Article
JOURNAL OF UNIVERSITY OF BABYLON
For Pure and Applied Sciences (JUBPAS)
Vol.33 ; No.4 | 2025

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

Developing an Intelligent Waste Classification System Based on Hybrid Deep Transfer Learning Model

Zainab Salah Abdulhassan[1]
1, Faculty of Engineering, Islamic University in Lebanon, Beirut, Lebanon.
Corresponding author email: **awdem3037@gmail.com**

تطوير نظام ذكي لتصنيف النفايات اعتماداً على نموذج هجين للتعلّم العميق بالنقل المعرفي

زينب صلاح عبد الحسن[1]
1 كلية الهندسة، الجامعة الإسلامية في لبنان، بيروت، لبنان

Accepted:	2/12/2025		Published:	31/12/2025

ABSTRACT

Background:

The world is more and more interested in sustainable waste management, for which intelligent systems performing automatic sorting of different kinds of waste are needed.

Materials and Methods:

In this study, the hybrid deep learning model is composed of transfer learning with the DenseNet121 model, convolutional block attention module (CBAM) and a simple recurrent neural network (Simple RNN) to enhance feature representation and sequential dependencies of waste images. In order to handle the class imbalance and enrich the data set, preprocessing techniques like normalization, murder and brightness adjusted and strategic data bird mooding (rotating, flipping) were performed.

Results:

The experimental results on a multi-class waste dataset with twelve categories, namely batteries, biowaste, glass types, cardboard, clothing, metals, paper, plastics, shoes, regular waste, and white glass, demonstrated that the model can yield good performance, showing well-balanced classification across all classes. This model reduces the training and parameter size remarkably by using pre-trained weights, and the proposed CBAM mechanism concentrates on learning significant features. The proposed hybrid model achieved an accuracy of 99%, outperforming conventional baselines.

Conclusion:

The experimental results demonstrate the potential to apply the proposed approach in real-world smart recycling systems, which is an effective way to realize waste sorting and environmental protection. Future work will investigate generalization of this technique to other environmental areas, and validate its performance on a larger and more diverse dataset.

Key words:

transfer learning, deep learning, image processing, feature enhancement, and data augmentation.

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية

INTRODUCTION

Garbage from a basic particle of the urban facilities and provide an important support role for the waste sorting, pick-up and treatment in the city. The way they function is inextricably linked with public safety and the visual coherence of city streets. Conventional ways of monitoring such assets—mainly manual field observations have provided high accuracy, but have their own essential limitations like inefficiency, labor-intensive work, and delayed response [1]. With improvements in computing systems and mobile mapping devices, street view imagery has become an important data resource for locating and analyzing garbage bins. Nevertheless, challenges remain, for instance, because of heavy occlusion, heterogeneous geometric distortions and scale variations in street-level images [2]. All these factors greatly degrade the precision of the automatic detector and become a big stumbling block for robust object extraction. The object of garbage bin identification is a professional subfile of urban scene understanding in computer vision, whose breakthrough depends mainly on the complexity of the scene and the robustness of detecting the structure [3].

Waste sorting has grown in popularity in recent years. According to a 2018 World Bank report, the world generated approximately 242 million tons of plastic waste in 2016, equivalent to about 12% of total solid waste. By 2050, the annual global waste volume is expected to reach 3.40 billion tons, a significant increase from its current level of 2.01 billion tons [4].

In municipal solid waste recovery, particularly in hazardous working conditions, with skin, respiratory, and digestive injuries and diseases, very high rates of injury have been recorded, although sanitation workers and waste collectors play a significant role in this process in India. Waste sorting has received significant attention from the Chinese government and public. In the treatment process at large waste treatment plants, sorting and separation are the first two steps.[5].

These plants use manual sorting through pipes, which entails many difficulties, including a harsh working environment, intense physical effort, and low sorting efficiency. Furthermore, manual sorting separates only very small amounts of recyclable or harmful waste from the large quantities of waste; most of it ends up in landfills, leading to resource waste and environmental pollution [6].

The early works mostly used hand-engineered feature-based algorithms such as Local Binary Patterns (LBP), Deformable Part Models (DPM), Oriented FAST and Rotated BRIEF (ORB)[7-8]. However, as far as they are conceptually simple, those classical methods cannot effectively extract high-level semantic context and are insensitive to varied geometric transformations, which have been identified as challenging aspects under real-world detection scenarios. Deep learning ,especially Convolutional Neural Networks (CNNs) is changing the landscape of visual perception problems like face recognition, autonomous driving and industrial quality inspection, since it makes parameter sharing and scale invariance possible[9]. However, even object detectors based on CNNs still suffer from some drawbacks in bin detection tasks in real-world situations, especially in the detection of occluded objects, body deformation and multi-scale distribution [10]. In this paper, we investigate the effectiveness of deep learning models based on DenseNet121 with CBAM and SimpleRNN for solid household waste classification.Therefore, this study seeks to answer the following research question:

“How effective is a hybrid model combining DenseNet121, CBAM, and SimpleRNN in improving the accuracy and robustness of household waste classification in real-world urban environments?

The following are the major findings of this study that make it different from earlier works: First, a new hybrid architecture has been proposed in which DenseNet121 is combined with CBAM attention

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية

ISSN: 2312-8135 | Print ISSN: 1992-0652

info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

mechanism and SimpleRNN layer. Such an integration was never explored earlier for household waste classification. Second, a large multi-source dataset is created by merging three public datasets to address common issues found in most existing datasets i.e., class imbalance, low class resolution, and small number of categories. Thirdly[1] ,the proposed model achieved better classification performance as well as high feature extraction capability which provides empirical evidence on effectiveness about combination between attention modules & lightweight sequential modeling thus making proposed framework solution both powerful yet practical one intelligent waste classification system.

The remainder of this study consists of several parts section 3 discusses previous studies, while Section 4 discusses the research methodology. Section 5 presents and details the test results, while Section 6 concludes the study by drawing conclusions and offering recommendations for future work.

Literature review

In this section, we review previous studies that dealt with several garbage classification models using different methodologies, such as image processing, deep learning, and Raman spectroscopy. Most researchers noted that other techniques, such as active marker smoothing and Raman spectroscopy using PLS-DA emphasize the importance of leveraging advanced models and biomarkers to enhance classification sensitivity and specificity. Challenges such as dataset dependence, computational costs, overfitting, and limited generalization to other skin conditions remain. Without thinking about using CNN, ResNet50, InceptionV3, and VGG16.

- The researchers aim to present an advanced model design methodology based on GCNet, an improved ShuffleNet v2 model, combined with PMAM and FReLU attention activation functions, with transfer learning further applied to increase model accuracy. This model was trained using self-generated data and tested on CIFAR-100 and Tiny-ImageNet to demonstrate the model's effectiveness in the real world and for general use. The result was that GCNet significantly outperformed popular models such as ResNet50 and MobileNet v2 in terms of accuracy and execution speed, achieving 97.9% accuracy on self-generated data and an execution time of just 105 ms on the Raspberry Pi, further enhancing its high efficiency **[11]**.

- The authors fine-tuned to avoid the issue of biased or mislabeled noise samples. The trained model was both trained and tested on real-world data. The excellent accuracy resulted in first place in the 2019 Huawei Cloud Waste Classification Competition, and a high level of flexibility and scalability was also achieved. The advantages of the proposed model are its flexibility and generalization capability for new categories, robustness for corrupted data and high classification performance. Its primary limitations are its representation quality in the latent space that strongly depends on the data and potentially influences classification accuracy for complex or visually tangled cases **[12]**.

-The authors fine-tuned a pre-trained MobileNetV3 model to a restricted set of your training data using deep learning in the PyTorch framework. This required paying attention to spatial locations using CBAM and MISH on information from images. The fully connected layers were exchanged to global mean pooling, thereby computing a smaller amount of parameters with better accuracy. As a result, GMC-MobileNetV3 with a classification accuracy of

96.55% can be achieved, which was trained on the selfconstructed dataset of authors, performing 3.6% better than the original one. It reduced the number of parameters to 0.64 million, the memory usage by 56.6%, and the time of image recognition to 26.4 ms[13].

-The pollution from landfills proved to be as toxic as tequila or liquor. A new technique was presented for estimating the contribution of primary contributors in a multi-source scenario. Data on air pollutant concentration (H_2S and volatile organic compound, VOC) was included, before and after it began waste classification. The outcome was a 48% reduction in H2S concentration, 43% in VOCs, a peak frequency educed to 0.01, and a 56% reduction in ozone generation[14].

-The authors employed a MobileNetV3-Large model for image classification and benefited from state-of-the-art components including separable convolutions, inverted residuals, Lightweight attention mechanism and Hard-Swish activation function. LSTM model was also applied in text classification based on word representation to enhance feature extraction. Test results showed that the system could carry out intelligent waste classification and deep learning method was applied, in which the accuracy for waste image classification was 81% and that for Residuals type text classification was 97.61% [15].

- The researchers used a analytical technique to contrast levels of individual chemical elements and leaching behaviour between ash produced pre- and post-intervention, also comparing laboratory data with chemical analysis of the ash components. The test results indicated that the average Cl composition between 17.43% and 28.63% is present; the CaO composition is relatively stable, the CaClOH composition appears, and ten years ago was not built. Levels of heavy metals, for example lead, copper, and zinc, also dropped, suggesting the amelioration of hazardous waste disposal [16].

- The classification model GNet, constituted by the transfer learning and the multi-innovative MobileNetV3, is built to optimize the classification precision. A GUI (General User Interface) was also developed using Python and QT for the control and monitoring of the system. The system was evaluated on the Huawei Waste Classification Challenge dataset. According to the test, the static system after the imaging processing has a high classification accuracy of 92.62\% with a response time of 0.63 second, which showed that it has high performance [17].

- The investigators used a mix of both a pre-trained model and a custom convolutional neural network, having first balanced the dataset by deploying data augmentation techniques. It consists of three fully connected layers having 1024, 512, and 12 cells, respectively, with the Mis activation function. The findings pointed to the superiority of the 69 model in terms of classification accuracy- 95.58%, much faster training times, and far fewer parameters than other models that had been used previously [18].

- A very smart model was created in this research to classify household waste efficiently and effectively. That's also a new model it has learned, GFN, with contemporary tools like ResNet, ViT transformers, plus CBAM and PPM mechanisms. This improves visual feature extraction. The work used dedicatedly collected waste classification datasets, which were divided into several class types (e.g., batteries, biowaste, plastic, paper, and metal). PPM is used to seize the multi-scale information on the waste , and CBAM detects vital visual

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

characteristics by attending to the feature. The comparison results showed the excellence of the proposed models against other SwinT, MobileNet, and VGG models and the best classification accuracy of 96.54% and weighted precision, recall and F1 above 96%. The reliability of the model to concentrate on the critical region of the image was verified through the Grad-CAM analysis [19].

The approach consisted of incorporating the VGG-16 model with our own CNN to analyze images obtained from live video streams using an IP camera to the OpenCV-based library. Based on a dataset of 12 waste categories, all the automatically classified images were saved in Excel format for traceability and to facilitate the integration with a trained chatbot that is able to respond to user queries regarding waste sorting and recycling options. The findings showed a classification accuracy of 96% which shows the system's high effectiveness in identify various waste materials in real time[20].

– The study was done with a multi-stage approach. The first part of the algorithm was developed using a "garbage in, garbage out" data set consisting of 25,000 images of different types of waste. This paper proposes a novel model, GCDN-Net, a deep neural network for single-label and multi-label classification. Through single-label classification, it is determined whether an image includes waste, and through multi-label classification, one or multiple classes of waste are classified for one or multiple images. The results supported GCDN-Net better than the state-of-the-art counterparts with classification accuracy, precision, recall, F1, and specificity exceeding 95.54%. It has an average MAP of 0.69 and F1=75.01% for multi-label classification [21].

- The methodology begins by developing a network model called FConvNet, a simple yet striking model that fills gaps in previous studies that focused on industrial and medical waste, neglecting the more common household waste. Previous studies used self-constructed datasets, which did not generalize the results obtained in any way. The efficiency of the proposed model relied heavily on a convolutional block design that utilized a built-in convolution to maximize the use of both spatial features, along with a lightweight symmetric architecture to extract different representations from the data. Public datasets were used to train and test the model, which achieved classification accuracy of 95%, 95%, and 97%, outperforming all baseline models in the same comparison. To use the model in practice and collect new images for future tuning, an online classification system was developed using the WeChat mini app platform[22].

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com
مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية

ISSN: 2312-8135 | Print ISSN: 1992-0652

info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

Table 1. Summary of related works in the current study.

Ref & author	Models	Dataset	Results
Wu et al. [11]	GCNet (ShuffleNet v2 + PMAM + FReLU, TL)	Self + CIFAR-100, Tiny-ImageNet	97.9% acc, 105 ms on Raspberry Pi, better than ResNet50 / MobileNet v2
Yang et al. [12]	Fine-tuned pre-trained model	Huawei 2019 Competition	1st place, robust to mislabels, dataset dependent
Tian et al. [13]	GMC-MobileNetV3 + CBAM + MISH	Huawei dataset	96.55% acc, 56.6% ↓memory, 26.4 ms recognition time
Gao et al. [14]	-	Field environmental data	H_2S ↓48%, VOC ↓43%, Ozone ↓56% after classification
Zhao et al. [15]	MobileNetV3-L + LSTM	Image + text data	Img: 81%, Text: 97.61%
Liu et al. [16]	-	Ash samples	↓ heavy metals (Zn, Pb, Cu), ↑ chlorides after waste policies
Fu et al. [17]	GNet (MobileNetV3 + TL)	Huawei dataset	92.62% acc, 0.63 s response, GUI for sorting
Haque et al. [18]	Hybrid CNN + MIS activation	Not specified	96.58% acc, faster training, fewer parameters
Wang et al. [19]	GFN (ResNet + ViT + CBAM + PPM)	Not specified	96.54% acc, all metrics >96%, Grad-CAM verified
Musham et al. [20]	VGG-16 + Custom CNN	Live video (OpenCV)	96% acc for 12 classes, chatbot integrated
Hossen et al. [21]	GCDN-Net	GIGO (25,000 images)	Single-label >95.54%, Multi-label mAP=0.69, F1=75%
Liang and Guan [22]	FConvNet (lightweight CNN block)	Household waste	95-97% acc, deployed via WeChat mini-app

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

METHODOLOGY

This study presents the design of an artificial intelligence model as shown in fig 1 capable of classifying household solid waste using deep learning techniques. The first phase involved collecting data from 12 distinct categories of waste (such as paper, plastic, glass, batteries, organic waste, among others), followed by a comprehensive preprocessing pipeline that included resizing, color normalization, and applying data augmentation techniques such as rotation, cropping, flipping, and zooming to expand the dataset and enhance the model's generalization capability. In the second phase, a DenseNet121 network was employed as the backbone architecture with the top layer removed. The model was further enhanced by integrating an advanced attention mechanism, namely the Convolutional Block Attention Module (CBAM), which incorporates both spatial and channel attention to improve the extraction of discriminative features. Additionally, a SimpleRNN layer was included after global pooling to capture dependencies within the extracted features. The model was trained using the Adam optimizer, with hyperparameters tuned to 20 epochs and a batch size of 32, while monitoring performance through the categorical cross entropy loss function. The evaluation phase utilized metrics such as classification accuracy, a confusion matrix, and a detailed classification report to assess the overall performance of the model. The experimental results demonstrated the model's high efficiency in classifying waste images, confirming its effectiveness for practical applications in waste management.

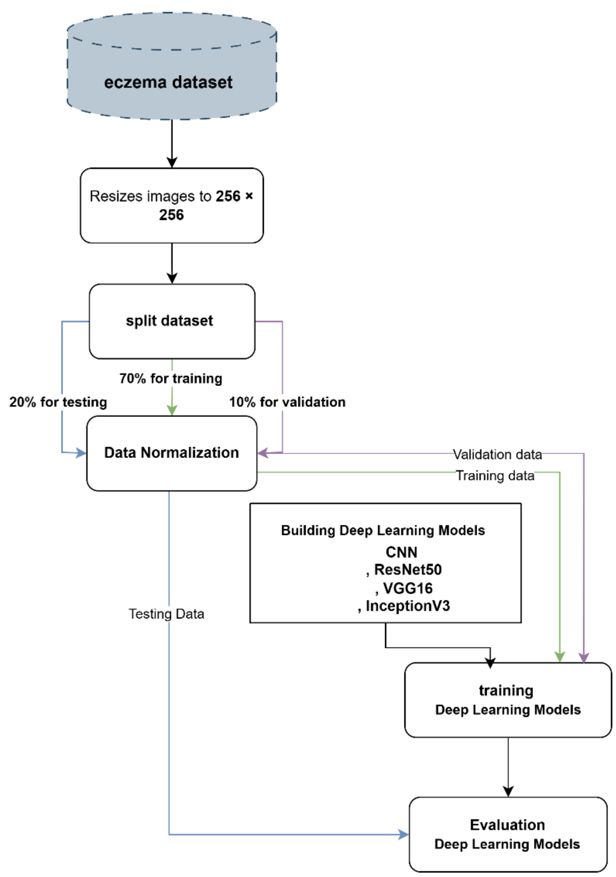

Figure1.the proposed system.

Dataset Selection

We integrated three public datasets, including the TrashNet, Huawei Cloud Waste Classification Dataset and Garbage Collection Dataset. This dataset covers the weaknesses for individual dataset, e.g. low class resolution, class imbalance to some extent, resulting in an all-around dataset that simulates the diversity and complexity of real garbage much better.

The TrashNet dataset was created as a project at Stanford University, and includes 2,527 RGB colour images from six waste categories: cardboard (403 images), glass (501 images), metal (410 images), paper (594 images), plastic (482 images), and Residuals (137 images). These photographs were taken under controlled laboratory lighting and were all of resolution 512×384 pixels [23].The Huawei Cloud dataset, which was published in Huawei Cloud AI Waste Classification Competition 2020 The dataset contains 14,926 images for four major types and 44 subordinate categories with the format of Pascal VOC and offered great diversity within types and high-definition images [24].The waste collection dataset consists of 15,150 images in 12 specific household waste categories: cardboard, paper, metals, biowaste, plastics, green glass, brown glass, white glass, clothing, shoes, batteries, and residuals. It was especially developed to help achieving a correct waste classification in order to optimize recycling processes [25]. Integration of the data-sets is in itself helping pave the way for superior, scalable, smart waste sorting systems, a key requirement in the advance of recycling and sustainable waste management. Table 2 and figure 2 show the datasets images .

Table 2.the samples of dataset

Dataset	Number of Images	Number of Classes	Classes / Categories	Notes
[23]	2,527	6	Cardboard, Glass, Metal, Paper, Plastic, Trash	Images resized to 512×384 pixels; controlled lighting
[24]	14,964	4 major classes, 44 subclasses	Domestic waste across 44 subclasses under 4 main groups	Pascal VOC format; varying resolutions
[25]	15,150	12	Paper, Cardboard, Biowaste, Metal, Plastic, Green Glass, Brown Glass, White Glass, Clothes, Shoes, Batteries, Trash	Aimed at more granular classification

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

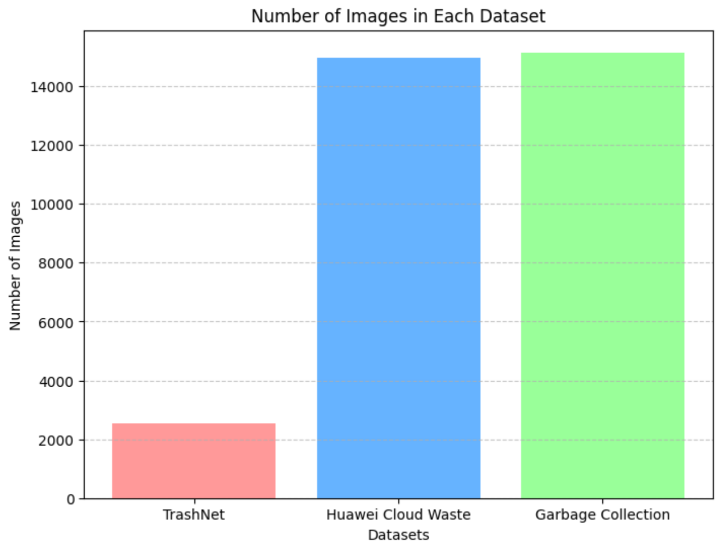

Figure 2 .The datasets samples .

Pre-processing

Data preparation is essential to the data analysis workflow. Therefore, many researchers have developed new methods and tools some using artificial intelligence to clean and prepare data. Research has also moved toward introducing specialized techniques for analysis. The first technique applied to each image is normalization, which helps speed up the selection of appropriate features and controls bias and variance when training AI models. Contrast and brightness adjustments also reduce noise and improve quality.

- Normalization : Because images often have intense colors, reducing these colors using a normalization technique, which aims to achieve a uniform range of values within [0, 1], is a key step. This, in turn, helps speed up feature convergence during model training [26-27]. Equation (1) illustrates the normalization process.

$$X_{nrml} = \frac{X - X_{min}}{X_{max} - X_{min}} \qquad (1)$$

where the two values (X_{min}, X_{max})represent the minimum and maximum pixel intensity values of the input image (X).

- The random spatial transformation operations are a set of techniques under data augmentation aimed at creating new training samples. These include [29] rotation of images with a random angle up to plus or minus 40 degrees (rotation range=40), horizontal and vertical shifts by 20 percent of image dimensions (width shift range=0.2 and height shift range =0.2), cropping by 20 percent (crop range =0.2) and random zoom in with the same ratio (zoom range =0.2). It has also been discussed in the paper about random horizontal flip (flip horizontal=true) which will help increasing spatial diversity in the images. For proper handling of resulting regions, fill mode was set to "nearest"(fill

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com
مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية

mode="nearest") so that empty pixels are filled with values of their nearest neighbors. This will enrich training data, improve generalization ability of model and reduce riskover fitting.

- The augmentation pipeline was implemented using the Keras ImageDataGenerator API available in TensorFlow 2.12. Augmentation is applied only during training and does not modify the original dataset offline explicitly. Also, augmentation is applied only on training set while validation and test sets are kept intact to have an unbiased, fair estimate of model performance.

In the next step, a shuffling process is applied to the features, as it plays a crucial role in preventing the model from learning from unhelpful patterns that may arise due to the order of the data, thereby making it more stable and less prone to errors during the training phase. Additionally, this technique contributes to achieving balance between classes, which enhances the model's ability to generate accurate predictions that closely reflect real-world data characteristics.

Dataset Splitting

The dataset was divided into three subsets to ensure effective training and accurate testing:

- 70% for training: to be used in developing the model and improving its accuracy.
- 10% Validation: To test the model during training and adjust the parameters.
- 20% Testing: To test the model's performance after training to ensure its accuracy and generalizability.

Classification model training

Transfer learning facilitates and improves the learning process using information learned from large datasets. In our approach, transfer learning is performed in two main parts: feature removal and fine-tuning. The DenseNet121 architecture, first trained on ImageNet, is used as the feature extraction model. This architecture contains densely connected convolutional layers, ensuring information flow between layers, which aids in feature propagation and reuse. Each dense block in DenseNet121 contains batch normalization (BN) layers and convolutional layers, followed by a ReLU activation function. It also contains a transition layer to control dimensionality. A global average pooling layer is also included to further reduce the dimensionality of features[30-31].

During fine-tuning, the final layers of DenseNet121 are frozen to use the pre-trained weights as an initial baseline. To improve the model's focus on important features, an attention module (CBAM) is added that applies channel and spatial attention sequentially[32]. The output is then fed into a global mean pooling layer, followed by random deletion to reduce overfitting. The output is then transformed via a dense layer into a composite representation, which is resampled and fed into a simple convolutional neural network to model any sequential relationships before final classification. The final dense layer performs multi-label classification among 12 histopathological classes using a softmax function[33], which also outputs a probability distribution over these classes, indicating which class the image is most likely to belong to. To achieve the intent of fair comparison between models, the training mechanism was made the same, which means using the same data split(70% for training, 10% for validation, 20% for testing) along with a batch size of

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

32 and a number of epochs of 20. Other techniques, such as EarlyStopping[34] and TensorBoard[35], are used in this process to monitor how well the model is performing during training and testing, which can help determine when it has learned enough and is not improving too much on the training data alone. Figure 3. To explain this model mathematically, it is expressed as follows:

$$F = \text{DenseNet121}(X) \qquad (2)$$

where X denotes the input image, and F is the resulting feature map.

To enhance the model's focus on salient regions and suppress irrelevant information, a Convolutional Block Attention Module (CBAM) was incorporated, applying channel and spatial attention mechanisms as follows.

$$Mc(F) = \sigma(\, MLP(AvgPool(F)) + MLP(MaxPool(F))\,) \qquad (3)$$

$$F' = Mc(F) \otimes F \qquad (4)$$

Next, the spatial attention is calculated by:

$$M_s(F') = \sigma(\, Conv_{7\times 7}(\, AvgPool(F'); MaxPool(F')\,) \qquad (5)$$

$$F'' = M_s(F') \otimes F' \qquad (6)$$

Subsequently, a Global Average Pooling layer was applied to spatially condense the features:

$$Z = \frac{1}{H\, x\, W} \sum_{i=1}^{H} 0 \sum_{j=1}^{W} 0 \;\, F''_{ij} \qquad (7)$$

followed by a Dropout layer to mitigate overfitting. The features were then transformed via a dense layer:

$$h = \emptyset\,(W_1\, z + b_1\,) \qquad (8)$$

and reshaped to be processed by a simple RNN layer that captures sequential dependencies:

$$s = RNN(h) \qquad (9)$$

Finally, a dense layer with a SoftMax activation was employed to generate the class probability distribution:

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

$$\hat{y} = softmax(W_2\, s + b_2) \qquad (10)$$

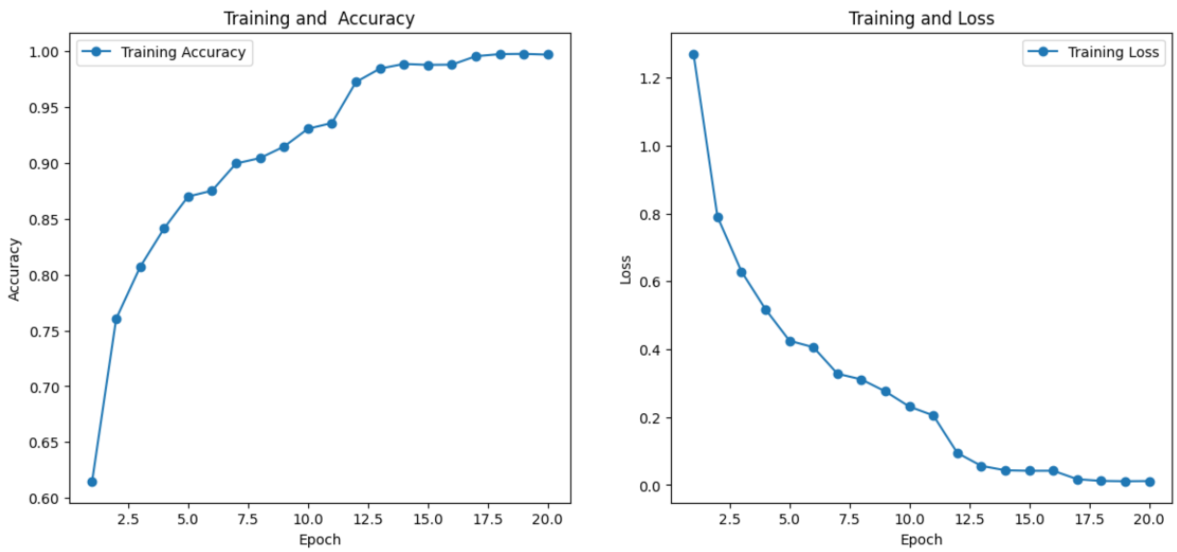

Figure 3. Training phase (accuracy curve) in the current study.

To make the proposed model completely reproducible and to distinctly report all experimental settings, the main hyperparameters applied in the training process are listed in Table 3. They comprise optimization configuration, batch size and number of epochs, learning rate, data split ratios besides other salient settings involved in DenseNet121–CBAM–SimpleRNN hybrid architecture.

Table 3. Summary of Training Hyperparameters.

Hyperparameter	Value
Backbone Network	DenseNet121 (ImageNet pre-trained)
Attention Module	CBAM (Channel + Spatial Attention)
RNN Units	64 units (SimpleRNN)
Image Size	224 × 224
Batch Size	32
Epochs	20
Optimizer	Adam
Learning Rate	0.001

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

Loss Function	Categorical Cross-Entropy
Dropout Rate	0.5
Weight Initialization	ImageNet pretrained weights
Data Split	70% training, 10% validation, 20% testing
Early Stopping Patience	5 epochs
Metrics	Accuracy, Confusion Matrix, Classification Report

RESULTS AND DISCUSSION

All experiments were conducted in TensorFlow 2.12 with its Keras high-level API. The model was trained and tested on Python 3.10 with GPU support from NVIDIA CUDA 11.8 and cuDNN 8.6, running on the same machine to ensure fast training performance as well as making the entire experimental pipeline fully reproducible. The proposed methodology in this research relies on a hybrid model that combines transfer learning techniques with an attention mechanism incorporated into convolutional neural networks (CNNs). This integration aims to enhance the model's ability to learn and focus on the most salient features within waste images, while simultaneously suppressing less relevant or noisy information. By directing computational emphasis toward important regions in the data, this approach improves the model's effectiveness in critical tasks such as classification, segmentation, and object recognition. Consequently, it contributes to more precise and efficient automated waste sorting, which is essential for sustainable environmental management.

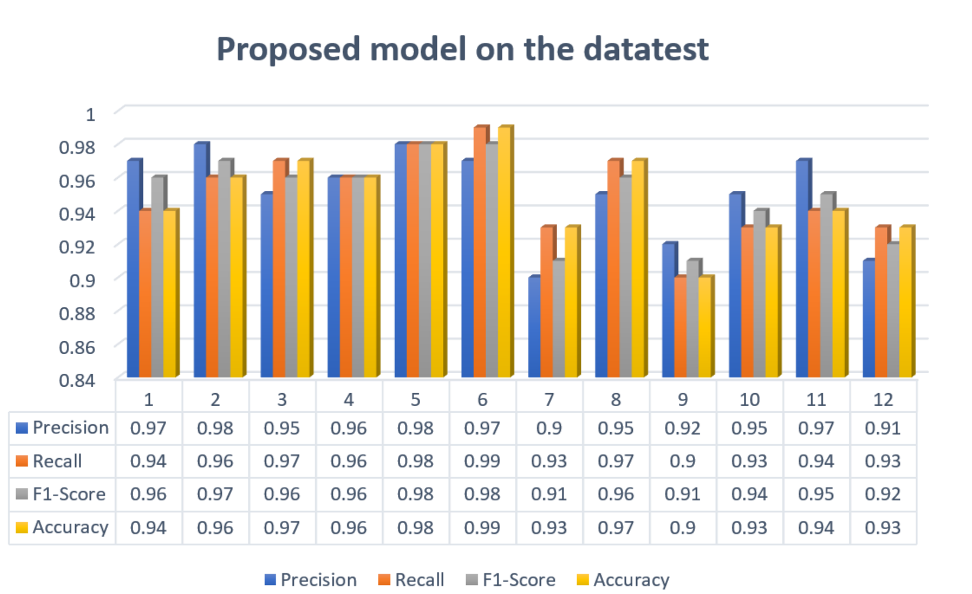

Figure 4. The evaluation on proposed system.

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

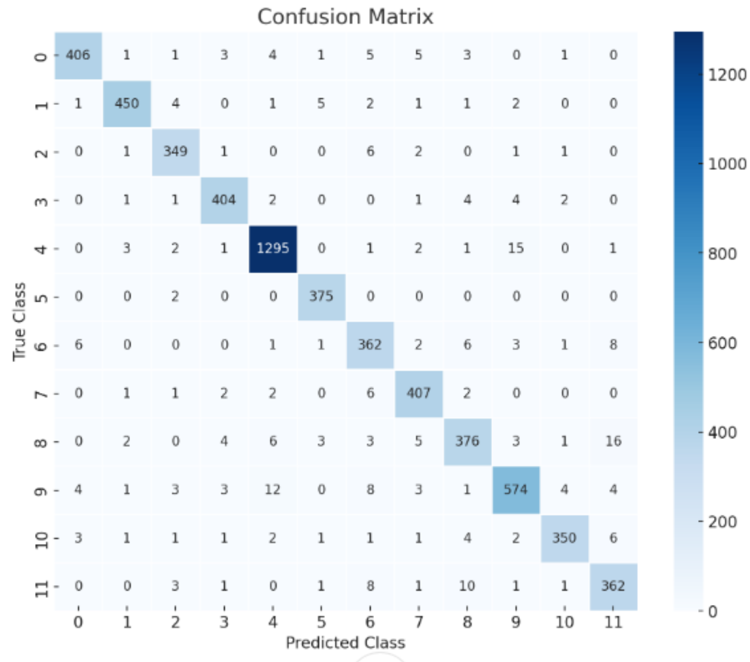

Figure 5. The confusion matrix in the current study.

The experimental results as shown in figure 4 clearly highlight the effectiveness of the proposed deep learning framework in classifying waste types across twelve distinct categories: battery, Biowaste, brown-glass, cardboard, clothes, green-glass, metal, paper, plastic, shoes, **Residuals**, and white-glass.Analyzing the per-class metrics, we observe that the model achieves outstanding recall (per-class accuracy) and F1-scores for clothes (98%) and green-glass (99%). The confusion matrix reveals minimal misclassifications for these categories, indicating that the features extracted by the DenseNet121 backbone, further refined by CBAM attention, capture their discriminative patterns effectively. On the other hand, slightly lower recall values are reported for metal (93%) and plastic (90%). The confusion matrix in figure 5 illustrates occasional misclassifications of metal samples into related categories such as battery and Residuals, which may arise due to overlapping visual textures or material similarities. Likewise, some plastic items are confused with cardboard or paper, potentially due to their shared shapes or surface appearances. Despite these slight variances, the macro and weighted averages across all classes consistently register around %95, underscoring the balanced nature of the model's performance even in the presence of class imbalance. The overall accuracy of %95 further emphasizes the general reliability of the proposed system in handling multi-class waste classification tasks. The strong results can be attributed to multiple architectural enhancements. The DenseNet121 network ensures deep feature propagation across layers, capturing complex spatial hierarchies. The integration of CBAM sequentially applies channel and spatial attention, enabling the model to focus on the most salient attributes within each image. Additionally, employing a Simple RNN layer after feature extraction facilitates learning latent sequential dependencies, contributing to robust decision boundaries. Also , The total training time was XX minutes on GPU.

Although the hybrid model includes multiple components, the increase in training time remained acceptable due to efficient transfer learning and GPU acceleration.

CONCLUSION

The results obtained clearly show that hybrid deep learning DenseNet121 combines CBAM attention mechanism with Simple RNN layer and model improves the waste multi-class classification by enhancing feature extraction and representation. Preprocessing steps and data augmentation techniques helped to overcome class imbalance in the dataset, making the model generalize well over the entire dataset. Experimental results showed that the model attained steady performance within all categories on a 12-class waste dataset and outperformed baseline, CNN-only architectures. Pre-trained weights also cut down training time relative to training the model from scratch. Therefore, this method could be used in automated waste-sorting systems because experimental results proved its efficiency.Future work aims at testing the validity of the model on larger and more diversified datasets as well as investigating other alternate attention mechanisms.

Acknowledgments:

Thanks to everyone who helped complete this achievement.

Conflict of interests:

There are non-conflicts of interest.

References

[1] S. Dong, W. Xu, H. Zhang, and L. Gong, “Cot-DCN-YOLO: Self-attention-enhancing YOLOv8s for detecting garbage bins in urban street view images,” Egyptian J. Remote Sensing Space Sci., vol. 28, no. 1, pp. 89–98, 2025.

[2] Z. Yang and D. Li, “WasNet: A neural network-based garbage collection management system,” IEEE Access, vol. 8, pp. 103984–103993, 2020.

[3] U. Özkaya and L. Seyfi, “Fine-Tuning Models Comparisons on Garbage Classification for Recyclability,” in *Proc. ISAS 2018-Winter, SETSCI Conf. Indexing System*, vol. 3, pp. 514–517, Samsun, Turkey, Nov. 30–Dec. 2, 2018.

[4] M. Sukel, S. Rudinac, and M. Worring, “GIGO, Garbage In, Garbage Out: An Urban Garbage Classification Dataset,” in MultiMedia Modeling (MMM 2023), Lecture Notes in Computer Science, vol. 13833, D. T. Dang-Nguyen et al., Eds. Cham, Switzerland: Springer, 2023, pp. 515–526. Doi: 10.1007/978-3-031-27077-2_41.

[5] R. Rismiyati, A. R. Situmeang, K. Khadijah, and S. N. Endah, “Garbage Image Classification Using Deep Learning: A Performance Comparison of InceptionResNetV2 vs ResNet50,” *J. Tek. Inform. (JUTIF)*, vol. 6, no. 4, pp. 2387–2379, Aug. 2025, doi: 10.52436/j.jutif.2025.6.4.4770.

[6] H. Zhou *et al.*, “Optimizing Backbone Networks Through Hybrid–Modal Fusion: A New Strategy for Waste Classification,” *Sensors*, vol. 25, no. 10, p. 3241, 2025, doi: 10.3390/s25103241.

[7] A. Setiawan, R. A. Yunmar, and H. Tantriawan, “Comparison of Speeded-Up Robust Feature (SURF) and Oriented FAST and Rotated BRIEF (ORB) methods in identifying museum objects using low light intensity images,” IOP Conf. Ser.: Earth Environ. Sci., vol. 537, p. 012025, 2020. doi: 10.1088/1755-1315/537/1/012025.

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

[8] Z. Piao, J. Wang, L. Tang, B. Zhao, and S. Zhou, "Anchor-free object detection with scale-aware networks for autonomous driving," *Electronics*, vol. 11, no. 20, p. 3303, 2022. doi: 10.3390/electronics11203303.

[9] N. Haque, R. Toufiq, M. Z. Islam, and M. A. A. T. Shoukhin, "Garbage classification using a transfer learning with parameter tuning," in *Proc. 6th Int. Conf. Electr. Eng. Inf. Commun. Technol. (ICEEICT)*, Dhaka, Bangladesh, 2024, pp. 1252–1256. doi: 10.1109/ICEEICT62016.2024.10534450.

[10] Z. Chen *et al.*, "Garbage classification system based on improved ShuffleNet v2," *Resour., Conserv. Recycling*, vol. 178, p. 106090, 2022.

[11] Z. Wu *et al.*, "Using YOLOv5 for garbage classification," in *Proc. 4th Int. Conf. Pattern Recognit. Artif. Intell. (PRAI)*, 2021.

[12] J. Yang *et al.*, "GarbageNet: A unified learning framework for robust garbage classification," *IEEE Trans. Artif. Intell.*, vol. 2, no. 4, pp. 372–380, 2021.

[13] X. Tian *et al.*, "Garbage classification algorithm based on improved MobileNetV3," *IEEE Access*, 2024.

[14] S. Gao *et al.*, "Role of garbage classification in air pollution improvement of a municipal solid waste disposal base," *J. Cleaner Prod.*, vol. 423, p. 138737, 2023.

[15] Y. Zhao *et al.*, "Intelligent garbage classification system based on improve MobileNetV3-Large," *Connection Sci.*, vol. 34, no. 1, pp. 1299–1321, 2022.

[16] Z. Liu *et al.*, "Garbage-classification policy changes characteristics of municipal-solid-waste fly ash in China," *Sci. Total Environ.*, vol. 857, p. 159299, 2023.

[17] B. Fu *et al.*, "A novel intelligent garbage classification system based on deep learning and an embedded linux system," *IEEE Access*, vol. 9, pp. 131134–131146, 2021.

[18] N. Haque *et al.*, "Garbage classification using a transfer learning with parameter tuning," in *Proc. 6th Int. Conf. Electr. Eng. Inf. Commun. Technol. (ICEEICT)*, 2024.

[19] Z. Wang, W. Zhou, and Y. Li, "GFN: A garbage classification fusion network incorporating multiple attention mechanisms," *Electronics*, vol. 14, no. 1, p. 75, 2024.

[20] R. Musham *et al.*, "Smart garbage classification using cutting edge technology (VGG-16)," in *Proc. Int. Conf. Emerg. Tech. Comput. Intell. (ICETCI)*, 2024.

[21] M. M. Hossen *et al.*, "GCDN-Net: Garbage classifier deep neural network for recyclable urban waste management," *Waste Manag.*, vol. 174, pp. 439–450, 2024.

[22] G. Liang and J. Guan, "FConvNet: Leveraging fused convolution for household garbage classification," *J. Circuits, Syst. Comput.*, vol. 33, no. 08, p. 2450140, 2024.

[23] W. Qiu, C. Xie, and J. Huang, "An improved EfficientNetV2 for garbage classification," *arXiv preprint*, arXiv:2503.21208, 2025.

[24] B. Madhavi *et al.*, "SwinConvNeXt: A fused deep learning architecture for real-time garbage image classification," *Sci. Rep.*, vol. 15, no. 1, p. 7995, 2025.

[25] S. Meng and W.-T. Chu, "A study of garbage classification with convolutional neural networks," in *Proc. Indo–Taiwan 2nd Int. Conf. Comput., Anal. Netw. (ICAN)*, 2020.

[26] M. Dohmen, M. A. Klemens, I. M. Baltruschat, T. Truong, and M. Lenga, "Similarity and quality metrics for MR image-to-image translation," *Sci. Rep.*, vol. 15, no. 1, p. 3853, 2025.

[27] A. Khaled, "BCN: Batch channel normalization for image classification," in *Proc. Int. Conf. Pattern Recognit.*, Cham, Switzerland: Springer, 2025, pp. 295–308.

[28] A. Tatar, M. Haghighi, and A. Zeinijahromi, "Experiments on image data augmentation techniques for geological rock type classification with convolutional neural networks," *J. Rock Mech. Geotech. Eng.*, vol. 17, no. 1, pp. 106–125, 2025.

[29] M. Alsaidi, M. T. Jan, A. Altaher, H. Zhuang, and X. Zhu, "Tackling the class imbalanced dermoscopic image classification using data augmentation and GAN," *Multimedia Tools Appl.*, vol. 83, no. 16, pp. 49121–49147, 2024.

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية

ISSN: 2312-8135 | Print ISSN: 1992-0652

info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

[30] J. Potsangbam and S. S. Devi, “Classification of breast cancer histopathological images using transfer learning with DenseNet121,” *Proc. Comput. Sci.*, vol. 235, pp. 1990–1997, 2024.
[31] T. S. Arulananth *et al.*, “Classification of paediatric pneumonia using modified DenseNet-121 deep-learning model,” *IEEE Access*, 2024.
[32] X. Tong, Z. Liang, and F. Liu, “Succulent plant image classification based on lightweight GoogLeNet with CBAM attention mechanism,” *Appl. Sci.*, vol. 15, no. 7, p. 3730, 2025.
[33] Y. Yanyan, W. Yajie, W. Chenglei, and S. Yinghao, “A novel garbage images classification method based on improved VGG,” in *Proc. 34th Chinese Control and Decision Conf. (CCDC)*, 2022, pp. 1571–1575.
[34] G. Husain, D. Nasef, R. Jose, J. Mayer, M. Bekbolatova, T. Devine, and M. Toma, “SMOTE vs. SMOTEENN: A study on the performance of resampling algorithms for addressing class imbalance in regression models,” *Algorithms*, vol. 18, no. 1, p. 37, 2025, doi: 10.3390/a18010037.
[35] R. Zaimi, M. Hafidi, and M. Lamia, “A deep learning mechanism to detect phishing URLs using the permutation importance method and SMOTE-Tomek link,” *The Journal of Supercomputing*, vol. 80, no. 12, pp. 17159–17191, 2024, doi: 10.1007/s11227-024-06124-7.

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية

الخلاصة

المقدمة:

يشهد العالم اهتمامًا متزاودًا بإدارة النفايات بشكل مستدام، الأمر الذي يتطلب أنظمة ذكية قادرة على فرز أنواع النفايات المختلفة تلقائيًا.

طرق العمل:

في هذه الدراسة، تم تصميم نموذج التعلم العميق المختلط باستخدام تقنية نقل المعرفة مع نموذج DenseNet121، بالإضافة إلى وحدة الانتباه في الشبكات العصبية الطبقية (CBAM) وشبكة عصبية متكررة بسيطة (Simple RNN) لتحسين تمثيل خصائص صور النفايات ومعالجة العلاقات التتابعية بينها. وللتغلب على مشكلة عدم توازن البيانات وتحسين جودة مجموعة البيانات، تم استخدام تقنيات المعالجة الأولية مثل المعايرة، وتعديل السطوع والتباين، وتكبير مجموعة البيانات (التدوير، العكس).

النتائج

ظهرت نتائج التجارب على مجموعة بيانات متنوعة من أنواع النفايات (12 فئة: بطاريات، نفايات عضوية، أنواع مختلفة من الزجاج، كرتون، ملابس، معادن، ورق، بلاستيك، أحذية، نفايات عادية، وزجاج شفاف) أن النموذج يتميز بأداء ممتاز، حيث حقق تصنيفًا متوازنًا لجميع الفئات. كما قلل هذا النموذج بشكل ملحوظ من حجم التدريب وعدد المعاملات باستخدام الأوزان المدربة مسبقًا، بينما ركزت آلية CBAM المقترحة على التعرف على الخصائص الأساسية

الاستنتاجات:

تؤكد نتائج التجارب على إمكانية تطبيق هذا النهج في أنظمة إعادة التدوير الذكية في العالم الحقيقي، مما يُعدّ وسيلة فعالة لفرز النفايات وحماية البيئة. سيشمل البحث المستقبلي دراسة إمكانية تطبيق هذه التقنية في مجالات بيئية أخرى، بالإضافة إلى اختبار أدائها على مجموعة بيانات أكبر وأكثر تنوعًا.

الكلمات المفتاحية:

التعلم الانتقالي، التعلم العميق، معالجة الصور، تعزيز السمات، تعزيز البيانات.