Z. (2023). Garbage detection and classification using a new deep learning-based machine vision system as a tool for sustainable waste recycling. *Waste Management*, *162*, 123–130. https://doi.org/10.1016/j.wasman.2023.02.014

Li, X., Li, T., Li, S., Tian, B., Ju, J., Liu, T., & Liu, H. (2023). Learning fusion feature representation for garbage image classification model in human–robot interaction. *Infrared Physics & Technology*, *128*, 104457. https://doi.org/10.1016/j.infrared.2022.104457

Lou, L., & Gou, N. (2023). Domestic waste image classification algorithm based on improved EfficientNetV2. *2023 8th International Conference on Intelligent Computing and Signal Processing (ICSP)*, 1374–1377. https://doi.org/10.1109/icsp58490.2023.10248442

Majchrowska, S., Mikołajczyk, A., Ferlin, M., Klawikowska, Z., Plantykow, M. A., Kwasigroch, A., & Majek, K. (2022). Deep learning-based waste detection in natural and urban environments. *Waste Management*, *138*, 274–284. https://doi.org/10.1016/j.wasman.2021.12.001

Malik, M., Prabha, C., Soni, P., Arya, V., Alhalabi, W. A., Gupta, B. B., Albeshri, A. A., & Almomani, A. (2023). Machine Learning-Based Automatic Litter Detection and Classification Using Neural Networks in Smart Cities. *International Journal on Semantic Web and Information Systems*, *19*(1), 1–20. https://doi.org/10.4018/ijswis.324105

Meng, S., & Chu, W.-T. (2020). A Study of Garbage Classification with Convolutional Neural Networks. *2020 Indo Taiwan 2nd International Conference on Computing, Analytics and Networks (Indo-Taiwan ICAN)*, 152–157. https://doi.org/10.1109/indo-taiwanican48429.2020.9181311

Mohamed, M. (2021). *Garbage Classification (12 classes)* [dataset].

Nnamoko, N., Barrowclough, J., & Procter, J. (2022). Solid Waste Image Classification Using Deep Convolutional Neural Network. *Infrastructures*, *7*(4), 47. https://doi.org/10.3390/infrastructures7040047

Rismiyati, Endah, S. N., Khadijah, & Shiddiq, I. N. (2020). Xception Architecture Transfer Learning for Garbage Classification. *2020 4th International Conference on Informatics and Computational Sciences (ICICoS)*, 1–4. https://doi.org/10.1109/icicos51170.2020.9299017

Song, Y., He, X., Tang, X., Yin, B., Du, J., Liu, J., Zhao, Z., & Geng, S. (2024). DEEPBIN: Deep Learning Based Garbage Classification for Households Using Sustainable Natural Technologies. *Journal of Grid Computing*, *22*(1), 2. https://doi.org/10.1007/s10723-023-09722-6

Verma, K., Bhardwaj, S., Arya, R., Islam, M. S. U., Bhushan, M., Kumar, A., & Samant, p. (2019). Latest Tools for Data Mining and Machine Learning. *International Journal of Innovative Technology and Exploring Engineering*, *8*(9S), 18–23. https://doi.org/10.35940/ijitee.i1003.0789s19

Wang, Q., & Wen, B. (2024). Design and implementation of garbage classification and detection system based on YOLOv8. *Ninth International Symposium on Sensors, Mechatronics and Automation System (ISSMAS 2023)*, 456–460. https://doi.org/10.1117/12.3015103

Wang, W. (2024). Garbage Image Classification based on Deep Learning. In *Proceedings of the 2023 International Conference on Data Science, Advanced Algorithm and Intelligent Computing (DAI 2023)* (pp. 400–410). Atlantis Press. https://doi.org/10.2991/978-94-6463-370-2_42

Yang, J., Zeng, Z., Wang, K., Zou, H., & Xie, L. (2021). GarbageNet: A Unified Learning Framework for Robust Garbage Classification. *IEEE Transactions on Artificial Intelligence*, *2*(4), 372–380. https://doi.org/10.1109/tai.2021.3081055

Yuan, J., Nan, X., Li, C., & Sun, L. (2020). Research on Real-Time Multiple Single Garbage Classification Based on Convolutional Neural Network. *Mathematical Problems in Engineering*, *2020*(1), 1–6. https://doi.org/10.1155/2020/5795976

Zhang, L., Liu, G., Li, S., Yang, L., & Chen, S. (2022). Model framework to quantify the effectiveness of garbage classification in reducing dioxin emissions. *Science of the Total Environment*, *814*, 151941. https://doi.org/10.1016/j.scitotenv.2021.151941

Zhang, Q., Zhang, X., Mu, X., Wang, Z., Tian, R., Wang, X., & Liu, X. (2021a). Recyclable waste image recognition based on deep learning. *Resources, Conservation and Recycling*, *171*, 105636. https://doi.org/10.1016/j.resconrec.2021.105636

Zhang, Q., Yang, Q., Zhang, X., Bao, Q., Su, J., & Liu, X. (2021b). Waste image classification based on transfer learning and convolutional neural network. *Waste Management*, *135*, 150–157. https://doi.org/10.1016/j.wasman.2021.08.038

Zhao, Y., Huang, H., Li, Z., Yiwang, H., & Lu, M. (2022). Intelligent garbage classification system based on improve MobileNetV3-Large. *Connection Science*, *34*(1), 1299–1321. https://doi.org/10.1080/09540091.2022.2067127