Automatic Garbage Classification

Abstract

The increasing volume of waste causes significant environmental challenges, and health problems and adversely affects the economy. Waste management through recycling is the way to reduce the growing waste generation. Thus, efficient and accurate classification of waste is important. Separating waste manually is time-consuming, necessitating the need to have an automated classification system for garbage to improve the recycling process, reduce waste generation, and maintain sustainability. Various models were developed for garbage classification purposes however those models need more computational resources and require hours of training time. Thus, the problem addressed in this research is developing an efficient and accurate automated system for garbage classification using an ensemble of CNN-ELM models that requires less time to train on the dataset.

The study aims to develop an ensemble model that combines CNNs with ELMs for the base models. Five CNN architectures were used as base models including ResNet50, InceptionV3, DenseNet121, VGG19, and EfficientNetB0. The main objective is to define a model that uses less training time but achieves an accuracy similar to the previous state-of-the-art. CNN-ELMs which are the base models were ensembled using the voting ensemble method. The predictions from the 5 models were aggregated and the prediction with the highest vote was set as the final answer. The models were trained and tested on a garbage classification dataset named ‘Trashnet’ [20]. The models were trained with and without data augmentation techniques to evaluate the performance. The model classifies the waste into 6 categories, paper, plastic, metal, trash, and glass. Accuracy, F1 Score, and confusion matrix were used to evaluate the performance of models.

The ensemble model attained an accuracy of 94.79% and an F1 Score of 94.78% when trained on non-augmented data. The performance of models decreased when trained on the augmented data. The model succeeds in achieving a good performance in a single epoch. The research successfully demonstrated that the ensemble approach enhances classification performance, meeting the objective of improving waste classification accuracy with less training time. However, the study also identified some limitations related to data augmentation.

Attestation

I understand the nature of plagiarism and am aware of the University's policy on this.

I certify that this dissertation reports original work by me during my University project except for the following:

- The code used in section 3.3 for pre-processing was largely taken from [21].
- The feature extraction part as discussed in section 3.5.1 was created using Keras Applications [33].
- The ELM model as discussed in section 3.5.2 was created using the library 'hpelm' [31].

Date: 30/08/2024

Acknowledgements

I would like to express my gratitude to those who have guided and supported me throughout this dissertation. Their guidance and help made it possible to complete this project.

I am grateful to my supervisors, Kevin Swingler, Keiller Nogueira, and Nora Tanner for their continuous guidance and feedback for the development of this work. Their years of experience and shared knowledge were important in shaping the direction and success of this project.

On a personal note, I would like to thank my family and friends for their emotional support and encouragement.

Table of Contents

List of Figures/ Tables

1 Introduction

1.1 Background and Context

Waste management is one of the growing problems that the world is facing. The rapid increase in population, urbanization, and economic development, leads to an increased waste generation at an alarming rate. The world produces over two billion metric tons of solid waste annually, a figure expected to increase by approximately 70% by 2050 [1]. Such rapid growth poses significant challenges to the ecosystem and human health. Waste can be toxic or hazardous, and if improperly treated can contaminate soil, air, and water leading to pollution. Thus, it is important to have effective waste management and disposal practices to minimize these negative impacts and to promote a sustainable future.

Landfills, Incineration, and Recycling are the most popular disposal techniques. Incineration and Landfills are the most commonly used disposal methods worldwide, but both pose severe negative impacts on human health and the environment. These methods have several significant problems such as producing greenhouse gases leading to climate change, land and soil degradation, etc. These methods are economically not sustainable [2].

Recycling is the most effective and environmentally friendly method [3]. Sorting of waste materials into different categories also contributes to waste reduction. Sorting the waste can help to understand how to reduce, reuse, or recycle the garbage, the 3 R's of recycling. Accurate classification is important as it determines the efficiency of recycling processes and the overall effectiveness of waste management systems. Manual separation is the traditional method of waste classification. However, this method is ineffective due to its time-consuming nature and high potential for human error [4]. This labor-intensive method has a high risk of long-term health problems due to exposure to hazardous substances. Another way of classification involved the use of magnets, one of the earliest technological innovations. The methods removed metallic items like iron or steel from the waste. Though effective, the technology could be applied to certain materials only. The introduction of automated sorting systems based on machine learning and computer vision can overcome these issues with manual sorting. Machine Learning has made remarkable progress across many fields in recent years. Leveraging machine learning and computer vision technology can be extremely helpful for such automation tasks due to their scalability and performance [5].

Automated systems with machine learning algorithms can be used to recognize and classify different types of waste to achieve high accuracy. Such systems can facilitate faster and more accurate segregation of waste, leading to more efficient recycling processes and reducing the environmental burden of waste. CNN or Convolutional Neural Networks, a deep learning neural network that excels at recognizing patterns and features in images has proven effective in image-based classification, including garbage classification. Bircanoglu et al. [6] in their experiment built RecycleNet (their own CNN model) which achieved an accuracy of 81% with augmented data. Jin S [32] proposed an improved MobileNetV2 version. To improve recognition accuracy, the attention mechanism is introduced in the first and last layer. N. Ramsurrun [2] explores a CNN-based system for classification on the Trashnet dataset. The method uses 12 different versions of CNNs with three types of classifiers: Support Vector Machine, Sigmoid, and Softmax. Current CNN garbage classification models using transfer learning and data augmentation have proved to be an effective way of yielding high accuracy. However, the CNN models often require high computational resources and high training time. Bircanoglu et al. [6] reported in their paper that their model required the longest training time, completing 300 epochs in 3 hours. Yusiong [7] in his paper used an ensemble approach (with CNN-ELMs as base models) that requires less training time but results in the same performance as others. This paper approaches the same method, exploring different CNN architectures and using augmentation techniques. This research focuses on the development of an ensemble model for garbage classification, with CNN (Convolutional Neural Networks)-ELMs (Extreme Learning Machines) as the base models. By integrating multiple CNN architectures with ELMs, the study aims to explore the strength of ensemble learning to overcome the limitations of individual models, ultimately leading to more accurate and reliable garbage classification systems.

1.2 Objectives

The primary objective of this study is to develop an ensemble model for garbage classification with CNN-ELMs as base models that improve garbage classification over existing models but with less training time. Combining the strength of CNN (Convolutional Neural Networks) and Extreme Learning Machines (ELM) to achieve better accuracy. The model classifies the waste materials into 6 categories: paper, glass, cardboard, metal, and trash.

The study aims to explore how different CNN architectures improve the accuracy and efficiency of garbage classification systems. Applying data augmentation to make the model more generalized and to assess the change in performance with augmented techniques. This research aims to critically analyze the performance, accuracy, and efficiency of the ensemble-based waste classification in comparison to the existing methods listed in the literature review section.

1.3 Achievements

The research advances the application of ensemble learning techniques and hybrid models in garbage classification. The CNN-ELM models were able to achieve good performance comparable to existing state-of-the-art with less training time. This research demonstrated the efficiency of using a CNN-ELM model by leveraging the strengths of both models. The ensemble model achieved superior accuracy and robustness compared to individual CNN-ELM models.

1.4 Overview of Dissertation

The dissertation explores the development of automatic garbage classification using an ensemble model with CNN-ELMs as base models. The dissertation is organized into 5 chapters:

- Chapter 1: INTRODUCTION lays the foundation of the project by introducing the need for waste management, different waste management techniques, and their limitations. It highlights the need to use Machine Learning and Computer vision for garbage classification. This chapter clearly states the research problem, objective, and achievements.
- Chapter 2: LITERATURE REVIEW investigates the existing research done related to CNNs, ELMs, and ensemble models for classification tasks, particularly in garbage classification.
- Chapter 3: This chapter includes the METHODOLOGY/RESULT part, explaining the methods used for the development of the model. This chapter includes details about the dataset and pre-processing. It details the used CNN architectures and the implementation of the model. The result and model performance are also analyzed and evaluated.
- Chapter 4: The CONCLUSION part summarizes all the findings, reflecting on how the research findings are met. It also highlights the limitations and potential of future works related to this research.

2 Literature Review

The rapid increase in waste generation has intensified the need for a proper and accurate classification system. An accurate and efficient garbage classification system is important for public health and effective environmental management. Proper classification can enhance resource recycling and reuse [8]. The current manual waste classification process is hazardous to human health and time-consuming. This leads to a need for an automated solution. This shift has been significantly driven by technological advancements, particularly in sensor-based systems. These systems utilize various sensors, and systems to detect physical and chemical properties of waste enabling more accurate classification compared to manual sorting. Recently, Machine Learning and Computer Vision have emerged as a powerful tool for enhancing the accuracy and efficiency of garbage classification. Machine learning systems can automatically learn from data patterns and improve their performance over time, leading to more accurate and efficient waste sorting. Below are some of the existing works done to improve the garbage classification accuracy:

- **CNN vs. traditional Machine Learning Algorithm**

  Bernado S. Costa [9] investigated various machine learning algorithms, including VGG16, AlexNet, SVM, K-Nearest Neighbors (KNN), and Random Forest, to determine the most effective model for waste classification. The study found that the pre-trained VGG16 model outperformed all other models, achieving an accuracy of 93%. This finding proved that deeper, more complex CNN architectures like VGGs are better suited for image classification tasks, especially when fine-tuned.

- **CNN – SVM Hybrid Models**

  Olugboja Adedeji [10] advanced the field by integrating CNNs with SVMs, leveraging the strengths of both models. They implemented a 50-layer ResNet Convolutional Neural Network (CNN), known for its deep learning capabilities and robust feature extraction, in combination with SVM for the final classification step. This hybrid approach has demonstrated a promising result and achieved an accuracy of 87% on the Trashnet dataset. This result demonstrates the effectiveness of combining the powerful feature extraction abilities of deep CNNs like ResNet-50 with the precise classification power of SVMs. The success of this model indicates that hybrid approaches can overcome some of the limitations faced by standalone CNN or SVM models.

- **Comparison of Different CNN architectures**

  The use of transfer learning has been widely adopted in the field of waste classification, as demonstrated by Sujan Poudel and Prakash Poudyal [4]. They used various CNN-based algorithms, including DenseNet201, ResNet50, InceptionResNetV2, MobileNet, VGG19, Xception, and InceptionV3, fine-tuning them on the Trashnet dataset. The transfer learning approach yielded high validation accuracies of 92%, 93%, and 93% for DenseNet201, ResNet50, and InceptionResNetV2, respectively. These results proved the effectiveness of using transfer learning in improving model performance by leveraging the knowledge gained from large, diverse datasets to enhance classification accuracy on smaller, task-specific datasets.

- **Own CNN Versions**

  Various deep-learning models were developed and evaluated for their effectiveness in garbage classification. Bircanoglu et al. [6] developed RecycleNet, a custom Convolutional Neural Network (CNN) model for waste classification. By using data augmentation, their model reached an accuracy of 81%. This shows that customized CNN models, along with enhanced training data, can significantly improve classification results.

  Jin S [32] made further progress by improving the MobileNetV2 architecture. To better recognize real-world garbage, he added an attention mechanism at the beginning and end of the network. This change significantly enhanced the model's accuracy to 89.26% in real-world tests. Adding attention mechanisms seems to help neural networks focus better, improving their ability of classification.

- **Ensemble Models for garbage Classification**

  Hua Zheng and Yu Gu [11] proposed an ensemble model called EnCNN-UPMWS, which combines multiple CNN architectures (GoogLeNet, ResNet-50, and MobileNetV2) with a unique strategy known as unequal precision measurement weighting (UPMWS). This ensemble approach employs the strengths of each CNN model to achieve more accurate and reliable waste classification. Ensemble models, by integrating multiple classifiers, can mitigate the weaknesses of individual models and provide more robust predictions. The development of ensemble models has shown promise in further improving waste classification accuracy.

- **CNN-ELM Framework**

  Huang et al. 2006 introduced a machine learning algorithm named ELM (Extreme Learning Machine) which has gained widespread recognition because of its good generalization performance, extremely fast learning speed, and ease of implementation. The algorithm provides results in a more efficient and fast manner compared to other neural networks [17][18]. Incorporating ELM (for classification) with CNN (for feature extraction) can make a strong hybrid model for classification with less training time [12], [13], [14], [15],[16].
  The CNN-ELM framework has proven its versatility and effectiveness in solving a variety of real-world classification problems across different domains. The CNN-ELM approach has been successfully applied to tasks such as DNA damage classification [12], age and gender classification [13], cervical cancer classification [14], electrocardiogram (ECG) signal classification [15], and accident image classification [16]. In these studies, the CNN-ELM models consistently outperformed traditional CNN-only models. Given its success in diverse applications, the CNN-ELM model could offer significant advantages in garbage classification. This is proved by John Paul Yosiong [7] in his paper.

- **Ensemble of CNN-ELM**

  Building on the idea of ensemble models and the proven success of CNN-ELM models, John Paul Yosiong [7] introduced LitterNet, an ensemble of different CNN-ELM models. This approach employs various pre-trained CNNs for feature extraction and an Extreme

Learning Machine (ELM) for the classification. LitterNet achieved results comparable to existing models but with the added advantage of reduced training time. The use of CNN-ELMs in ensemble models represents a novel approach, combining the feature extraction power of CNNs with the fast-learning capabilities of ELMs, making it a promising direction for future research in waste classification.

Recent studies have shown that CNNs are the most promising method for automatically sorting garbage. CNNs can automatically learn hierarchical features from the raw image data, unlike traditional methods that rely on handcrafted features. Research has shown that CNN-ELM models are faster and more efficient because they train in a single epoch. Combining multiple CNN-ELM models can make the predictions even more reliable and accurate. Using the same ensemble approach of CNN-ELM used by J.P. Yosiong [7], this paper explores different CNN architectures and utilizes data augmentation techniques to assess the performance of the models.

3 Methodology/Results

3.1 Research Design

The idea used in this research is to develop an ensemble model that incorporates CNN-ELMs as a base model using different CNN architectures for automatic waste classification. The ensemble model leverages the strength of both the CNN and ELM. The CNN model is used for its deep feature extraction capability and the ELM model because of its rapid learning ability.

The integration of CNNs with ELMs offers a dual advantage. CNNs are known for their complex feature extraction due to their deep architecture, ELMs on the other hand offer fast training and good generalization with their single-layer feedforward networks. The voting ensemble approach was used to combine the predictions from multiple models. Using an ensemble approach helps reduce each model's individual biases and enhances the overall classification [19]. The process was conducted initially on non-augmented data and then with the augmented data to assess the impact of augmentation on model performance.

3.2 Data Collection

The dataset used in this study was obtained from Kaggle – Garbage Classification [20]. The dataset was chosen because it contains various images for each label, ensuring a diverse representation of real-world scenarios, and is easily available without complex access restrictions. The data is stored in a directory structure in which images are sorted into subdirectories according to the class labels in the '.jpg' format. Figure 1 likely illustrates a subset of these images.

Figure 1: Data images

The dataset is loaded using TensorFlow into batches of 32 images. TensorFlow uses batching to stream the data from disk, making it easy to handle large datasets [21]. The dataset contains a total of 2527 files belonging to 6 different classes. The distribution of images for each label is as follows: 403 for cardboard, 501 for glass, 410 for metal, 594 for paper, 482 for plastic, and 137 for trash. The distribution is also represented in a bar plot in Figure 2.

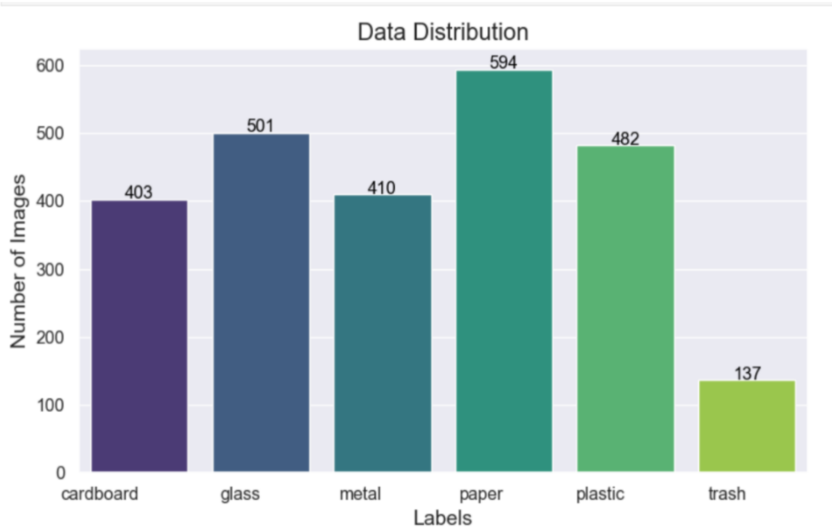

Figure 2: Bar graph displaying the distribution of data

3.3 Pre-Processing

Image pre-processing is a crucial step to ensure that the data is in optimal format for the model. Image preprocessing helps to improve the image quality and to analyze it more effectively.

- **Resizing:** To improve the response time during training, each image was programmatically resized to 256x256 pixels, the standard input size for many pre-trained CNN models.
- **Splitting:** The dataset was randomly split into training (80%), validation (10%), and testing (10%) subsets. Figure 3 represents the training, validation, and test dataset. The training dataset was used to train the models, the validation dataset was used for hyperparameter tuning, and the testing dataset was used to evaluate the performance of the final model.

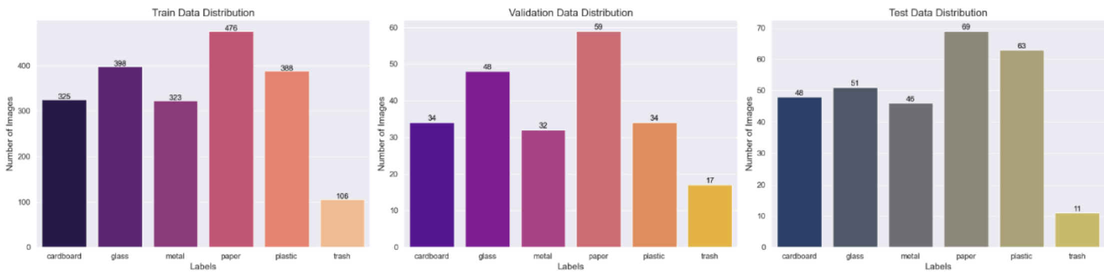

Figure 3: Data Distribution for train, validate and test dataset

- **Normalization:** The pixel intensity value of an image was adjusted to a range between 0 to 255, to standardize the pixel value and to enhance the contrast of an image.
- **Data Augmentation:** Finally, data augmentation was performed on the training dataset to artificially increase its diversity and create a more robust model for generalization. Positional augmentation, such as flipping and rotation, was used. A separate variable was defined for the augmented data to compare the result with and without augmentation.

3.4 CNN Architectures

Five CNN models with diverse architecture namely, ResNet50, Inception, EfficientNet, DenseNet, and VGG19 were used for the feature extraction stage facilitating the transfer learning approach. The transfer learning approach utilizes the knowledge (of identifying complex patterns and features) learned from the large and diverse dataset, in this case, 'imagenet'. Pre-trained base models were loaded using the 'tensorflow.keras.applications' library with the weights trained on the imagenet dataset and with the top layer excluded. The base model was frozen to prevent its weights from being updated during training, thereby preserving the learned features. The pre-processed images were passed through the base model to extract features and then an average pooling layer was applied to convert the feature maps into a single feature vector. Average pooling is commonly used with pre-trained models as the final layer when the classification layer is avoided [22].

CNN Architectures	**Strengths**
ResNet50	• Efficient at extracting abstract and complex features because of its deep layers and residual connections [23][10].
InceptionV3	• Good at extracting features at different scales (locally and globally) due to the multi-branch inception modules [24]. • Provides effective feature extraction for diverse datasets.
VGG19	• Simple and effective for extracting hierarchical features. • The model is capable of achieving optimal values [25].
DenseNet	• More diversified and rich pattern feature extraction due to its dense connections. [26]. • Can achieve similar or better performance with fewer parameters [26].
EfficientNet	• Optimized for balancing model accuracy and efficiency using compound scaling [27]. • Provides excellent feature extraction with fewer parameters [28].

Table 1: Used CNN architectures

3.5 Implementation of CNN-ELM

The CNN-ELM is a hybrid model that combines the strength of two powerful machine learning techniques. The model integrates the Convolutional Neural Networks (CNNs) described above with Extreme Learning Machine (ELM) to leverage their strengths. Figure 4 illustrates the design of the CNN-ELM model.

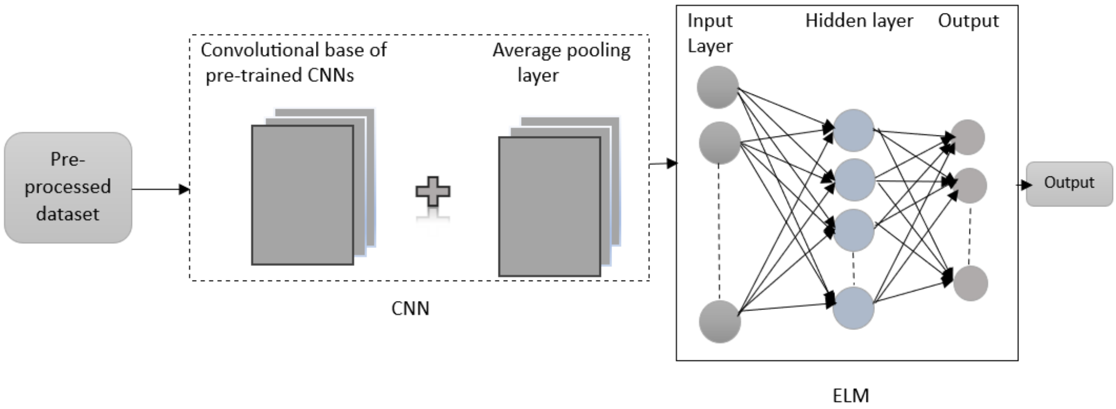

Figure 4 : CNN-ELM Framework

The implementation of the model is as follows:

3.5.1 Feature Extraction with CNNs

CNN architecture has multiple layers, such as the input, convolutional layer, pooling layer (downsample layer), and the fully connected layer or classification layer. The convolutional layer is responsible for the feature extraction, and the pooling layer is used to reduce the spatial dimensions of those extracted features [29][30]. For the first step, these two layers, including the input layer, are required. The last layer, the fully connected classification layer, is removed. Only the convolutional base (convolutional layer + pooling layer), which acts as a feature extractor, is retained. The feature map obtained from the last convolutional layer was used as input to the ELM.

3.5.2 Classification using ELM and its Configuration

ELM model consists of three layers, the input, a single hidden layer, and the output layer. As shown in Figure 5 The number of features corresponds to the number of inputs to the ELM, while the number of labels determines the number of neurons in the output layer.

ELMs were configured with a single hidden layer. The hidden layer is assigned a specific number of neurons and an activation function, which can vary between different CNN models. The number of neurons in this hidden layer was tuned based on validation performance and the activation function was set to ‘tanh’. The number of neurons in the hidden layer is a critical hyperparameter that can significantly impact model performance, potentially leading to overfitting if not carefully chosen.

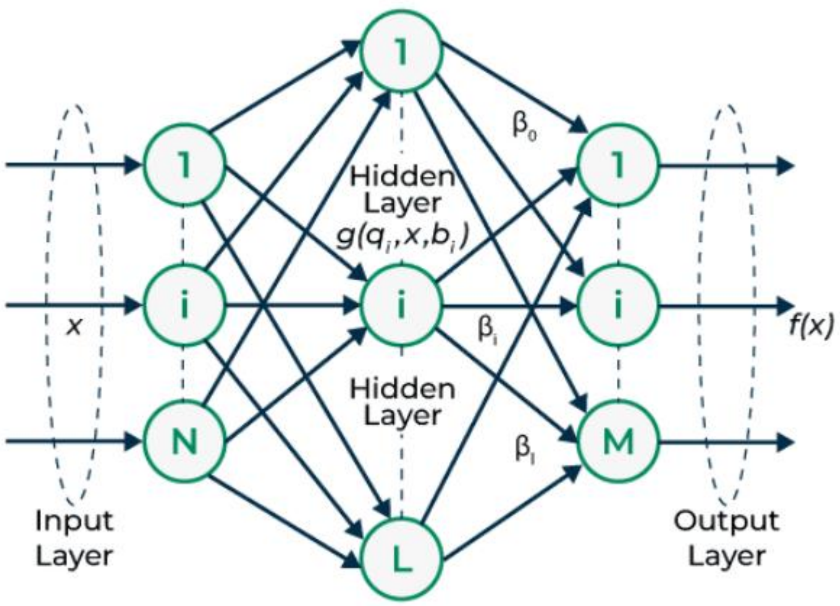

Figure 5: ELM Architecture [17]

Consider,

- **X:** vector of input features
- **N:** total number of features
- **J:** number of data points
- **M:** number of output variables
- **L:** number of neurons in the hidden layer
- **W:** input weight matrix of size L
- **b:** hidden layer bias vector (L,1)
- **H:** output of hidden layer (J, L)
- **g(.):** activation function
- **β:** Output weight matrix (L, M)
- **f(x):** output prediction matrix (J, M)

The output of the hidden layer, **H** is calculated by:

$$\boldsymbol{H} = \boldsymbol{g}(\boldsymbol{W} * \boldsymbol{X} + \boldsymbol{b})$$

, and the output of the network is calculated by:

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{H} * \boldsymbol{\beta}$$

The ELM training involves a unique process, the input weights and bias are randomly assigned which are fixed and do not change [17]. And for the output weights, it doesn't use gradient-based backpropagation, but the Moore-Penrose generalized inverse method [17]. This avoids the need for iterative optimization, resulting in rapid model training.

3.5.3 Training Process

The CNN-ELM models were trained in two stages. First, the pre-trained CNNs were used on training data for extracting high-level features from the image data. The extracted features were scaled using Standard Scaler to enhance the performance of the ELM model. Next, the feature maps generated by CNNs were passed to the ELM model for classification purpose. The

ELMs were trained using the Moore-Penrose pseudoinverse to compute the output weights, allowing fast learning.

3.5.4 Validation Process

To improve the performance of the ELM model, the number of neurons which is a hyperparameter is hyper-tuned on the validation dataset. ELM doesn't support standard hyperparameter techniques like Grid Search or Randomized Search CV because of its unique architecture and training process. The Model Structure Selection (MSS) method from the 'hpelm' library to determine the best number of optimal neurons and to avoid overfitting [31]. The optimal number of neurons for each model is listed in Table 2.

CNN-ELM Model	Number of neurons (when trained on non-augmented data)	Number of neurons (when trained on augmented data)
ResNet50-ELM	750	750
InceptionV3-ELM	859	427
VGG19-ELM	937	362
DenseNet121-ELM	812	684
EfficientNetB0-ELM	886	375

Table 2: Number of optimal neurons associated with every model

3.5.5 Testing Process

The CNN-ELM models were tested on the test dataset to evaluate their performance. The results were compared among various CNN-ELM models.

3.6 Ensemble Model

The ensemble model was developed to classify different waste materials. Ensemble learning enhances the result by combining various independent models through voting, stacking, bagging, or boosting [19].
Five trained CNN-ELM models with different CNN architectures were used as base models. The results from the base models were combined using the voting ensemble approach. The prediction with the maximum vote will be considered the final output for that particular input. The model workflow is illustrated in Figure 6.

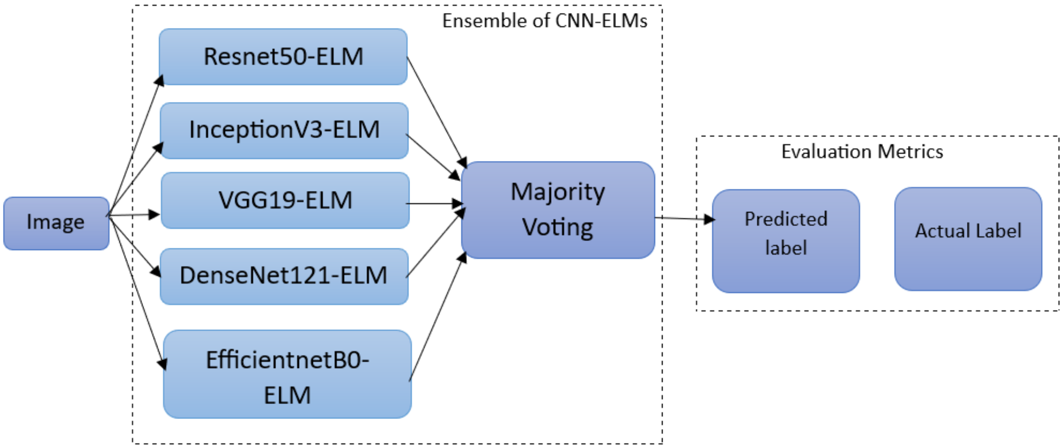

Figure 6: Ensemble model framework

3.6.1 Voting Strategy

- Each CNN-ELM model votes for a class for a particular image and the class with the majority of votes is chosen as the final prediction.
- Voting strategy reduces individual model biases and provides a more robust model.

3.6.2 Training and Testing

- After training the individual CNN-ELMs, their predictions were combined using the voting strategy.
- The ensemble model was evaluated on the test set to determine its effectiveness compared to individual models.

3.7 Evaluation Metrics

The models' performance was evaluated using various metrics such as accuracy, f1 score, and confusion matrix.

- Accuracy: The accuracy defines the proportion of correctly classified instances (correctly classified waste images of test dataset) to the total number of instances (total waste images of test dataset). Using accuracy as an evaluation metric is a common strategy many researchers apply.

$$Accuracy = \frac{Correct\ number\ of\ prediction}{Total\ number\ of\ prediction}$$

- F1 Score: The F1 score provides a harmonic balance between precision and recall. In the case of garbage classification, it is important to consider having fewer misclassified instances because you cannot apply a specific recycling type to another type of waste.

$$F1\ Score = 2 * \frac{Recall * Precision}{Recall + Precision}$$

- Confusion Matrix: A confusion matrix was used to provide a detailed breakdown of the model's performance across all classes, showing the true positives, true negatives, false positives, and false negatives for each class.

3.8 Result

3.8.1 Performance of Individual CNN-ELMs on test data

When examining the results without data augmentation, the model with Densenet and InceptionV3 CNN architecture stood out achieving an accuracy and f1 score of around 93%. This is because of their dense architecture, making them suitable for more diversified and rich pattern extraction. VGG19-ELM and EfficientNetB0-ELM also performed better attaining an accuracy of 85.07% and 79.51% respectively. ResNet50-ELM, despite having deep architecture suitable for extracting complex features attained the lowest accuracy of 57.64%.

In contrast, when data augmentation was applied, the overall performance of all models declined significantly. The DenseNet121-ELM model attained an accuracy of 79.51%, which is the highest. Though the accuracy decreased but the model is still the best performed model. InceptionV3-ELM achieved the second highest accuracy among all the models which is 72.92%, followed by VGG19-ELM with an accuracy of 65.62%. Despite excelling with non-augmented data, DenseNet121-ELM and InceptionV3-ELM suffered a sharp drop in accuracy, a decrease of over 14% and 20% respectively. ResNet50-ELM and EfficientNetB0-ELM were the most affected by augmentation, with their accuracies dropping to 32.29% and 45.14%, respectively, indicating a significant struggle to adapt to the augmented data.

CNN-ELM Model	**Accuracy**	**F1 Score**
ResNet50-ELM	57.64%	57.43%
InceptionV3-ELM	93.75%	93.73%
VGG19-ELM	85.42%	85.30%
DenseNet121-ELM	93.75%	93.76%
EfficientNetB0-ELM	79.51%	79.56%

Table 3: Performance of CNN-ELMs without data augmentation

CNN-ELM Model	**Accuracy**	**F1 Score**
ResNet50-ELM	32.29%	31.68%
InceptionV3-ELM	72.92%	71.95%
VGG19-ELM	65.62%	64.34%
DenseNet121-ELM	79.51%	79.50%
EfficientNetB0-ELM	45.14%	44.87%

Table 4: Performance of CNN-ELMs with data augmentation

3.8.2 Performance of Ensemble Model on Test Data

- Without Data Augmentation, the ensemble model achieved an accuracy of 94.79% and an F1 score of 94.78%. This performance is significantly higher compared to any individual CNN-ELM model, demonstrating the effectiveness of the ensemble approach in combining predictions from multiple models to enhance overall accuracy.
- With Data Augmentation, the ensemble model's accuracy dropped to 76.04%, with a corresponding F1 score of 75.73%. Although this still represents a strong performance, it is noticeably lower than the results obtained without augmentation. This decline indicates that individual models struggled to generalize from augmented data, thus, negatively impacting the ensemble's overall performance.

Ensemble Model	**Accuracy**	**F1 Score**
Without augmentation	94.79%	94.78%
With augmentation	76.04%	75.73%

Table 5: Performance of Ensemble models with and without augmentation

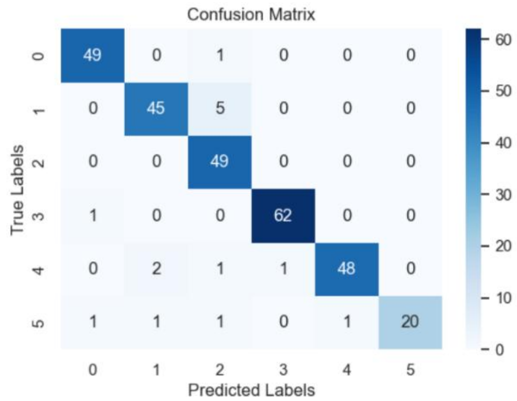

Figure 7: Confusion Matrix for ensemble model (without augmentation)

Figure 8: Confusion Matrix for ensemble model (with augmentation)

3.8.3 Comparison with Existing Studies

The proposed model achieved almost similar accuracy as the previous state-of-the-art. The accuracy is slightly lower than some of the high-performing models but still demonstrates a strong performance and potential for future applications.

Author	Model	Accuracy
Bernado S. Costa [9]	VGG16	93%
Olugboja Adedeji [10]	ResNet50-SVM	87%
Sujan Poudel and Prakash Poudyal [4]	DenseNet201 ResNet50 InceptionResNetV2	92% 93% 93%
Hua Zheng and Yu Gu [11]	EnCNN-UPMWS	93.50%
John Paul Yosiong [7]	LitterNets	93.97%
Proposed model	Ensemble of CNN-ELMs	94.79%

Table 6: Comparative Analysis

4 Conclusion

4.1 Summary

The research evaluated the effectiveness of using an ensemble model for garbage classification. Five CNN architectures integrated with ELM were used as base models. The CNN architectures used for the hybrid model include ResNet50, InceptionV3, VGG19, DenseNet121, and EfficientNetB0. The architectures were chosen because of their capabilities of extracting rich and sophisticated features from the image data. The algorithms were proved to be effective in various classification problems. The research also evaluated the performance of models based on augmentation. First, the non-augmented data were trained and tested using the independent CNN-ELM model and the ensemble model on a diverse dataset of garbage images. Then the augmented data were tested and trained using the models. The models were evaluated using Accuracy, F1 Score, and Confusion Matrix.

The results indicate that though most of the CNN-ELMs were able to achieve good accuracy using less training time indicating strong individual performance, the ensemble model outperformed all the base models in both the cases of using augmented and non-augmented data. The ensemble model achieved higher accuracy and an F1 Score demonstrating the strength of combining different CNN-ELM through voting ensemble approach. The augmentation techniques when applied to this dataset degrades the performance of the model.

The ensemble model achieved comparable accuracy to the previous state-of-the-art while requiring less training time.

4.2 Future Work

This study demonstrates the effectiveness of using an ensemble model and explores different CNN architectures. Though the model achieved good accuracy, some limitations are impacting the model's performance. Data Augmentation which was supposed to improve the performance, resulted in a significant decrease in accuracy. The models might need fine-tuning to improve the performance. The study did not cover all possible CNN architectures or ensemble methods. The focus was on the selected models and a specific technique. Another drawback is that the findings will be specific to the particular dataset and configurations used.

Future work includes the optimization of the ensemble model and trying different CNN architectures to make it more accurate and effective. The study employed a voting ensemble approach, although numerous alternative methods for combining predictions could be explored. The next step is to implement the model in real-world applications and assess its effectiveness in these settings.

References

[1] B. Alves, "Topic: Waste generation worldwide," *Statista*, Feb. 08, 2023. https://www.statista.com/topics/4983/waste-generation-worldwide/#topicOverview
[2] N. Ramsurrun, G. Suddul, S. Armoogum and R. Foogooa, "Recyclable Waste Classification Using Computer Vision And Deep Learning | IEEE Conference Publication | IEEE Xplore," *ieeexplore.ieee.org*. https://ieeexplore.ieee.org/document/9499291
[3] D. Team, "5 Types of Waste Disposal Methods | PROS Services," *PROS Services*, May 20, 2022. https://prosservices.com/5-types-of-waste-disposal-methods/

[4] S. Poudel and P. Poudyal, "Classification of Waste Materials using CNN Based on Transfer Learning," *Proceedings of the 14th Annual Meeting of the Forum for Information Retrieval Evaluation*, Dec. 2022, doi: https://doi.org/10.1145/3574318.3574345.

[5] D. G. Solla, "Advanced waste classification with Machine Learning," *Medium*, Jan. 14, 2022. https://towardsdatascience.com/advanced-waste-classification-with-machine-learning-6445bff1304f

[6] Bircanoğlu, Cenk & Atay, Melvin Selim & Beser, Fuat & Genc, Ozgun & Kizrak, Merve Ayyuce. (2018). RecycleNet: Intelligent Waste Sorting Using Deep Neural Networks. 10.1109/INISTA.2018.8466276. https://www.researchgate.net/publication/325626219_RecycleNet_Intelligent_Waste_Sorting_Using_Deep_Neural_Networks?enrichId=rgreq-7933a4cca61acb464b6a3d5c401dedab-XXX&enrichSource=Y292ZXJQYWdlOzMyNTYyNjIxOTtBUzo2Njc2MTQwMzI0NTM2MzZAMTUzNjE4Mjk5NzgxOQ%3D%3D&el=1_x_3#full-text

[7] Yusiong, John Paul. (2022). An Ensemble of CNN-ELM Models for Trash Classification. ICIC Express Letters. 16. 943-950. 10.24507/icicel.16.09.943. RecycleNet: Intelligent Waste Sorting Using Deep Neural Networks (researchgate.net)

[8] Son Vu, "Waste: Origin, Classification, Processing, and Recycling," *Shopequo.com*, 2023. https://shopequo.com/blogs/blog/waste-classification?srsltid=Afm-BOoosu44aGIpIMk6lbmQNsJ84-eJGY2VXmsSAw1LFOWQu6Q0WszhM

[9] Costa, Bernardo & Bernardes, Aiko & Pereira, Julia & Zampa, Vitoria & Pereira, Vitoria & Matos, Guilherme & Soares, Eduardo & Soares, Claiton & Silva, Alexandre. (2018). Artificial Intelligence in Automated Sorting in Trash Recycling. 198-205. 10.5753/eniac.2018.4416. https://www.researchgate.net/publication/330350735_Artificial_Intelligence_in_Automated_Sorting_in_Trash_Recycling

[10] O. Adedeji and Z. Wang, "Intelligent Waste Classification System Using Deep Learning Convolutional Neural Network," *Procedia Manufacturing*, vol. 35, pp. 607–612, 2019, doi: https://doi.org/10.1016/j.promfg.2019.05.086.

[11] H. Zheng and Y. Gu, "EnCNN-UPMWS: Waste Classification by a CNN Ensemble Using the UPM Weighting Strategy," *Electronics*, vol. 10, no. 4, p. 427, Feb. 2021, doi: https://doi.org/10.3390/electronics10040427.

[12] Y. T. Hafiyan, Afiahayati, R. D. Yanuaryska, E. Anarossi, V. M. Sutanto, J. Triyanto and Y. Sakak ibara, A hybrid convolutional neural network-extreme learning machine with augmented dataset for DNA damage classification using comet assay from Buccal Mucosa sample, International Journal of Innovative Computing, Information and Control, vol.17, no.4, pp.1191-1201, 2021. https://keio.elsevierpure.com/en/publications/a-hybrid-convolutional-neural-network-extreme-learning-machine-wi

[13] M. Duan, K. Li, C. Yang, and K. Li, "A hybrid deep learning CNN–ELM for age and gender classification," *Neurocomputing*, vol. 275, pp. 448–461, Jan. 2018, doi: https://doi.org/10.1016/j.neucom.2017.08.062.

[14] A. Ghoneim, G. Muhammad and M. S. Hossain, Cervical cancer classification using convolutional neural networks and extreme learning machines, Future Generation Computer Systems, vol.102, pp.643-649, 2020. https://www.sciencedirect.com/science/article/pii/S0167739X19306065

[15] S. Zhou and B. Tan, Electrocardiogram soft computing using hybrid deep learning CNN-ELM, Applied Soft Computing, vol.86, DOI: 10.1016/j.asoc.2019.105778, 2020. https://www.sciencedirect.com/science/article/pii/S1568494619305599

[16] A. Pashaei, M. Ghatee, and H. Sajedi, "Convolution neural network joint with mixture of extreme learning machines for feature extraction and classification of accident images," *Journal of Real-Time Image Processing*, vol. 17, no. 4, pp. 1051–1066, Feb. 2019, doi: https://doi.org/10.1007/s11554-019-00852-3.

[17] "Extreme Learning Machine," *GeeksforGeeks*, Oct. 30, 2023. https://www.geeksforgeeks.org/extreme-learning-machine/

[18] K. Erdem (burnpiro), "Introduction to Extreme Learning Machines," *Medium*, May 29, 2020. https://towardsdatascience.com/introduction-to-extreme-learning-machines-c020020ff82b

[19] Awan-Ur-Rahman, "Understanding Soft Voting and Hard Voting: A Comparative Analysis of Ensemble Learning Methods," *Medium*, Aug. 02, 2023. https://medium.com/@awanurrahman.cse/understanding-soft-voting-and-hard-voting-a-comparative-analysis-of-ensemble-learning-methods-db0663d2c008

[20] "Garbage Classification," *www.kaggle.com*. https://www.kaggle.com/datasets/asdasdasasdas/garbage-classification

[21] codebasics, "Deep learning project end to end | Potato Disease Classification Using CNN - 1 : Problem Statement," *YouTube*, Aug. 22, 2021. https://www.youtube.com/watch?v=dGtDTjYs3xc&list=PLeo1K3hjS3ut-JFNGyBpIvjWgSDY0eOE8S

[22] A. Rosebrock, "Convolutional Neural Networks (CNNs) and Layer Types," *PyImageSearch*, May 14, 2021. https://pyimagesearch.com/2021/05/14/convolutional-neural-networks-cnns-and-layer-types/

[23] "Weights & Biases," *W&B*, 2024. https://wandb.ai/mostafaibrahim17/ml-articles/reports/The-Basics-of-ResNet50---Vmlldzo2NDkwNDE2#:~:text=The%20complete%20architecture%20of%20ResNet50

[24] K. Yan, "These are the 5 best pre-trained neural networks," *Medium*, Jul. 11, 2023. https://medium.com/@kyan7472/these-are-the-5-best-pre-trained-neural-networks-23798e61a043

[25] Stateczny, Andrzej & Kiran, Goru & Bindu, Garikapati & Kanegonda, Ravi Chythanya & Swamy, Kondru. (2022). Spiral Search Grasshopper Features Selection with VGG19-ResNet50 for Remote Sensing Object Detection. Remote Sensing. 14. 5398. 10.3390/rs14215398. https://www.researchgate.net/publication/364975950_Spiral_Search_Grasshopper_Features_Selection_with_VGG19-ResNet50_for_Remote_Sensing_Object_Detection

[26] RAVI SHEKHAR TIWARI, "Transfer Learning -Part -7.0 !! Dense net - RAVI SHEKHAR TIWARI - Medium," *Medium*, Jul. 22, 2022. https://tiwari11-rst.medium.com/transfer-learning-part-7-0-dense-net-bd23d9f6d4fd

[27] "EfficientNet," *Huggingface.co*, 2020. https://huggingface.co/docs/timm/en/models/efficientnet

[28] S. Herath, “EfficientNet,” *Data Science and Machine Learning*, Jan. 03, 2024. https://medium.com/image-processing-with-python/efficientnet-a835427acf8b
[29] GeeksforGeeks, “Introduction to Convolution Neural Network,” *GeeksforGeeks*, Mar. 14, 2024. https://www.geeksforgeeks.org/introduction-convolution-neural-network/

[30] “Pooling Layers | Dremio,” *www.dremio.com*. https://www.dremio.com/wiki/pooling-layers/#:~:text=What%20is%20Pooling%20Layers%3F

[31] “Model Structure Selection — hpelm 1.0.3 documentation,” *Readthedocs.io*, 2014. https://hpelm.readthedocs.io/en/latest/api/mss.html

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

[33] “Applications - Keras 2.1.3 Documentation,” *Faroit.com*, 2017. https://faroit.com/keras-docs/2.1.3/applications/