Organic and recyclable waste classification using deep learning methods

MSc Research Project
MSCAI1 JAN23l – Research in Computing CA1

Rashi Dabhane
Student ID: x21176321

School of Computing
National College of Ireland

Supervisor: Mayank Jain

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Rashi Dabhane
Student ID:	x21173621
Programme:	MSCAI1 JAN23I – Research in Computing CA1
Year:	2023
Module:	MSc Research Project
Supervisor:	Mayank Jain
Submission Due Date:	14/12/2023
Project Title:	Organic and Recyclable waste classification using deep learning methods
Word Count:	6746
Page Count:	26 (Including references)

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Rashi Sunil Dabhane
Date:	14/12/2023

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	✓
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	✓
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	✓

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Organic and recyclable waste classification using deep learning methods

Rashi Dabhane
x21173621

Abstract

This study addresses the pressing global challenge of waste management and the crucial task of organic and recyclable waste classification. Specifically investigating transfer learning with the VGG-16 deep learning architecture, the approach integrates multiple diverse datasets into a comprehensive dataset. The paper offers a comparative analysis of models, including DenseNet121, DenseNet169, VGG16, VGG19, and ResNet18, highlighting innovations in architecture customization and strategic techniques like learning rate scheduling and early stopping, etc. The study not only holds practical significance in automating waste classification processes but also reduces reliance on manual sorting, thereby promoting sustainable waste management practices. Rigorous experimentation underscores achieving a peak accuracy of 95.59% with VGG-16 and transfer learning, contributing to substantial enhancements in model performance, reliability, and generalizability. This contribution aligns with the broader global agenda of sustainable development.

Keywords: Waste management, deep learning, vgg16, image classification, sustainable development.

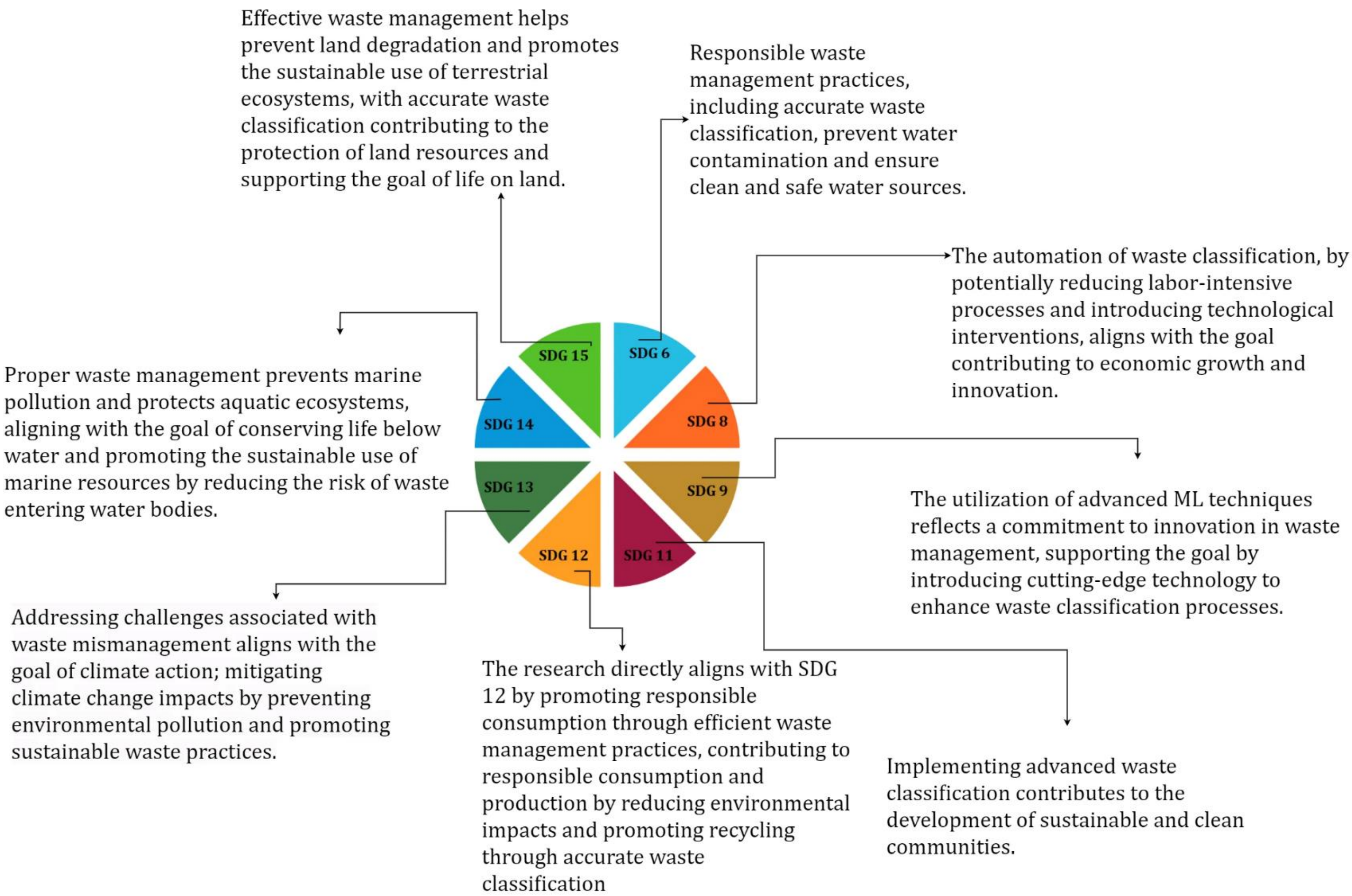
1 Introduction

Effective waste management is critical in the ever-increasing concern for environmental sustainability. The importance of this topic arises from the escalating environmental challenges posed by improper waste disposal. With conventional methods characterized by labour-intensive and error-prone processes proving insufficient, leveraging advanced technologies, particularly convolutional neural networks (CNNs), stands out as a promising solution to enhance the efficiency of waste classification.

This research was motivated by the pressing need to address inefficiencies in existing waste management systems. As environmental concerns grow, the need for clever, waste management systems become evident. The proposed waste classification model is a technological intervention that would improve waste sorting accuracy, reduce labour, and lead to more sustainable and environmentally friendly waste management ecosystem.

Beyond its practical applications in waste management, this research holds substantial merit from an academic and learning standpoint. Developing a deep learning-based waste classification model provides an opportunity for learners and researchers to delve into the complexities of advanced machine learning techniques. The study encompasses data preparation, model architecture development, training, optimisation, and evaluation. The application of deep learning methods in a real-world scenario not only improves technical skills but also fosters a deeper understanding and awareness of the convergence between technology and environmental sustainability.

Contribution to Sustainable Development Goals (SDGs):



This research aligns with several Sustainable Development Goals (SDGs), providing a comprehensive approach to address environmental challenges.

Fig. 1. Sustainable Development Goals (SDGs) relevant to the waste management and Image classification

This research focuses on waste classification, specifically differentiating between organic and recyclable waste using deep learning methods.

The concept of iterative experimentation resonates with the decision to revisit the challenge of Organic and recyclable waste classification using deep learning, wherein we continuously refine the research based on insights gained from previous experiments, showcasing commitment to enhancing the accuracy and efficiency of waste classification models. The choice to revisit this challenge stems from the belief that initial solutions, while effective, can benefit from refinement to address emerging complexities and boost overall performance.
The primary contribution of this research lies in the development and implementation of an enhanced waste classification model using advanced machine learning techniques. We employ a Convolutional Neural Network (CNN) architecture, specifically VGG16, known for its

success in image classification tasks. Transfer learning from pre-trained models and data augmentation techniques contribute to the technical robustness of the proposed WasteVGG16 model along with some architecture customizations, aiming to overcome the limitations of traditional and existing methods.

CNNs are well-suited for waste classification, excelling in image analysis with the ability to automatically learn spatial hierarchies and distinguish intricate features within diverse waste images. Their feature abstraction capability ensures accurate categorization of materials in a waste stream, particularly related to organic and recyclable materials.

The rest of the structure of the paper is organized as follows: The Literature Review that provides a comprehensive overview of existing knowledge and research relevant to the study. The Methodology section introduces the proposed research methodology, outlining the step-by-step approach taken in the study. The Design and Implementation specifications section explores the technical aspects of the study, detailing design considerations and implementation specifics. The Evaluation presents experiments and results analysis, offering insights into the outcomes and model performance. Finally, the Conclusion and Discussion summarizes findings, draws conclusions, and engages in a discussion about the implications and significance of the study.

2 Literature review

The rapidly increasing volume of urban household waste poses a critical environmental and resource threat, necessitating effective waste management strategies. Notably, garbage recycling and automatic sorting have emerged as practical solutions. While conventional image classification methods have succeeded in addressing waste image classification issues, they often overlook spatial relationships between features, resulting in misclassifications.

The authors [4] highlight the increasing urban waste production and stress the importance of waste separation. Deep learning, particularly convolutional neural networks (CNNs), has significantly enhanced image classification accuracy, enabling automated garbage sorting using vision technology.

Despite progress, existing CNN studies exhibit limitations, primarily related to spatial relationships between features. CNNs tend to lose information about spatial relationships during pooling operations, leading to false positives. Addressing these limitations of CNNs, the authors propose the ResMsCapsule waste image classification algorithm, incorporating a residual module and a multi-dimensional capsule module in the capsule network structure.

The ResMsCapsule model is evaluated on the TrashNet dataset, achieving a classification accuracy of 91.41%. Comparative analyses with other models, including AlexNet, Vgg16, and ResNet18, reveal superior performance and fewer parameters in ResMsCapsule. Notably, the proposed model outperforms SVM and Inception-ResNet, achieving the highest accuracy.

The authors address misclassifications, particularly between plastic and glass objects, highlighting the challenge posed by the similarity in material and shape, especially in light backgrounds. Future work is proposed to focus on additional data preprocessing steps to further condense features, reduce model parameters, and enhance performance.

The research [5] addresses a critical issue in waste management, emphasizing the potential of automation to enhance efficiency. The study focuses on the development and evaluation of a waste sorting system employing Convolutional Neural Networks (CNNs), specifically ResNet18, for image classification. By utilizing a dataset comprising images from TrashNet and the authors' collection, the researchers aim to automate waste segregation, reducing the need for human intervention.

The authors have employed transfer learning, utilizing pre-trained models like ResNet and VGG, to address the challenges associated with training deep neural networks on limited data. Results showcase the effectiveness of the ResNet18 model, exhibiting a validation accuracy of 87.8%. However, the study acknowledges the need for further research, emphasizing data augmentation and model fine-tuning as possibilities for improvement.

The paper [6] introduces a Waste classification method employing a Multilayer Hybrid Convolutional Neural Network (MLH-CNN) to enhance waste categorization accuracy and efficiency. Experimentation on the TrashNet dataset illustrates MLH-CNN's superiority over Vgg16, AlexNet, and ResNet50, achieving a notable 92.6% classification accuracy, surpassing existing methods by 4.18% to 4.6%. Despite these advancements, several limitations and research gaps are the dependence on the TrashNet dataset raises concerns about limited diversity in waste scenarios. With only 2527 images, the dataset's small size prompts questions about generalizability.

The simplified occlusion tests may not capture real-world complexities, and the absence of real-world deployment considerations, including computational constraints and hardware scalability, is a notable limitation. While reporting high accuracy, a detailed analysis of metrics such as precision, recall, and F1-score is crucial. Additionally, the lack of comparison with non-neural network approaches and unexplored real-time processing considerations further highlights possibilities for future research, aiming to fortify the proposed method's real-world applicability and performance robustness.

This study [7] explores the state of the art in waste classification systems, with a focus on intelligent systems using deep learning, particularly Convolutional Neural Networks (CNN) and Support Vector Machines (SVM). The paper proposes an automated waste material classification system to streamline waste separation processes, reduce human involvement, and enhance the efficiency of waste management. The proposed solution involves intelligent waste material classification, utilizing a combination of ResNet-50 pre-trained CNN as an extractor and SVM for waste classification. The paper employs a trash image dataset, consisting of four waste categories: glass, paper, metal, and plastic. The small dataset size is acknowledged, emphasizing the need for data augmentation techniques to maximize diversity.

Results from the experiments are presented, showcasing an 87% accuracy on the trash image dataset. The training process using Stochastic Gradient Descent with Momentum is outlined, along with key parameters and their impact on the system's performance. The authors suggest possibilities for future improvements, including expanding the dataset for enhanced accuracy and the potential for categorizing additional waste items.

Waste management is a critical global challenge due to the escalating volume of waste, especially in urban areas. This paper [8] proposes an automated waste classification system using deep convolutional neural networks (CNNs) to categorize waste into organic and solid classes (glass, metal, plastic). The OrgalidWaste dataset, comprising 5600 images from various waste datasets, facilitates model training. The study compares the performance of different CNN architectures, including 3-layer CNN, VGG16, VGG19, Inception-V3, and ResNet50.

This paper contributes by focusing on a four-class classification (organic, glass, metal, plastic) and using a novel dataset. The proposed approach aims to address the challenges of manual waste sorting, offering a cost-effective solution through automation. The proposed OrgalidWaste dataset, augmented with real-world waste images, provides a valuable resource for training robust waste classification models. The study demonstrates the efficacy of VGG16, achieving an accuracy of 88.42%, and highlights the potential for automating waste management processes. Future work could involve expanding the dataset and integrating the developed model into practical waste management applications.

The research [9] underlines the pressing challenges of landfills due to population growth and stresses the need for robust waste management solutions. Existing research, such as the suggested waste classification method, which employs CNN, VGG16, and ResNet50 models, demonstrates breakthroughs in accurate waste classification. The paper specifically emphasizes the superior accuracy of ResNet50, reaching 93.35%. Furthermore, the literature reveals a gap in the integration of waste classification with broader waste regulation strategies. While the proposed models offer notable accuracies, there's a critical evaluation of the dataset's impact on training, suggesting the importance of dataset expansion. These results highlight the continuous search for effective waste management systems and point to future research possibilities in dataset quality enhancement, model integration, and overall waste regulation strategies.

The paper [12] presents a real-world dataset named GIGO, aiming to facilitate algorithm development and benchmarking for the multimodal classification of urban waste in street-level imagery. The dataset comprises 25k images from Amsterdam, collected to aid cities in sustainable garbage collection. The paper introduces challenges such as visually heterogeneous garbage categories and varying environmental conditions. The research emphasizes the importance of multimodal approaches in urban waste management, leveraging real-time street-level imagery. The authors provide open data statistics about the geographic area, encouraging experimentation with multimodal techniques. State-of-the-art baselines using ResNet, EfficientNet, VisionTransformers, and SwinTransformers are presented, with the VisionTransformer architecture yielding the best results.

The unique challenges of fine-grained multimodal garbage classification, including dynamic backgrounds, varying object sizes, and spatial-temporal dynamics, are discussed. The authors propose a multimodal graph neural network (GNN) that leverages geospatial information to enhance classification performance. The GIGO dataset is released with annotations, contextual data, and baseline experiments, fostering research in sustainable urban waste management.

The paper concludes by highlighting the potential for further improvements, suggesting the incorporation of additional information about the surroundings using techniques like graph neural networks and transformers.

3 Methodology

The research aims to develop a robust waste classification system using deep learning methods for distinguishing between organic and recyclable waste. The study includes a comprehensive comparison of various convolutional neural network (CNN) architectures, namely VGG16[22], ResNet18[23], VGG19[22], DenseNet169[24], DenseNet121[24], and a simple CNN for waste classification.

1. Data Collection

1.1 Dataset

- *The dataset is a combination of various datasets, resulting in a larger and more comprehensive dataset that consists of 27,603 images of organic and recyclable waste items collected from Kaggle.*
- *The dataset is split into training, validation and testing sets, ensuring a balanced distribution of classes to avoid bias in model training. The training set is used to train the models, the validation set helps in tuning hyperparameters, and the test set evaluates the model's performance on unseen data.*

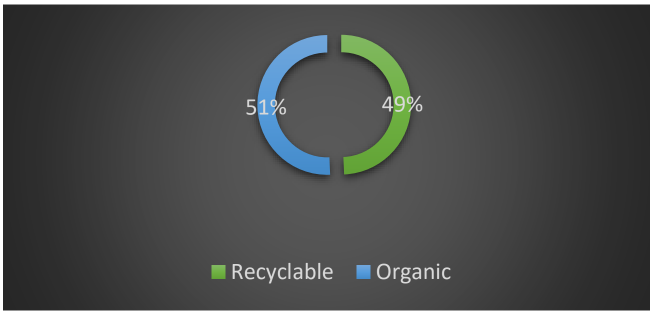


Fig. 2. Waste images distribution into recyclable and organic waste

1.2 Data Pre-processing Techniques

- *Dataset Loading:*
  Images of organic and recyclable waste items are sourced and organized into training and testing datasets.

- *Labelling:*
  Each waste image is associated with a corresponding label, indicating whether it belongs to the organic or recyclable category.

- *Data Transformation*
  Data Augmentation:

Data augmentation is applied to the training set using transformations such as random resizing, centre cropping, horizontal flipping, random rotation, and colour jittering. This enhances the model's ability to generalize by exposing it to diverse variations of the input data and the dataset is enriched with diverse representations of waste items. This variety is crucial for training a robust model capable of accurately classifying different types of organic and recyclable waste.

i. *Resizing the Image:* Images in the dataset are adjusted to have a consistent size of 256x256 pixels. This ensures uniformity and aids in model training.
ii. *Horizontal Flipping:* To diversify the dataset, some images are horizontally flipped randomly. This mimics the natural variability in the orientation of waste items.
iii. *Random Rotation:* Introducing a degree of randomness, images undergo random rotations of up to 15 degrees. This augmentation simulates different angles at which waste items might be captured.
iv. *Centre Cropping:* Images are centrally cropped to a size of 224x224 pixels. This focuses the model's attention on the central region, capturing the most relevant features.
v. *Conversion to RGB Format:* Ensuring a consistent colour representation, all images are converted to the RGB format. This helps maintain uniformity in colour channels.
vi. *Conversion to Tensor:* Images are transformed into a format compatible with deep learning models, specifically tensors. This facilitates efficient processing by the neural network.
vii. *Normalization:* Normalization adjusts the pixel values of the images to a standardized scale. This aids in stable and efficient training by bringing all features to a similar range.

	Waste Image Dataset	
Train set example images		
	Image 1. Organic waste 1	**Image 2. Organic waste 2**
		
	Image 3. Recyclable waste 1	**Image 4. Recyclable waste 2**

Test set example images		
	Image 5. Organic waste 1	**Image 6. Organic waste 2**
		
	Image 7. Recyclable waste 1	**Image 8. Recyclable waste 2**

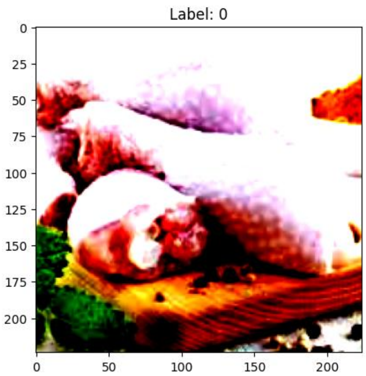


Image 9. Sample image 'Organic waste'

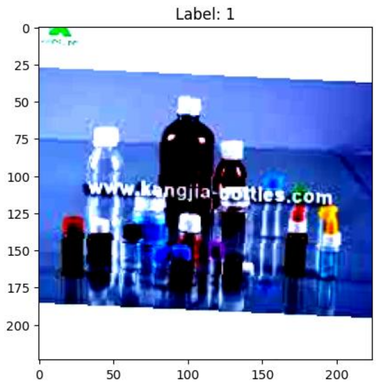


Image 10. Sample image 'Recyclable waste'

Visual inspection plays a crucial role throughout the development process. It ensures the effectiveness of data preprocessing and augmentation techniques by identifying and correcting anomalies during the data quality check. Examining augmented images aids model interpretability, providing insights into how the model generalizes to diverse data variations, showcasing its robustness. In instances of suboptimal model performance, visual analysis of sample images helps identify issues like misclassifications or unexpected patterns. Visual examples derived from sample images enhance understanding, making it easier to grasp the nature of problems and comprehend the model's performance across various scenarios. Visual inspection, integrated at different stages of model development, provides both qualitative and quantitative insights into data and model behavior.

2. Model Implementation

CNN Architectures

The study compares VGG16[22], ResNet18[23], VGG19[22], DenseNet169[24], DenseNet121[24], and a simple CNN for waste classification. The selection of each architecture is based on a trade-off between complexity, interpretability, and efficiency. DenseNet architectures are chosen for their parameter efficiency and feature reuse, ResNet for addressing vanishing gradient, and VGG16 as a straightforward baseline. The Simple CNN is included for simplicity and ease of understanding. Each architecture is implemented leveraging pre-trained models for feature extraction.

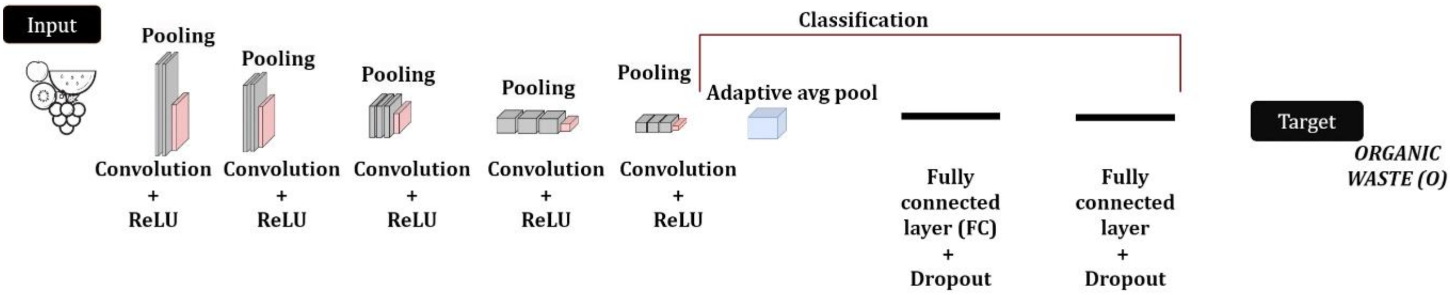


Fig. 3. A CNN architecture

A. VGG16

The adoption of the VGG16 architecture in this study is founded on its well-established reputation as a Convolutional Neural Network (CNN) design, widely acknowledged for its effectiveness in tasks related to image analysis.

The reason behind selecting VGG16 precedes the detailed exploration of its role in the study. This choice is rooted in the model's simplicity, making it an accessible and interpretable tool for waste classification. By incorporating VGG16 as a pivotal model, the study establishes a benchmark for evaluating the performance of more complex architectures. Furthermore, the application of transfer learning emphasizes on the knowledge embedded in VGG16 during its pretraining on ImageNet, thereby enhancing its proficiency as a feature extractor tailored to the specific challenges posed by waste classification.

B. Working

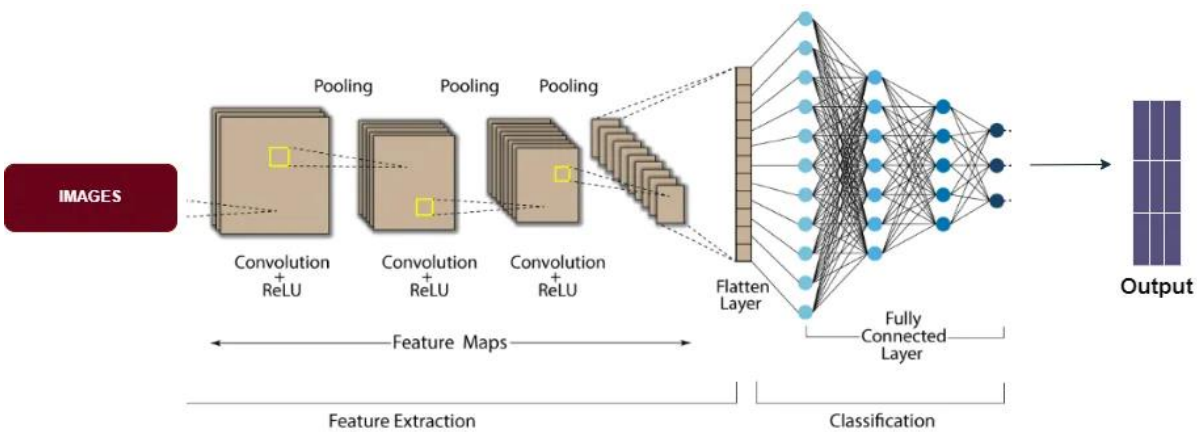


Fig. 4. VGG16 architecture (Customized)

Forward Pass

The pre-processing steps (refer section 1.2) ensure that the input adheres to the format expected by the model. The pre-processed image is then passed through the VGG16 backbone which consists of multiple convolutional layers that automatically learn hierarchical features from the image.

Feature extraction

Convolutional layers extract low-level features like edges and textures; Neurons in these layers respond to patterns within small receptive fields. As the filters build over the input image, they create feature maps that represent localized information. As the image data progresses through the network, it encounters deeper convolutional layers. These layers build on the low-level features detected in earlier layers and combine them to recognize more complex patterns and structures. Features extracted in deeper layers become more abstract and include larger spatial arrangements.

Pooling layers, usually max pooling, come after convolutional layers. Pooling aids in the achievement of translation invariance, which means that the model becomes less sensitive to the precise position of features. This improves the network's capacity to recognise patterns no matter where they appear in the input.

VGG16's final fully connected layers integrate the abstracted features from the convolutional layers to make predictions. The network learns to weight characteristics that are more representative of the target classes. The convolutional layer output is flattened and sent to a fully connected layer. This layer oversees integrating the obtained features and predicting. The model's prediction is generated by the final fully linked layer. Because the objective is binary classification (Organic or Recyclable), the output has two nodes, and raw scores are often converted into probabilities using a softmax activation function.

The output of the model is a probability distribution over the classes. For example, the model might output [0.8, 0.2], indicating an 80% probability of being Organic and a 20% probability of being Recyclable. A threshold is applied to convert the probabilities into class labels. For instance, if the threshold is set at 0.5, predictions with a probability above 0.5 are assigned to the positive class (O - Organic), and those below are assigned to the negative class (R - Recyclable). The threshold here has been chosen based on experimental observations, to make a balanced decision for both the classes (O and R) as well as using a threshold of 0.5 implies that both false positives and false negatives are considered equally important in the experimentation.

To compute the loss, the model's predictions are compared to the ground truth labels (the image's actual class). The loss indicates how well or poorly the model performs on the given task. The determined loss is utilised to update the weights and biases of the model. Backpropagation is utilised to do this, in which the gradients of the loss with respect to each parameter in the model are computed and used to alter the parameters.

Backpropagation and optimisation techniques (here, Adam optimizer) are used to change the weights of the convolutional filters and fully connected layers throughout the training phase.

This refines the network's ability to recognise characteristics relevant to the classification. The entire process (forward pass, backward pass, and parameter updates) is performed over and over for numerous epochs (the model is rigorously trained on multiple epochs ranging from 5 to 50 epochs). The model will learn and increase its ability to accurately classify images over time.

Following training, the model's performance is evaluated on a separate test dataset. Classes are predicted using the same forward pass approach. The trained model may be used to infer new, previously unseen pictures. The pre-trained weights enable the model to foresee images it hadn't encountered before during training.

This entire process of forward and backward passes, training, evaluation, and inference is orchestrated by the underlying deep learning framework. The model, through the learning process, becomes adept at distinguishing between Organic and Recyclable waste based on the features it has learned from the training data.

Customization to the architecture:

- The customization of the architecture for the waste classification task involves alterations to the last fully connected layer. The number of output features in the last fully connected layer has been changed to 2, indicating the model is designed for binary classification (two classes), specifically distinguishing 'Organic' from 'Recyclable.'
- Dropout Added (0.5): Dropout is a regularization technique where, during training, randomly selected neurons are ignored, or "dropped out."
- Training the model on the waste classification task leverages the CrossEntropyLoss in conjunction with the softmax activation function.
- The purpose of adaptive average pooling is to adaptively adjust the spatial dimensions of the data to a predefined size.

During the training process, several techniques and strategies have been implemented to optimize the performance of the neural network. These include the use of the Adam optimizer, a learning rate scheduler, and early stopping. The explanation behind these choices and detailed configurations will be elaborated in the next section.

4 Design and Implementation Specifications

A. Model Selection

The Waste Classification task utilizes the VGG16 architecture as the foundational Convolutional Neural Network (CNN). VGG16 is chosen for its simplicity, making it interpretable and serving as a baseline for performance comparison with more complex architectures like ResNet and DenseNet. Transfer learning is leveraged, with the pretraining of VGG16 on ImageNet, enhancing its capability as a feature extractor.

The table provides an overview of the architecture of the proposed Waste Classification model based on a customized VGG16 architecture. It outlines each layer's characteristics:

Layer	Input Size	Output Size	Kernel Size	Stride	Padding	Activation	Type
Conv1	3x224x224	64x224x224	3x3	1	1	ReLU	Convolution
Conv2	64x224x224	64x224x224	3x3	1	1	ReLU	Convolution
MaxPool1	64x224x224	64x112x112	2x2	2	0	-	Max Pooling
Conv3	64x112x112	128x112x112	3x3	1	1	ReLU	Convolution
Conv4	128x112x112	128x112x112	3x3	1	1	ReLU	Convolution
MaxPool2	128x112x112	128x56x56	2x2	2	0	-	Max Pooling
Conv5	128x56x56	256x56x56	3x3	1	1	ReLU	Convolution
Conv6	256x56x56	256x56x56	3x3	1	1	ReLU	Convolution
Conv7	256x56x56	256x56x56	3x3	1	1	ReLU	Convolution
MaxPool3	256x56x56	256x28x28	2x2	2	0	-	Max Pooling
Conv8	256x28x28	512x28x28	3x3	1	1	ReLU	Convolution
Conv9	512x28x28	512x28x28	3x3	1	1	ReLU	Convolution
Conv10	512x28x28	512x28x28	3x3	1	1	ReLU	Convolution
MaxPool4	512x28x28	512x14x14	2x2	2	0	-	Max Pooling
Conv11	512x14x14	512x14x14	3x3	1	1	ReLU	Convolution
Conv12	512x14x14	512x14x14	3x3	1	1	ReLU	Convolution
Conv13	512x14x14	512x14x14	3x3	1	1	ReLU	Convolution
MaxPool5	512x14x14	512x7x7	2x2	2	0	-	Max Pooling
AdaptiveAvgPool	512x7x7	512x7x7	Adaptive	-	-	-	Avg Pooling
FC1	512x7x7	4096	-	-	-	ReLU	Fully Connected
Dropout1	4096	4096	-	-	-	-	Dropout
FC2	4096	4096	-	-	-	ReLU	Fully Connected
Dropout2	4096	4096	-	-	-	-	Dropout
FC3	4096	2	-	-	-	Softmax	Fully Connected

Table 1. Customized VGG16 Architecture for Waste Classification

B. Optimization Parameters

The Adam Optimizer is a key component in enhancing model training. It is selected for its efficiency in convergence, requiring minimal hyperparameter tuning, suitable for tasks with substantial data and high-dimensional parameter spaces, well-suited in waste classification. For experimentation, Adam Optimizer is employed along with a loss function and customized beta values (0.95, 0.9995) for stability.

To dynamically adjust the learning rate throughout the training process, a Step Learning Rate Scheduler is implemented. This scheduler, reducing the learning rate by a factor of 0.5 every two epochs, strikes a balance between rapid convergence and fine-tuning. This adaptability contributes to optimal convergence during the training of waste classification models [30]

Furthermore, Early Stopping is incorporated as an important regularization strategy. By halting training when model performance on validation data deteriorates, overfitting is prevented. it set to 5 epochs without improvement. This strategy ensures that the model generalizes effectively to unseen waste data.

Binary Cross-Entropy Loss is selected as the loss function, suitable for binary classification tasks. Binary Cross-Entropy loss is mathematically defined as follows:

$$L_{BCE} = -\frac{1}{n}\sum_{i=1}^{n}(Y_i \cdot \log \hat{Y}_i + (1 - Y_i) \cdot \log(1 - \hat{Y}_i))$$

where,
n is the number of instances.
Yi is the true label (0 or 1) for the i-th instance.
$\hat{Y}_i$ is predicted probability of belonging to class 1 for the i-th instance

Image 11. Binary Cross-Entropy Loss [32]

In binary classification, the BCE loss encourages the predicted probabilities to be close to 0 for instances belonging to class 0 and close to 1 for instances belonging to class 1. The negative log-likelihood term penalizes deviations from the true labels.

In terms of experimentation, the application of Dropout (0.5) during training serves as a regularization technique. Randomly ignoring neurons during training helps prevent overfitting, contributing to the model's ability to generalize well to diverse waste images.

The determination of optimal parameters, including the choice of beta values, learning rate scheduling, and early stopping criteria, is rooted in iterative experimentation. Experimenting across multiple cycles of training and validation, modifying batch sizes, adjustments were made based on observed convergence patterns and performance metrics. These tailored parameter choices ensure that the model is finely tuned to the nuances of waste classification, demonstrating improved accuracy and robustness. Each iteration yields valuable insights into the model's behaviour, guiding further adjustments to achieve improved performance.

C. Visualization:

Seaborn and Matplotlib are used for visualizing results. The training and validation loss curves and the accuracy curve over epochs provide a concise overview of the model's learning process, showcasing trends in both training and validation performance. The loss curve helps identify convergence patterns, while the accuracy curve offers insights into the model's ability to correctly classify waste images. Confusion matrixes are visualized using Seaborn's heatmap functionality. The confusion matrix illustrates the model's performance on the test set, presenting a clear depiction of predicted versus true labels for both organic and recyclable waste classes.

D. Software Environment:

- Programming Language: Python 3.8
- Libraries and Frameworks:
    a. PyTorch, NumPy, Pandas: Essential for deep learning, numerical operations, and data handling.
    b. Matplotlib, Seaborn: Utilized for result visualization.
    c. ReportLab: Employed for PDF report generation.

E. Deployment:

- Environment: Jupyter Notebook, Google colab.
- Hardware: The model is designed to run on a GPU-enabled environment for accelerated training.

F. Computational Efficiency

- GPU Utilization:
  The model training was conducted on Jupyter and Google Colab as well as utilizing an RTX 3050 GPU for acceleration. The GPU acceleration significantly expedited the training process compared to CPU-only implementations.

- Training Time:
  The training time is influenced by several factors, including the complexity of the model architecture, dataset size, and the computational power of the hardware, etc.

i. *VGG16 with Transfer Learning and Customizations:*
   Training Time: Approximately 175 to 225 minutes (20-50 epochs). This approach significantly reduces training time compared to default VGG16 settings, making it a best suitable choice for image classification task.
ii. *Default VGG16 Configuration:* Training Time - Approximately 250-372 minutes.
iii. *ResNet18:* Training Time - Approximately 189 to 243 minutes.
iv. *Simple CNN*: Training Time - Approximately 90-170 minutes. A simplified CNN architecture demonstrates faster training times, albeit with potential trade-offs in complexity.

The choice of utilizing VGG16 with transfer learning and customizations enhances efficiency. The reduced training time makes it a practical solution for the waste classification task, striking a balance between model complexity and computational efficiency.

G. User Interaction

An interactive component allows users to upload images for real-time classification. The model predicts whether the waste in the images is recyclable or organic, providing probabilities for transparency.

H. Challenges and Constraints:

Training deep learning models, especially large architectures like VGG16, demands substantial computational power. The nature of waste datasets may lead to imbalances in the distribution of classes (Organic vs. Recyclable). Annotating large datasets for waste classification is somewhat time-consuming and resourceintensive. More complex architectures may require longer training times, impacting the speed of model development and iteration.

5 Evaluation

A. Model Performance

After rigorous testing and experimentation, evaluation of various models for waste classification, the customized VGG16 model demonstrated outstanding performance, achieving an impressive accuracy of 95.59%. The comparison with other models is presented below:

Model	Test accuracy
Vgg16 – customized with transfer learning (proposed)	95.59%
Vgg16 without transfer learning	84.63%
Vgg16 with default values and parameters	54%
ResNet18	92%
Vgg19	88%
DenseNet121	89.5%
DenseNet169	88%
Simple CNN	87%

Table 2. Test accuracies for all the evaluated models and the performance

B. Training and Validation Loss Analysis

After repeated experiments and carefully analysing the training and validation loss curves reveal a balanced fit of the model. The training loss consistently decreases, indicating effective convergence during the training process. Simultaneously, the validation loss remains low, demonstrating the model's ability to generalize well to unseen data. This balanced fit suggests that the model is not overfitting to the training set.

C. Overall Testing Accuracy

The customized VGG16 architecture demonstrates exceptional performance, achieving the highest testing accuracy of 95.5911%. Its simplicity enhances interpretability, making it suitable for waste classification. Leveraging transfer learning from ImageNet provides the model with valuable hierarchical features, benefiting various computer vision applications. The success of the customized VGG16 is attributed to its simplicity, tailored binary classification design, regularization techniques, and superior overall performance compared to alternative architectures. These factors collectively contribute to the model's effectiveness in classifying waste images into 'Organic' and 'Recyclable' categories.

```
Test Accuracy: 0.9559
Test Precision: 0.9813
Test Recall: 0.9554
Test F1 Score: 0.9682
```

Table 3. Metric Result for the proposed Vgg16 (Customized) model

D. Integration of Model Testing Results:

To assess the versatility and effectiveness of the proposed waste classification model, extensive testing was conducted on diverse datasets beyond the training and validation sets. The model demonstrated exceptional performance across different datasets sourced from various waste management scenarios. Notably, the accuracy in distinguishing between 'Organic' and 'Recyclable' waste remained consistently high.

The success of the model across varied datasets attests to its adaptability and potential for broader applicability in real-world waste management systems. Moreover, the accuracy achieved in predicting the waste category instils confidence in the model's reliability, substantiating its potential for deployment in practical waste management scenarios.

E. Evaluation Metrics

Confusion Matrix

A confusion matrix is a matrix that evaluates a machine learning model's performance on a set of test data. It is frequently used to assess the performance of classification models, which are designed to predict a category label for each input occurrence. The matrix shows how many true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) the model produced on the test data.

Accuracy: Accuracy is used to measure the performance of the model. It is the ratio of Total correct instances to the total instances [12].

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

Precision: Precision is a measure of how accurate a model's positive predictions are. It is defined as the ratio of true positive predictions to the total number of positive predictions made by the model [12].

$$\text{Precision} = \frac{TP}{TP+FP}$$

Recall: Recall measures the effectiveness of a classification model in identifying all relevant instances from a dataset. It is the ratio of the number of true positive (TP) instances to the sum of true positive and false negative (FN) instances [12].

$$\text{Recall} = \frac{TP}{TP+FN}$$

F1-Score: F1-score is used to evaluate the overall performance of a classification model. It is the harmonic mean of precision and recall [12].

$$\text{F1-Score} = \frac{2 \cdot Precision \cdot Recall}{Precision+Recall}$$

F. Results

Confusion Matric results for different models for which experiments have been performed

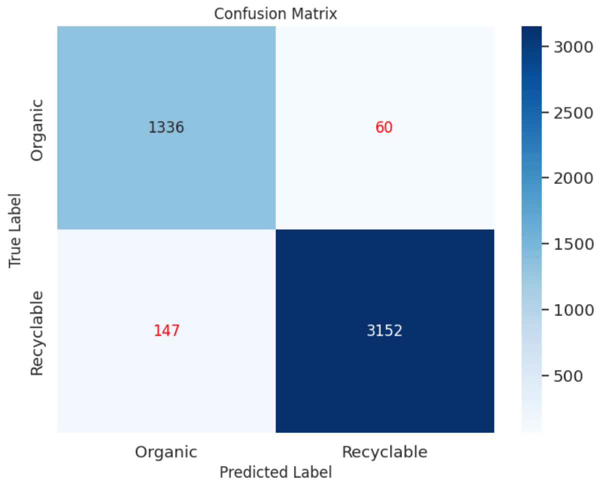


Fig. 5. Confusion matrix for Vgg16 – customized with transfer learning (proposed)

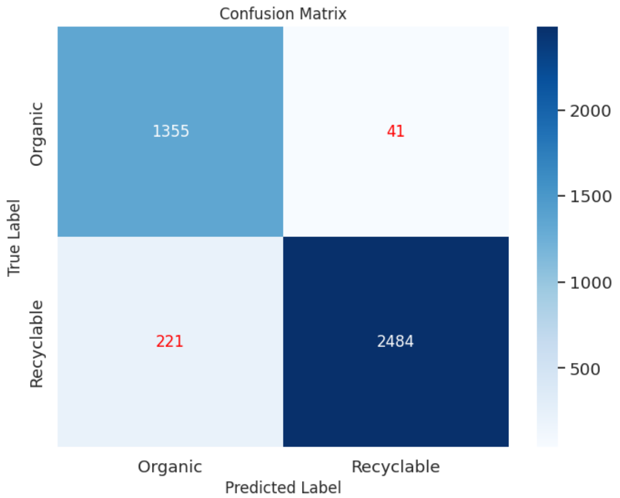


Fig. 6. Confusion matrix for ResNet18

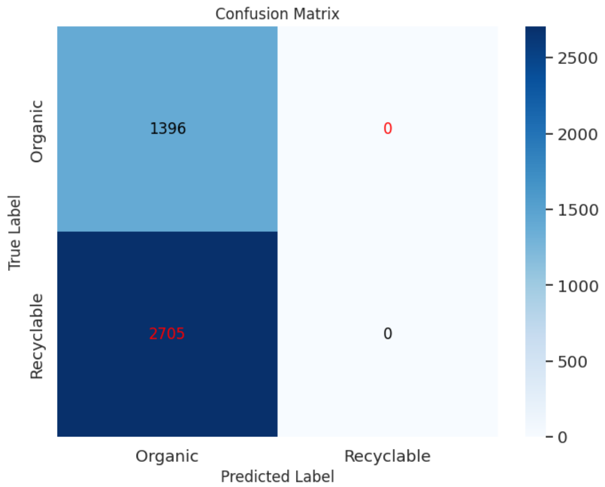


Fig. 7. Confusion matrix for Vgg16 with default values and parameters

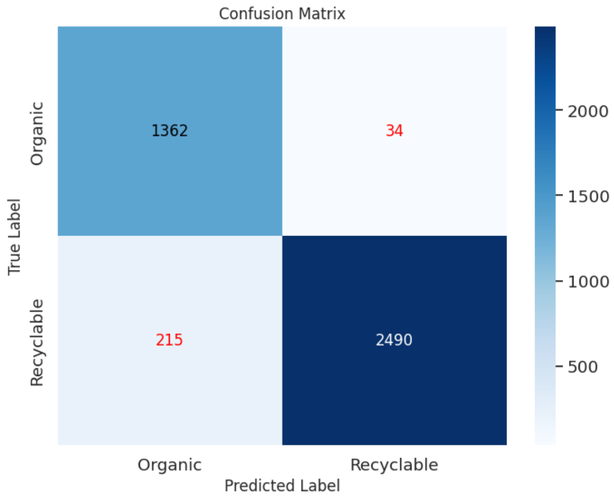


Fig. 8. Confusion matrix for Simple CNN

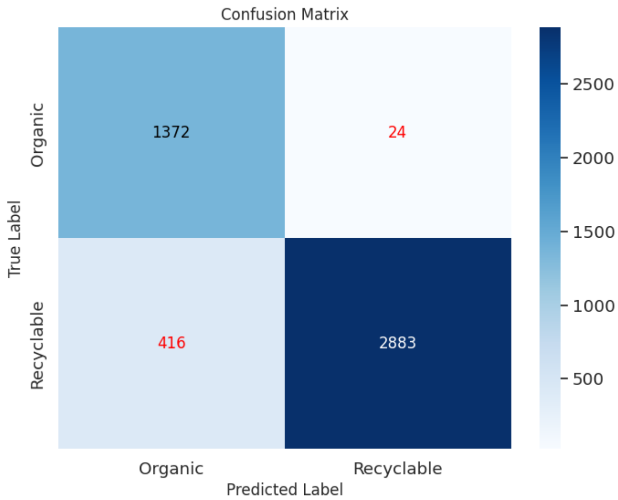


Fig. 9. Confusion matrix for DenseNet model

Training and Validation loss curves for different models that are experimented and implemented during the research

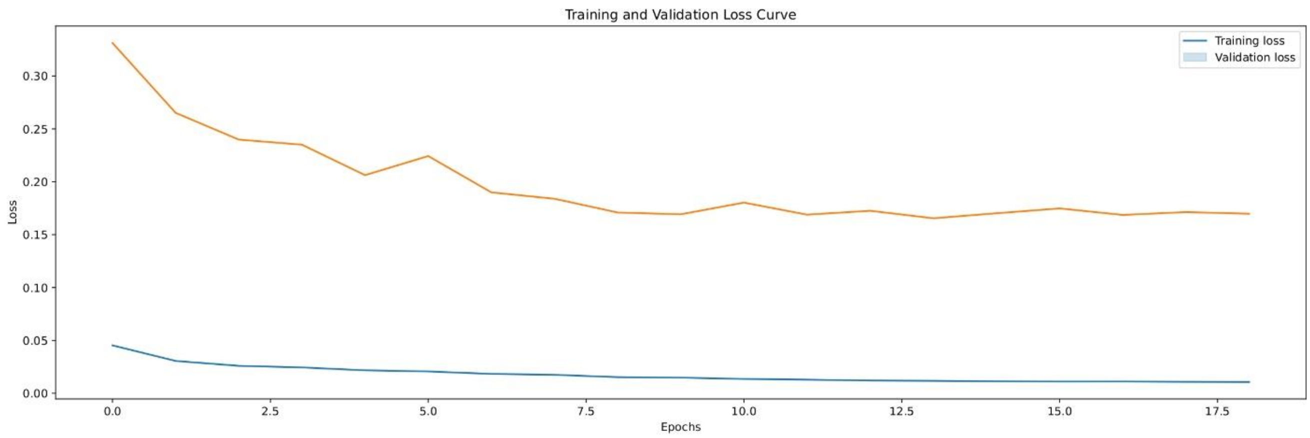


Fig 10. Training and Validation loss – Vg, g16 – customized with transfer learning (proposed)

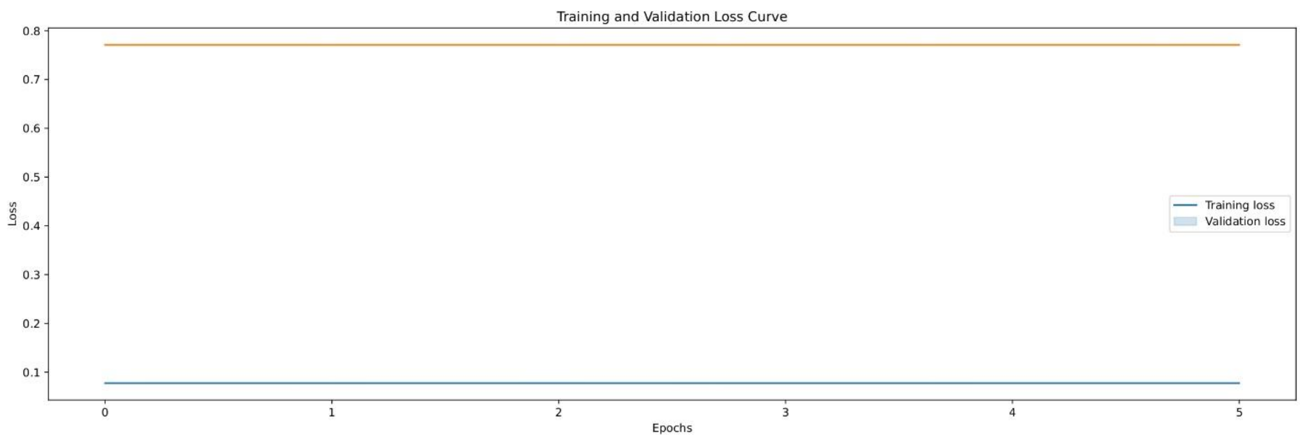


Fig 11. Training and Validation loss - Vgg16 with default values and parameters

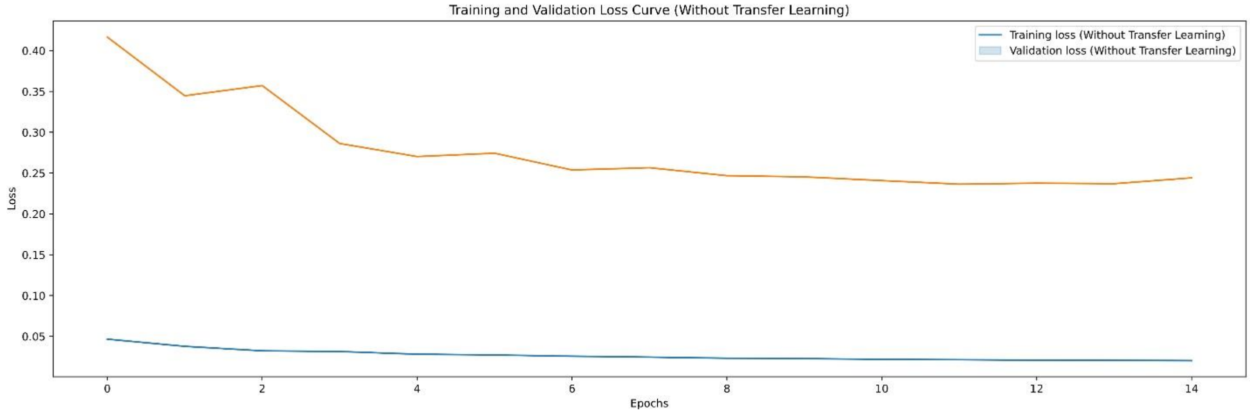


Fig 12. Training and Validation - Vgg16 without transfer learning

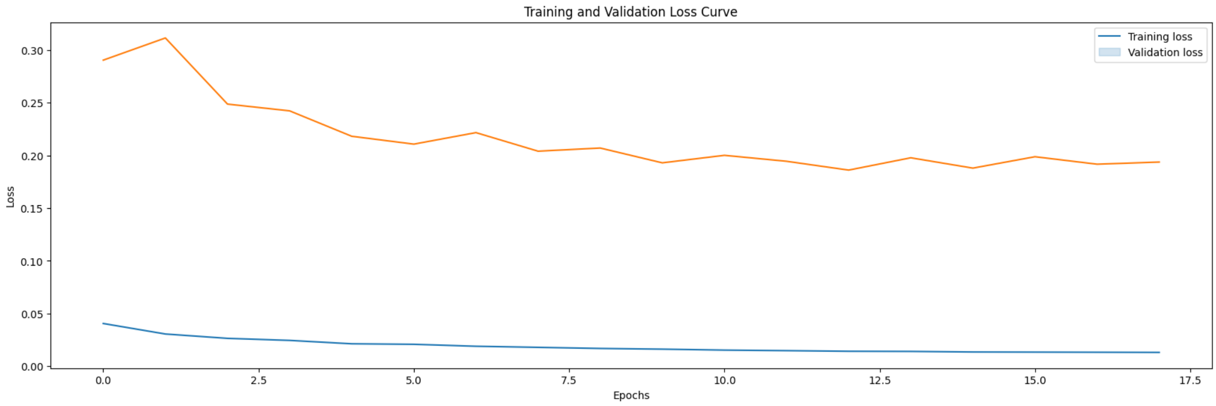


Fig 13. Training and Validation loss – ResNet18

Accuracy Curves for different models:

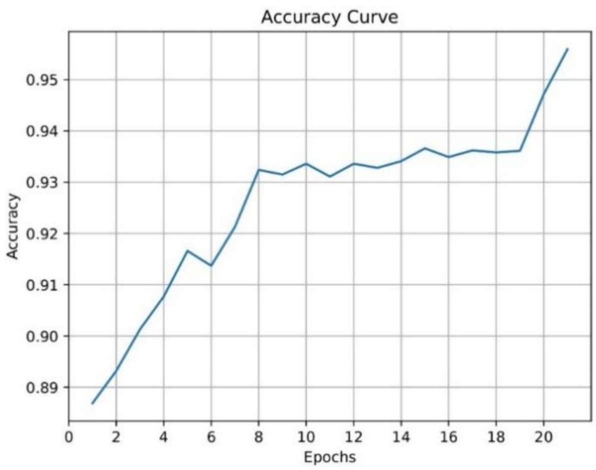


Fig 14. Accuracy curve
Vgg16 – customized with transfer learning

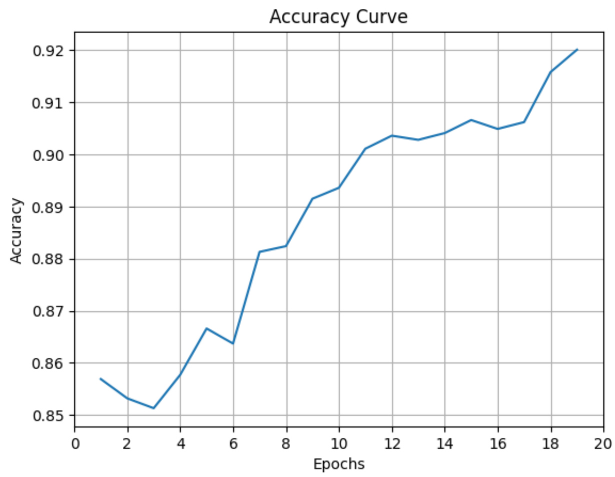


Fig 15. Accuracy curve
ResNet18

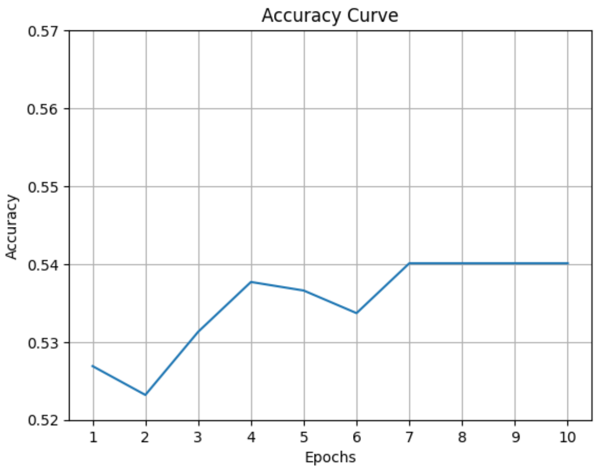


Fig 16. Accuracy curve
Vgg16 with default values and parameters

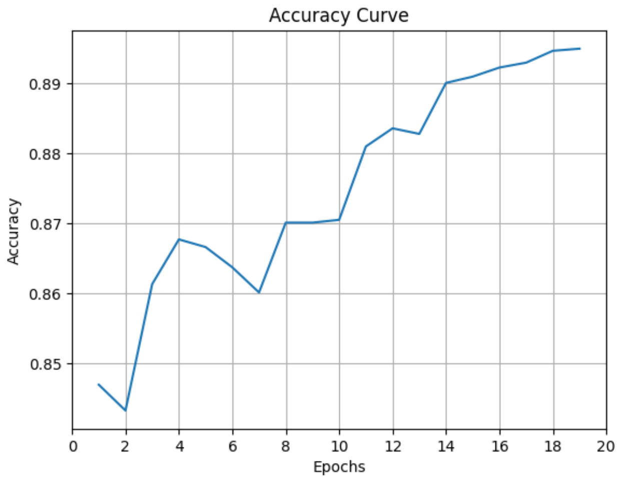


Fig 17. Accuracy curve
DenseNet121

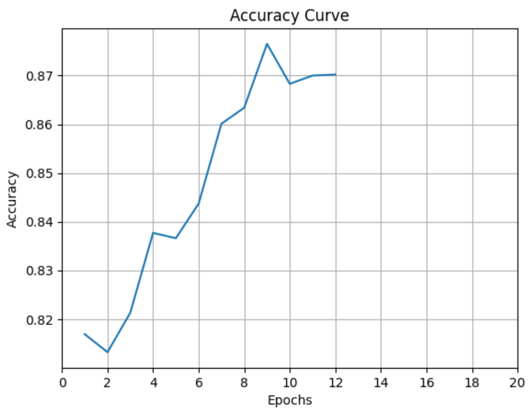


Fig 18. Accuracy curve
Simple CNN with few layers

5. Conclusion and Discussion

The research addresses limitations observed in prior studies on waste image classification, leveraging a dataset comprising over 27K images, significantly surpassing the sample sizes of previous studies. The model does not overfit and the accuracy of 95.59% is indicative of its robust generalization to unseen data, overcoming concerns related to dataset generalizability and diversity. To enhance spatial relationship handling, a customized VGG16 architecture is employed, incorporating transfer learning, early stopping, a learning rate scheduler, dropout layers, and other optimizations. The evaluation metrics such as precision, recall, and F1-score, contributes to a more thorough performance assessment. Notably, the experimentation phase involves the implementation and analysis of multiple architectures, aiming to select the most suitable model for the waste image classification task. These advancements highlight the novelty of this research in the field of waste management and classification.

The model's effectiveness across multiple data sets demonstrates its adaptability and the robust performance encourages the reuse of the model. While the model exhibits promising performance, it is crucial to acknowledge certain limitations. One notable consideration is the model's dependency on hardware resources, particularly a larger GPU for optimal performance.

This hardware requirement could pose a limitation for users with constrained computational resources. The model's performance heavily relies on the quantity and diversity of the training data. A larger and more diverse dataset could further enhance its generalization capabilities and performance accuracy. If the dataset has a significant class imbalance, the model may show a bias towards the majority class. Addressing class imbalance could improve overall model performance.

Future work could focus on hyperparameter optimization, utilizing larger datasets, and implementing advanced strategies such as ensemble learning techniques for ongoing improvements in waste classification systems, with possibilities for real-world applications.

References

[1] Sekar, S. (2019). Waste classification data. Kaggle. https://www.kaggle.com/datasets/techsash/waste-classification-data

[2] CCHANG. (2018). Garbage Classification. Kaggle. https://doi.org/10.34740/KAGGLE/DS/81794

[3] Ozkefe, F. (2021). TrashNet. Kaggle. https://www.kaggle.com/datasets/feyzazkefe/trashnet

[4] Huang, L., Li, M., Xu, T., et al. (2023). A waste classification method based on a capsule network. **Environmental Science and Pollution Research, 30**(86454–86462). https://doi.org/10.1007/s11356-023-27970-7

[5] Gyawali, D., Regmi, A., Shakya, A., Gautam, A., & Shrestha, S. (2020). Comparative Analysis of Multiple Deep CNN Models for Waste Classification. **arXiv:2004.02168**. https://doi.org/10.48550/arXiv.2004.02168.

[6] Shi, C., Tan, C., Wang, T., & Wang, L. (2021). A Waste Classification Method Based on a Multilayer Hybrid Convolution Neural Network. **Applied Sciences, 11**(88572). https://doi.org/10.3390/app11188572

[7] Olugboja Adedeji, Zenghui Wang. (2019). Intelligent Waste Classification System Using Deep Learning Convolutional Neural Network. **Procedia Manufacturing, 35**, 607-612. https://doi.org/10.1016/j.promfg.2019.05.086. https://www.sciencedirect.com/science/article/pii/S2351978919307231

[8] Faria, R., Ahmed, F., Das, A., & Dey, A. (2021). Classification of Organic and Solid Waste Using Deep Convolutional Neural Networks. In **2021 IEEE 9th Region 10 Humanitarian Technology Conference (R10-HTC)**, pp. 01-06. Doi: 10.1109/R10-HTC53172.2021.9641560.

[9] Bhadra, J., & Lawrence DLima, A. (2023). Classification of Organic and Recyclable Waste for Sustainable Development using Resnet50 Model. In **2023 International Conference on Advances in Electronics, Communication, Computing, and**

Intelligent Information Systems (ICAECIS), pp. 78-83. Doi: 10.1109/ICAECIS58353.2023.10170501.

[10] Ma'Rifah, P. N., Sarosa, M., & Rohadi, E. (2023). Garbage Classification using Faster R-CNN. In **2023 International Conference on Electrical and Information Technology (IEIT),** pp. 196-201. Doi: 10.1109/IEIT59852.2023.10335519.

[11] Mulim, W., Revikasha, M. F., Rivandi, & Hanafiah, N. (2021). Waste Classification Using EfficientNet-B0. In **2021 1st International Conference on Computer Science and Artificial Intelligence (ICCSAI)**, pp. 253-257. Doi: 10.1109/ICCSAI53272.2021.9609756.

[12] Sukel, M., Rudinac, S., Worring, M. (2023). GIGO, Garbage In, Garbage Out: An Urban Garbage Classification Dataset. In **MMM 2023: Multimedia Modelling,** 13833, 41-52. https://doi.org/10.1007/978-3-031-27077-2_41

[13] Poudel, S., & Poudyal, P. (2022). Classification of Waste Materials using CNN Based on Transfer Learning. In **Forum for Information Retrieval Evaluation (FIRE '22).** ACM. https://doi.org/10.1145/3574318.3574345

[14] Pandey, A., Jain, H., Raj, H., & Gupta, P. P. (2023). Identification and Classification of Waste using CNN in Waste Management. In **2023 IEEE 8th International Conference for Convergence in Technology (I2CT),** pp. 1-6. Doi: 10.1109/I2CT57861.2023.10126312.

[15] Girsang, S., Yunanda, R., Syahputra, M. E., & Peranginangin, E. (2022). Convolutional Neural Network Using Res-Net For Organic and Anorganic Waste Classification. In **2022 IEEE International Conference of Computer Science and Information Technology (ICOSNIKOM),** pp. 01-06. Doi: 10.1109/ICOSNIKOM56551.2022.10034869.

[16] Qiuhao, Z. (2021). Kitchen Waste Classification Based on Deep Residual Network and Transfer Learning. In **2021 6th International Symposium on Computer and Information Processing Technology (ISCIPT)**, pp. 625-629. Doi: 10.1109/ISCIPT53667.2021.00133.

[17] Verma, M., Kumar, A., & Kumar, S. (2022). Medical Waste Classification using Deep Learning and Convolutional Neural Networks. In **2022 IEEE Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)**, pp. 1-5. Doi: 10.1109/IATMSI56455.2022.10119431.

[18] Shah, J., & Kamat, S. (2022). A Method for Waste Segregation using Convolutional Neural Networks. In **2022 Second International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT)**, pp. 1-5. Doi: 10.1109/ICAECT54875.2022.9807969.

[19] S, L., & Usha, D. (2023). An Automated Approach to Waste Classification Using Deep Learning. In **2023 Fifth International Conference on Electrical,**

Computer and Communication Technologies (ICECCT), pp. 01-10. Doi: 10.1109/ICECCT56650.2023.10179743.

[20] Huynh, M. -H., Pham-Hoai, P. -T., Tran, A. -K., & Nguyen, T. -D. (2020). Automated Waste Sorting Using Convolutional Neural Network. In **2020 7th NAFOSTED Conference on Information and Computer Science (NICS),** pp. 102-107. Doi: 10.1109/NICS51282.2020.9335897.

[21] Rahman, N., & Das, S. K. (2022). A Fusion of Three Custom-Tailored Deep Learning Architectures for Waste Classification. In **2022 4th International Conference on Sustainable Technologies for Industry 4.0 (STI)**, pp. 1-6. Doi: 10.1109/STI56238.2022.10103297.

[22] Simonyan, K., & Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. **ICLR**. [ArXiv:1409.1556](https://arxiv.org/abs/1409.1556)

[23] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. **CVPR**, 770-778. [DOI:10.1109/CVPR.2016.90](https://doi.org/10.1109/CVPR.2016.90)

[24] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely Connected Convolutional Networks. **CVPR**, 4700-4708. [DOI: 10.1109/CVPR.2017.243](https://doi.org/10.1109/CVPR.2017.243)

[25] Geeks for Geeks. (2023, March 21). Confusion matrix in Machine Learning. **Geeks for Geeks**. [https://www.geeksforgeeks.org/confusion-matrix-machine-learning/]

[26] Abdu, H., & Mohd Noor, M. H. (2022). A Survey on Waste Detection and Classification Using Deep Learning. **IEEE Access, 10**, 128151-128165. Doi: 10.1109/ACCESS.2022.3226682.

[27] 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.

[28] Ba Alawi, E., Saeed, A. Y. A., Almashhor, F., Al-Shathely, R., & Hassan, A. N. (2021). Solid Waste Classification Using Deep Learning Techniques. In **2021 International Congress of Advanced Technology and Engineering (ICOTEN)**, pp. 1-5. Doi: 10.1109/ICOTEN52080.2021.9493430.

[29] Bhattacharya, S., Sai, K. B., Puvirajan, H. S., Peera, H., & Jyothi, G. (2023). Automated Garbage Classification using Deep Learning. In **2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)**, pp. 404-410. Doi: 10.1109/ICAAIC56838.2023.10141483.

[30] PyTorch. (n.d.). torch.optim.lr_scheduler.StepLR. Retrieved from https://pytorch.org/docs/stable/generated/torch.optim.lr_scheduler.StepLR.html

[31] Kingma, D. P., & Ba, J. (2015). Adam: A Method for Stochastic Optimization. **arXiv preprint arXiv:1412.6980 [cs.LG]**. https://doi.org/10.48550/arXiv.1412.6980

[32] Figure 11. Binary Cross Entropy loss calculation. Source: Twitter (https://pbs.twimg.com/media/FNBiaSRXEAAyrzC.png, accessed on December 10, 2023).