Engineering Applications of Artificial Intelligence 162 (2025) 112522

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

# PLDs-CNN-ridge-ELM: Interpretable lightweight waste classification framework

Mansura Naznine [a], Md. Nahiduzzaman [b], Md. Jawadul Karim [b], Md. Faysal Ahamed [b], Abdus Salam [b], Mohamed Arselene Ayari [c], Amith Khandakar [d], Azad Ashraf [e], Mominul Ahsan [f], Julfikar Haider [g,*]

[a] Department of Computer Science & Engineering, Rajshahi University of Engineering & Technology, Rajshahi, 6204, Bangladesh
[b] Department of Electrical and Computer Engineering, Rajshahi University of Engineering & Technology, Rajshahi, 6204, Bangladesh
[c] Civil and Environmental Engineering, College of Engineering, Qatar University, Doha, Qatar
[d] Department of Electrical Engineering, College of Engineering, Qatar University, Doha, Qatar
[e] Chemical Engineering Department, University of Doha for Science and Technology, Doha, Qatar
[f] Department of Computer Science, University of York, Deramore Lane, Heslington, York, YO10 5GH, UK
[g] Department of Engineering, Manchester Metropolitan University, Chester Street, Manchester, M1 5GD, UK

ARTICLE INFO

*Keywords:*
Waste classification
PLDs-CNN-Ridge-ELM
Hardware architecture
Graphical user interface
And shapley additive explanations

ABSTRACT

The accelerating global population growth and expanding economic activities have resulted in a notable increase in waste generation, necessitating accurate and efficient waste classification systems for sustainable waste management. This research presents a novel two-stage waste classification model leveraging a Lightweight Parallel Depth-wise Separable Convolutional Neural Network (PLDs-CNN), combined with a Ridge Regression Extreme Learning Machine (Ridge-ELM) classifier, using waste images as input. The proposed system efficiently classifies waste into four primary categories (hazardous, household, recyclable, and residual) in the first stage and further refines the classification into twelve subcategories in the second stage. Featuring a lightweight architecture of nine layers and about 1.09 million parameters, the PLDs-CNN model achieves high accuracy with substantially reduced computational overhead, outperforming many deeper networks. In the four-class classification stage, the system achieves an average accuracy of 99 %, with precision, recall, F1-score, and receiver operating characteristics (ROC)-area under the curve (AUC) values of 97.25 ± 0.02 %, 96 ± 0.03 %, 96.5 ± 0.01 %, and 99.28 %, respectively. In the twelve-class classification, the model continues to deliver superior results, with 96 % accuracy and equally strong precision, recall, and F1-score metrics. The system is supported by a real-time hardware architecture, featuring a user-centric Graphical User Interface (GUI), a webcam-enabled conveyor belt sorting mechanism, and a 2-axis pan-tilt system for automated waste sorting. Additionally, the model's interpretability is significantly improved through the integration of Shapley Additive Explanations (SHAP), which provides important perspectives into the decision-making process, increasing transparency and trustworthiness in real-world applications. The proposed framework not only surpasses conventional methods in both accuracy and computational efficiency but also emphasizes sustainability by facilitating cost-effective and scalable waste management solutions aimed at promoting recycling and resource reuse.

## 1. Introduction

The rapid growth of global populations and economies has led to a significant rise in resource consumption, resulting in a concerning increase in waste generation (Abuga and Raghava, 2021; Özkan et al., 2015). Based on prior studies, it is expected that the worldwide production of solid waste will reach 2.2 billion tons annually by 2025, which will require a budget of $375.5 billion for waste management (Hoornweg and Bhada-Tata, 2012). Thirty-four percent of the world's municipal solid garbage is produced by 16 % of the population in

* Corresponding author.
*E-mail addresses:* 1703008@student.ruet.ac.bd (M. Naznine), nahiduzzaman@ece.ruet.ac.bd (Md. Nahiduzzaman), 1710021@student.ruet.ac.bd (Md.J. Karim), faysal.ahamed@ece.ruet.ac.bd (Md.F. Ahamed), 1710035@student.ruet.ac.bd (A. Salam), ArslanA@qu.edu.qa (M.A. Ayari), amitk@qu.edu.qa (A. Khandakar), azad.ashraf@udst.edu.qa (A. Ashraf), mominul.ahsan@york.ac.uk (M. Ahsan), j.haider@mmu.ac.uk (J. Haider).

https://doi.org/10.1016/j.engappai.2025.112522
Received 29 November 2024; Received in revised form 1 June 2025; Accepted 24 September 2025
Available online 8 October 2025

developed countries (Jaunich et al., 2019). Compared to the global average waste of 0.74 kg per day, these countries generate almost 2.1 kg of waste per person on a daily basis (Jaunich et al., 2019). Waste production is anticipated to grow significantly, including in low- and middle-income regions such as Africa and certain areas of Asia. These countries produce approximately 35 % of the world's solid waste (Kaza and Bhada-Tata, 2018). Unfortunately, there is no efficient and automated system for waste disposal to address this ever-expanding problem (Li and Zhang, 2024). In cities, waste handling authorities gather waste materials from residential waste containers and manually sort them for recycling purposes or dispose of them in landfills. The United Nations Environment Program (UNEP) has identified this as a substantial issue that has negative effects on economic progress, human communities, and public well-being (Harris et al., 2021). The inadequate disposal of waste, specifically through landfills and the burning process, presents a substantial danger to urban ecosystems and the welfare of inhabitants (Wang et al., 2020). The harmful effects of unlimited waste production, such as the buildup of dangerous compounds and widespread plastic pollution, emphasize the need for scientific solutions in waste management. Recycling and composting are the primary methods of sustainable waste management. However, less than 19 % of waste is reused through recycling and composting globally, while approximately 40 % of waste ends up in landfills (Kaza et al., 2018). A thorough knowledge of waste classification is essential for the implementation of an efficient waste control system, as there exists numerous distinct types of waste. Therefore, many countries have initiated research on intelligent garbage classification and recycling technologies (Carrera et al., 2022; Cheah et al., 2022; Jiang et al., 2023; Kang et al., 2020; Lu et al., 2020).

The amount of waste produced in Bangladesh increased from 1100000 tons in 1970–14778497 tons in 2012, indicating a 134300 ton annual increase (Ashikuzzaman and Howlader, 2020; Kaza et al., 2018; Shams et al., 2017). According to recent data, the average amount of solid waste produced per person in different areas of Bangladesh varies between 0.2 and 0.56 kg (Ahsan et al., 2014). Dhaka, the capital city, generated an average of 6448.37 tons of solid waste each day from 2016 to 2017 (Jerin et al., 2022). Although the City Corporation collects approximately 50 % of Dhaka's waste, a substantial portion—estimated between 40 % and 60 %—remains uncollected and is subject to improper disposal practices. This uncollected waste consists of approximately 80 % organic material (Ahsan et al., 2014). By 2025, the urban population is expected to reach 78.44 million, and the rate of trash production is estimated to increase to 220 kg per capita per year (Habib et al., 2021). The government has initiated the implementation of the National 3R (Reduce, Reuse, and Recycle) Strategy to solve the waste management problem (Haque and Razy, 2021). Another organization, "Waste Concern", is a social business enterprise that has emerged to address the issue of municipal garbage accumulation by collaborating with families. The United Nations International Children's Emergency Fund (UNICEF) has also implemented recycling programs and waste management initiatives together with city corporations and municipalities. Nevertheless, there are currently insufficient efforts to improve these standards. Factors such as land scarcity and insufficient technical skills have further worsened the problem of dealing with large amounts of garbage. To effectively address the escalating waste management challenges in Bangladesh, the integration of advanced technologies such as automated sorting systems, sensor-based methods, and artificial intelligence (AI) presents a promising solution to improve waste separation and recycling accuracy. Bangladesh's significant constraints—characterized by limited financial and technical resources alongside inadequate waste sorting infrastructure—make it a critical context for assessing the feasibility and impact of cost-effective, AI-driven waste classification systems. Selecting Bangladesh as a case study not only reflects the urgent need for innovative waste management strategies within the country but also allows the findings to be generalized to other low- and middle-income nations facing similar environmental and infrastructural challenges.

As previously indicated, in most cases, manual waste sorting processes are still employed by authorities. However, in smart cities, the development of automated circular economy (CE) systems for waste management is essential to maintain sustainability. Automatic waste classification systems provide efficient and accurate waste categorization, significantly reducing manual labor and processing time. By integrating AI, this system can effectively recover valuable resources for recycling and reuse and minimize environmental impact while offering long-term cost savings and scalability. The advancement of AI has introduced novel concepts to this domain. Several researchers have successfully employed Convolutional Neural Networks (CNN) for precise waste categorization (Lin et al., 2022; Y.-L. Zhang et al., 2023; Lin et al., 2022b), resulting in a range of notable accomplishments. For waste detection and classification, some studies have employed You Only Look Once (YOLO)-based models (Y. Chen et al., 2023; Mao et al., 2022; Qiao et al., 2023; J. Yang et al., 2023; Q. Zhang et al., 2022; Qiao et al., 2023) Some also employed Transfer Learning (TL) based models (Mao et al., 2021; Q. Zhang et al., 2021), which have shown high accuracy performance, ranging from 87 % to 96 %. Nevertheless, these TL models possess higher parameters (1.2–1.3 million) (Z. Chen et al., 2022; Feng et al., 2022), and the process of determining them necessitates a substantial number of floating-point operations (Gaba et al., 2022). Additionally, majority of the current research used datasets with limited classes (3–6 classes) and a limited number of images (fewer than 5000 images). As a result, implementing real-world applications based on these models is a challenging task. Hence, researchers have begun investigating computationally efficient models for waste classification, specifically designed for deployment on resource-limited platforms such as embedded devices. Nevertheless, more research on lightweight waste classification models is needed. The existing lightweight models have shown lower classification accuracy (83–95 %) on large datasets, which can lead to the misclassification of waste elements. Additionally, many of the existing studies did not explore the potential for real-life implementation of the proposed models (Z. Chen et al., 2022; Feng et al., 2022; Gaba et al., 2022; Mao et al., 2021; Z. Yang and Li, 2020; Q. Zhang et al., 2021). To overcome these challenges, it is imperative to implement a lightweight model with minimal parameters, layers, and size.

The present study proposed a two-stage parallel lightweight depth wise separable CNN (PLDs-CNN) model. The motivation behind this study is that the existing manual waste classification system is not capable of rapidly categorizing waste materials and directing them toward the recycling process automatically. Hence, there is a need for an AI-based system that can perform this task accurately and rapidly. The PLDs-CNN feature extractor, which integrates depth-wise separable convolution layers, is both lightweight and demonstrates high performance with low computational requirements. All these characteristics make the proposed model faster and more accurate, which can improve the productivity of waste identification. The implementation of SHAP (Shapley Additive Explanations) and a laboratory-scale hardware system validates the application of the proposed model in real-world settings. The major contributions of this paper are outlined as follows.

1. **Efficient Parallel Lightweight CNN Architecture (PLDs-CNN):** A novel two-stage parallel CNN architecture has been designed using depthwise separable convolutions and multiscale kernels to capture diverse spatial features while significantly reducing parameter count and computational complexity. This allows for fast, low-resource inference, making it well-suited for real-time applications, particularly on resource-constrained edge devices.
2. **Enhanced Classifier Integration via Modified Ridge-ELM:** The traditional softmax classifier is replaced with a Ridge-ELM, which provides significantly faster training (no backpropagation), better generalization, and superior multiclass classification performance. Ridge regularization improves stability and robustness in learning, especially when dealing with complex or high-dimensional features.

3. **Interpretability with SHAP in Real-Time:** For the first time in waste classification, the model integrates Shapley Additive Explanations (SHAP) to offer real-time, pixel-level interpretability. This contributes to transparency throughout the decision-making phase and increases trust in practical deployment settings.
4. **Hardware Implementation with GUI for Real-World Validation:** The proposed model has been deployed on hardware with a real-time graphical user interface, validating its operational usability. Despite being a lightweight model, it demonstrates high accuracy and fast response, confirming its practical viability.

Section 2 of this research presents an in-depth overview of the prior relevant studies conducted on this topic. Section 3 presents the proposed methodology, which includes a comprehensive framework, a description of the dataset, feature extraction methods, and performance metrics. Section 4 provides a detailed presentation of the comprehensive classification results, along with a detailed explanation of the interpretability of the proposed framework using SHAP and hardware-software structure for real-life implementation. Concluding remarks are detailed in Section 5.

## 2. Overview of previous work

To date, researchers have proposed several lightweight and deep learning (DL) models for automated classification of wastes (Q. Zhang et al., 2021; S. Zhang et al., 2021). Mao et al., developed an enhanced DenseNet121 model by applying a genetic algorithm to optimize its fully connected layer. Their work utilized the TrashNet dataset, which includes 2527 images categorized into six classes (Mao et al., 2021). To improve classification accuracy, data augmentation techniques were applied to expand the training set. The resulting model achieved an impressive accuracy of 99.60 %, with a training duration of 5542 s. Another method for waste classification was proposed by Zhang et al. (Q. Zhang et al., 2021), who employed a DenseNet169 model with transfer learning. This study used the NWNU-TRASH dataset, consisting of 2528 images divided into five classes, with a 70/30 training-testing split. Their model attained an accuracy exceeding 82 %, outperforming multiple previous approaches. Nevertheless, this method faced limitations due to the relatively small dataset, imbalanced class distribution, and the model's high complexity from a large parameter count, which posed challenges for practical deployment. Khan et al. (2022) proposed another waste classification approach, called recycling waste classification using emperor penguin optimizer with DL (RWC-EPODL), which utilizes the emperor penguin optimizer model to generate bioenergy from recyclable garbage. This model employed AX-RetinaNet for object identification and used a stacked auto-encoder (SAE) for classification. The study achieved a success rate of 98.96 % on Kaggle's garbage categorization dataset, which comprises 750 images categorized into six distinct classes. Lin et al. (Lin et al., 2022) employed different Residual Neural Network (ResNet) architectures derived from TL models to classify waste on the TrashNet dataset. Although they achieved 88.8 % accuracy, the extensive parameter count negatively impacts the model's performance. On average, it took approximately 700 s to train for one epoch.

Using modern technologies, some researchers have created intelligent waste categorization tools that have developed automated and effective garbage classification systems in real life. These solutions are intended to improve garbage management procedures using advanced algorithms to achieve higher classification accuracy and promote a sustainable environment. Chen et al. (Z. Chen et al., 2022) proposed a garbage classification network (GCNet) based on improved ShuffleNetv2. By employing the parallel mixed attention mechanism (PMAM), incorporating novel activation functions Rectified linear unit (FReLU), and leveraging TL, they enhanced the model's performance and achieved an exceptional accuracy of 97.9 % on their custom dataset. The dataset comprises 4256 photos, which were divided into 14 distinct subcategories. The proposed model can categorize waste into four separate groups: recyclable waste, wet waste, hazardous waste, and dry waste. The categorization process required 0.88 s, utilizing 1.3 million parameters. Similarly, Fan et al. (2023) designed an intelligent garbage bin that separates regular household waste into four categories. The system consists of an automated image classification system that utilizes a Raspberry Pi unit, a digital camera, and three rotating plates. The image classification technique used the EfficientNetB2 model in combination with the PMAM to obtain high accuracy. Additionally, researchers have proposed a background noise removal (BNR) approach to address the impact of environmental factors on garbage recognition. Their results showed a classification accuracy of 93.38 % on the Huawei Cloud Garbage Classification dataset. Feng et al. (2022) employed the Generalized Error Correction Model (GECM)-EfficientNet model for effective waste classification to create an intelligent waste bin and achieved high accuracy (94.54 % and 94.23 %) on self-built and TrashNet datasets with 1.23 million parameters. Based on EfficientNet, GECM-EfficientNet uses TL, efficient channel attention (ECA) and coordinate attention (CA) modules and streamlining techniques to achieve better accuracy and real-time performance. The waste bin had a camera and servos for sorting waste into fan-shaped bins. It was operated by a Raspberry Pi 4B.

Jin et al. (2023) designed a device utilizing DL techniques to facilitate sustainable garbage recycling. Their model utilized MobileNetV2 as the main framework, incorporating one convolutional block attention module (CBAM), one principal component analysis (PCA) module, and one fully connected classification layer. The proposed approach significantly decreases the time required for garbage identification by 170 ms compared to the conventional MobileNetV2 network and effectively categorizes garbage into four distinct categories. Using the Huawei Cloud Garbage dataset, which includes 14683 pictures, the recommended approach achieved 90.7 % accuracy. Similarly, Zhang et al. (S. Zhang et al., 2021) developed an automated waste sorting device for categorizing domestic waste. The authors presented a two-step trash recognition-retrieval technique adopting the Visual Geometry Group (VGG16) model. A dataset including 1040 waste images was created, and various data augmentation technique was employed. To mitigate the issue of overfitting, they utilized the ten-fold cross-validation technique. The model categorized 13 different forms of waste into four distinct categories. According to the experimental results, the model's average accuracy was 94.71 %.

Abdulkareem et al. (2024) introduced a two-stage intelligent waste decision framework using DL models and Multi-Criteria Decision Making (MCDM). This work introduced a Multi-Fused Decision Matrix (MFDM) to evaluate and select optimal deep Waste Sorting Models (WSMs) based on different fusion rules by experimenting on a dataset of 1451 waste pictures across four classes. The results indicated that the hybrid Inception-Xception model works better than the other models, while the ResNet50-GoogleNet-Inception model achieved an outstanding accuracy of 98 %. Another study (Mohammed et al., 2023) presented a novel automated waste sorting and recycling classification system leveraging an Artificial Neural Network (ANN) and Feature Fusion (including color, Local Binary Patterns (LBP), Histogram of Oriented Gradients (HOG), and Uniform LBP). With an impressive accuracy of 91.7 %, the model effectively classified test waste images into three classes. Similarly, Kumar et al., introduced a system for automatically sorting and classifying COVID-19-related medical waste from other waste materials using an Artificial Neural Network (ANN) and Feature Fusion. Al-Mashhadani et al. (Al-Mashhadani, 2023) assessed the performance of ResNet50, GoogleNet, InceptionV3, and Xception DL models in waste classification and achieved excellent accuracy and precision. They utilized a dataset consisting of 1451 photos that were divided into four waste categories. ResNet50 achieved a classification accuracy and precision of 95 %, while InceptionV3 achieved a perfect classification result of 100 % in all categories. Another study (Rahman et al., 2023) discussed the difficulties related to the management of solid

**Table 1**
Literature review with research gaps.

| Ref. | Model | Dataset | Result | Limitation |
|---|---|---|---|---|
| Mao et al. (2021) | Optimized DenseNet121 with Genetic Algorithm | TrashNet (2527 images, 6 classes) | 99.60 % accuracy | Small dataset, large number of parameters, limited scalability |
| Q. Zhang et al. (2021) | DenseNet169 | NWNU-TRASH (2528 images, 5 classes) | >82 % accuracy | Small dataset, high model complexity |
| Khan et al. (2022) | AX-RetinaNet + Stacked Autoencoder (RWC-EPODL) | Kaggle Garbage Dataset (750 images, 6 classes) | 98.96 % accuracy | Limited dataset size, no explainability |
| Lin et al. (2022) | ResNet-based TL. | TrashNet (2527 images, 6 classes) | 88.8 % accuracy | Large parameter count, ~700s/ epoch training |
| Z. Chen et al. (2022) | GCNet (Improved ShuffleNetv2 + PMAM + FReLU) | Custom (4256 images, 14 subclasses, 4 main classes) | 97.9 % accuracy | Not publicly available dataset, lacks explainability |
| Fan et al. (2023) | EfficientNetB2 + PMAM | Huawei Cloud Garbage Dataset (14802 images) | 93.38 % accuracy | Large model (7.8M params) |
| S. Zhang et al. (2021) | VGG16 + Two-step Retrieval + Cross-validation | Custom (1040 images, 13 subclasses) | 94.71 % accuracy | Small dataset, prone to overfitting despite augmentation |
| Abdulkareem et al. (2024) | Inception-Xception + MCDM + MFDM | Custom (1451 images, 4 classes) | 98 % accuracy | Limited dataset, no real-time capability, no explainability |
| Mohammed et al. (2023) | ANN + Feature Fusion (LBP, HOG, color) | Custom (3 classes) | 91.7 % accuracy | Limited to 3 classes, lacks scalability |
| Kumar et al. (2021) | ANN + Feature Fusion for COVID Waste | Custom | Not specified | Designed for niche waste category (medical) only |
| Al-Mashhadani (2023) | ResNet50, InceptionV3, etc. | Custom (1451 images, 4 classes) | Up to 100 % accuracy | Extremely deep networks, not lightweight, no XAI |

waste in metropolitan areas because of substantial population growth.

According to the state-of-the-art models, most researchers have employed DL and TL models with high numbers of model parameters and layers. For instance, Feng et al. (2022) utilized GECM-EfficientNet with 1.23 million parameters, whereas the authors in (Z. Chen et al., 2022) employed GCNet with 1.3 million parameters. Al-Mashhadani et al. (Al-Mashhadani, 2023) employed ResNet50, GoogleNet, InceptionV3, and Xception with 50, 22, 48 and 71 layers, respectively. Thus, training these models on GPUs necessitates substantial time investment. The proposed method in (Mao et al., 2021) required 5542 s for training purposes. Similarly, in (Lin et al., 2022), the authors utilized various ResNet architectures that required 7000 s for training.

Undoubtedly, implementing these algorithms in real life is challenging. To develop a practical, cost-effective, intelligent waste sorting system, it is necessary to design a computationally efficient model having a reduced number of parameters and layers, enabling shorter training times compared to existing models. Again, the literature shows that certain studies were able to achieve higher classification accuracy (Mao et al., 2021). However, they used a dataset with a small number of classes and images to demonstrate their proposed model. Abdulkareem et al. (2024) utilized a dataset that consists of only 1451 images of four distinct classes. Similarly, in (Mohammed et al., 2023), the authors evaluated their proposed models' performances on a smaller dataset, which contains 2400 images from three classes. Although some datasets in the literature have 13, 14 or 18 subclasses, these datasets are not publicly accessible (Chen et al., 2022; Feng et al., 2022; S. Zhang et al., 2021). Furthermore, no studies have demonstrated the use of real-time explainable AI, such as SHAP or LIME, emphasizing the impact of individual features. To incorporate recent studies and better highlight existing research gaps, an updated comparative analysis of state-of-the-art waste classification models is presented in Table 1. This table outlines the model architectures, datasets used, achieved results, and limitations of various prominent works.

**Table 2**
Overall datasets on both four classes and twelve subclasses.

| Testing phase | Trash Type | | Training | Testing | Validation |
|---|---|---|---|---|---|
| First Stage: 4 classes | Hazardous Waste (0)[a] | | 766 | 94 | 85 |
| | Household Food Waste (1) | | 797 | 99 | 89 |
| | Recyclable Waste (2) | | 10439 | 1289 | 1160 |
| | Residual Waste (3) | | 564 | 70 | 63 |
| | **Total** | | **12566** | **1552** | **1397** |
| Second Stage: 12 subclasses | Battery (0)[a] | Hazardous Waste | 766 | 94 | 85 |
| | Expired Food (1) | Household Food Waste | 797 | 99 | 89 |
| | Brown Glass (2) | Recyclable Waste | 491 | 61 | 55 |
| | Cardboard (3) | | 722 | 89 | 80 |
| | Clothes (4) | | 4313 | 533 | 479 |
| | Green Glass (5) | | 509 | 63 | 57 |
| | Metal (6) | | 623 | 77 | 69 |
| | Paper (7) | | 851 | 105 | 94 |
| | Plastic (8) | | 701 | 86 | 78 |
| | Shoes (9) | | 1601 | 198 | 178 |
| | White Glass (10) | | 628 | 77 | 70 |
| | Trash (11) | Residual Waste | 564 | 70 | 63 |
| | **Total** | | **12566** | **1552** | **1397** |

[a] Here, 0–3 and 0–11 indicate class numbers for four-class and twelve-class in the first and second stage classifications, respectively.

## 3. Experimental design and methods

### 3.1. Data overview

The performance of learning models is critically influenced by the quality of the dataset employed. TrashNet, a widely used dataset for garbage classification research, has a limited six-category taxonomy. Jin et al. (2023) used the Huawei Cloud Garbage Classification dataset in their research on garbage identification and classification; however, this dataset is no longer available. The dataset utilized in this study was obtained from Kaggle's Garbage Classification dataset (Mostafa Mohamed, 2021). The database comprises 15150 images representing twelve distinct categories of domestic waste. All the images of the dataset were thoroughly classified into four groups according to characteristics such as origin, composition, and perceived hazard levels. Afterwards, the images were additionally categorized into twelve subclasses. Table 2 and Fig. 1 present a detailed overview of the dataset along with representative sample images.


**Fig. 1.** Garbage Classification dataset includes 4 classes: (1) Hazardous Waste, (2) Household Food Waste, (3) Recyclable Waste, (4) Residual Waste; and 12 sub-classes: (a) Battery, (b) Expired Food, (c) Brown Glass, (d) Cardboard, (e) Clothes, (f) Green Glass, (g) Metal, (h) Paper, (i) Plastic, (j) Shoes, (k) White Glass, (l) Trash (MOSTAFA MOHAMED, 2021). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

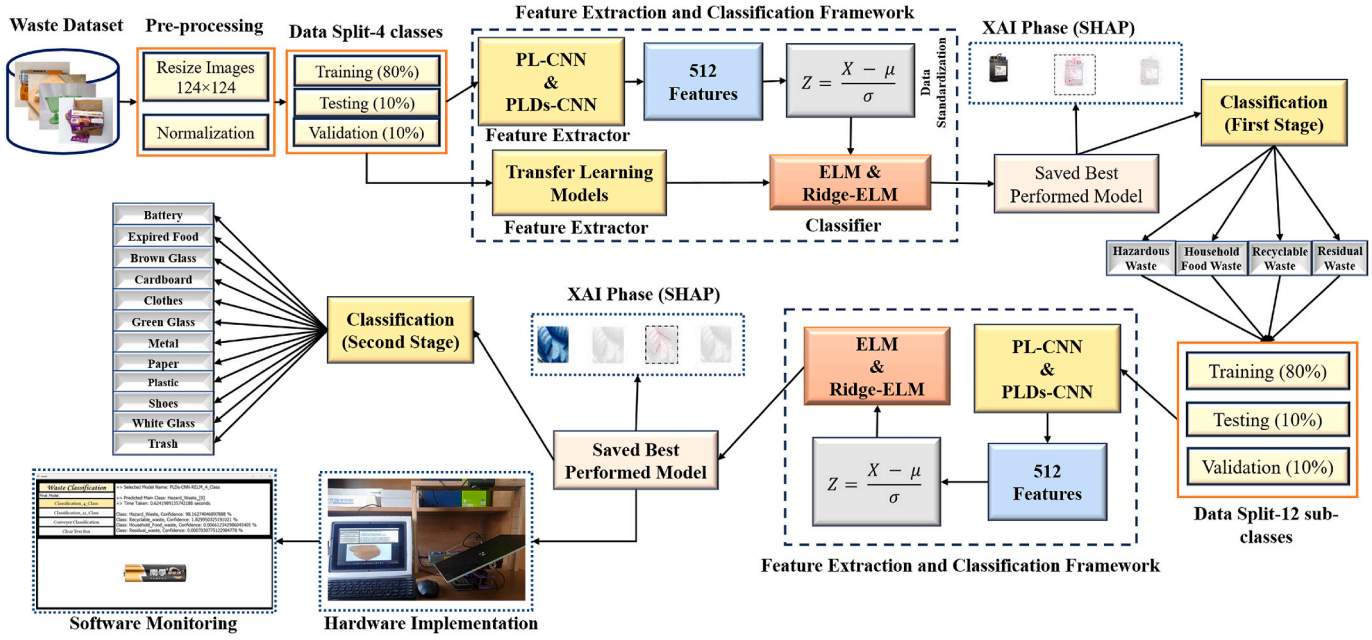


**Fig. 2.** Overview of the proposed two-stage framework for multiclass waste image classification.

### 3.2. Proposed framework

Fig. 2 illustrates the key phases of the deep learning (DL) framework introduced in this study, which is divided into two distinct stages to enhance classification accuracy. Preprocessing is applied to the dataset initially, which is then divided into training, testing, and validation sets comprising 80 %, 10 %, and 10 % images, respectively. In the first stage, two advanced neural network architectures—PLDs-CNN and parallel CNN (PL-CNN)—are employed to extract essential image features. For comparative analysis between the proposed models and leading transfer learning (TL) models, various TL-based feature extractors are also utilized. After standardizing the data, two classifiers, pseudo-Extreme Learning Machine (ELM) and Ridge-ELM, are developed to assess the class identification performance. The SHAP (SHapley Additive exPlanations) approach is used to explain the output generation process of the models. During the final step of the first stage, waste materials are systematically classified into four key categories: hazardous waste, household food waste, recyclable waste, and residual waste.

At the second stage, the classification is performed again, this time considering broader class categories. After splitting the dataset into training, testing, and validation subsets, the same PL-CNN and PLDs-CNN models are employed separately for feature extraction. ELM and

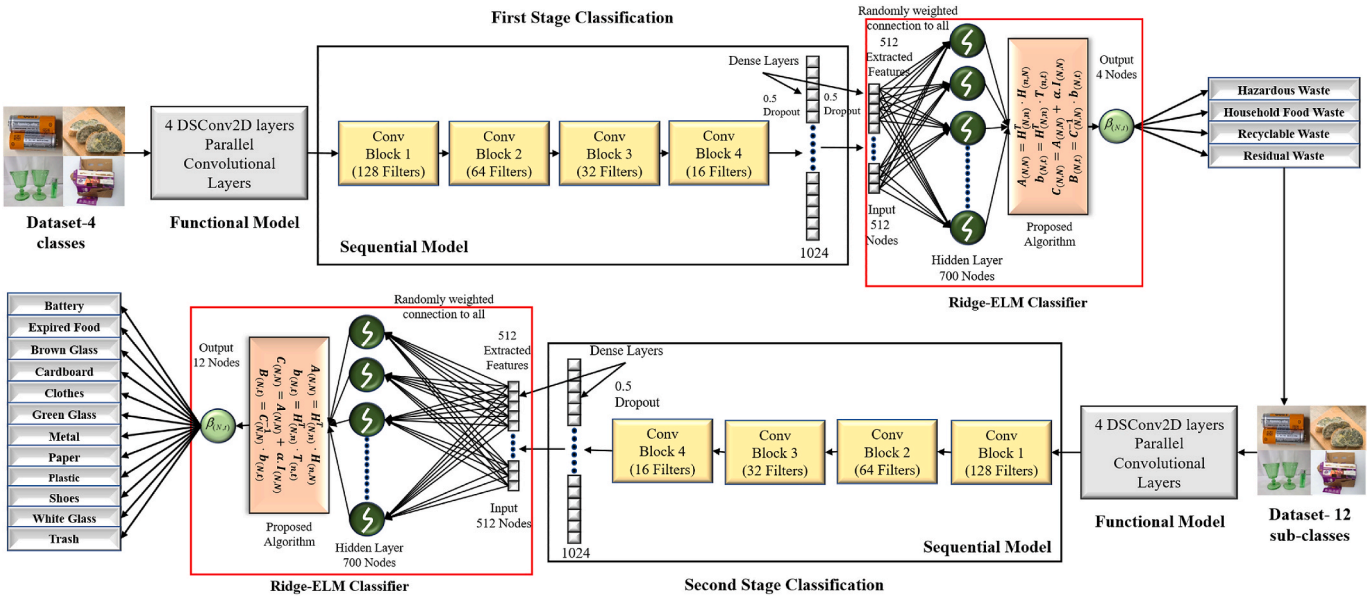


**Fig. 3.** Proposed PLDs-CNN-Ridge-ELM architecture.

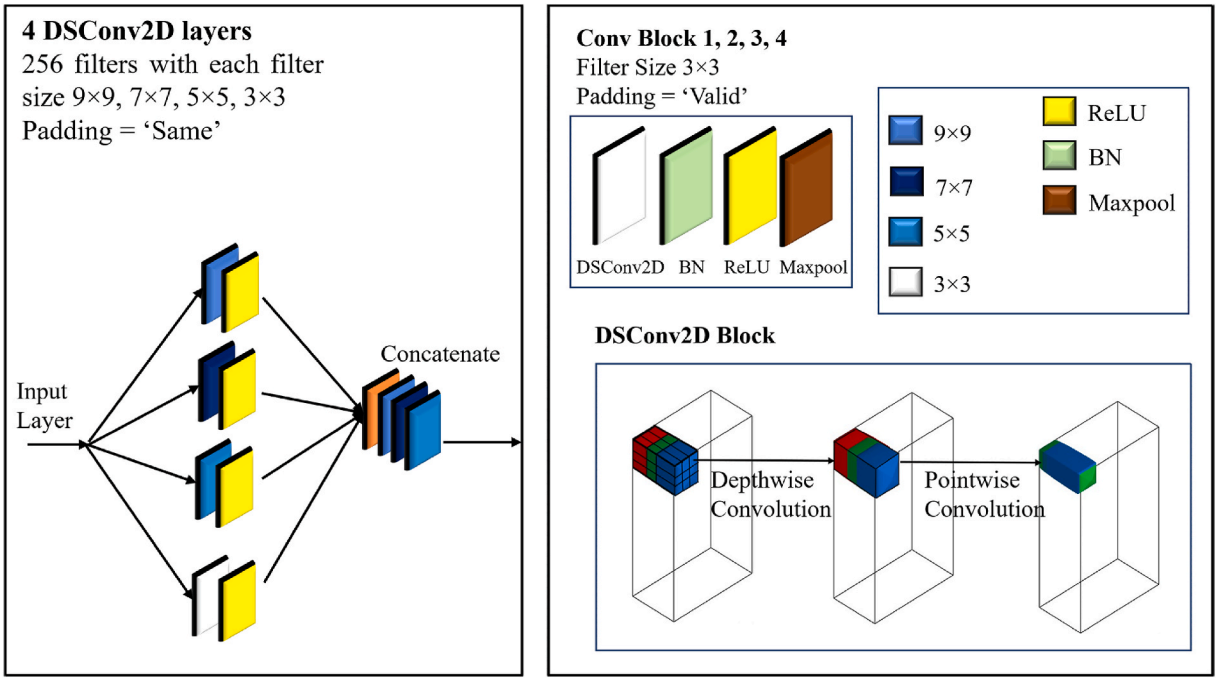


**Fig. 4.** Structural overview of the convolution block.

Ridge-ELM classifiers are then applied to classify the waste materials into twelve more specific subcategories. Once again, SHAP is utilized to enhance the interpretability of the second-stage classification results. Furthermore, a comprehensive hardware structure is designed, including a user-friendly Graphical User Interface (GUI) for rapid waste classification, a webcam-based conveyor belt sorting mechanism, and a 2-axis pan-tilt system for autonomous waste sorting.

The implementation of a two-stage classification model is motivated by the inherent complexities encountered in real-world industrial waste sorting scenarios. The diversity in physical attributes, material constituents, and recyclability among waste materials poses challenges for effective classification and processing. This variability makes it difficult for a single-stage, end-to-end classification model to achieve both high accuracy and robustness across a wide range of categories. By introducing a two-stage approach, it can be more effectively managed this complexity and progressively refine the classification task.

In the first stage, the model broadly classifies waste materials into four key categories: Hazardous Waste, Household Food Waste, Recyclable Waste, and Residual Waste. This initial categorization reduces the complexity of the second stage, making it easier for the model to focus on finer distinctions. The second stage then classifies these broadly categorized waste items into twelve specific subcategories, facilitating more precise and detailed waste categorization. This hierarchical structure reduces misclassification rates while improving overall accuracy by narrowing the focus at each stage.

While a two-stage classification system may intuitively seem to

**Table 3**
Lightweight parallel depthwise separable convolutional neural network (PLDs-CNN) model summary.

| Network Structure by Processing Blocks | | | |
|---|---|---|---|
| Processing Block | Layer Components | Feature Dimensions | Trainable Units |
| **Input Stage** | Input Layer | 124 × 124 × 3 | 0 |
| | Functional Model | 124 × 124 × 1024 | 4588 |
| **Primary Feature Extraction** | Depthwise Separable Convolution | 122 × 122 × 128 | 140,416 |
| | Normalization + Activation | 122 × 122 × 128 | 512 |
| | Feature Pooling | 61 × 61 × 128 | 0 |
| **Secondary Feature Extraction** | Depthwise Separable Convolution | 59 × 59 × 64 | 9408 |
| | Normalization + Activation | 59 × 59 × 64 | 256 |
| | Feature Pooling | 29 × 29 × 64 | 0 |
| **Tertiary Feature Extraction** | Depthwise Separable Convolution | 27 × 27 × 32 | 2656 |
| | Normalization + Activation | 27 × 27 × 32 | 128 |
| | Feature Pooling | 13 × 13 × 32 | 0 |
| **Final Feature Extraction** | Convolutional Layer | 11 × 11 × 16 | 816 |
| | Normalization + Activation | 11 × 11 × 16 | 64 |
| | Feature Pooling | 5 × 5 × 16 | 0 |
| | Regularization (Dropout) | 5 × 5 × 16 | 0 |
| **Dimension Reduction** | Flatten Operation | 400 | 0 |
| **Classification Head** | Dense Neural Network | 1024 | 410,624 |
| | Normalization | 1024 | 4096 |
| | Regularization (Dropout) | 1024 | 0 |
| | Output Dense Layer | 512 | 524,800 |
| **Parameter Distribution** | | | |
| **Network Section** | **Parameter Count** | **Percentage** | |
| Feature Extraction Layers | 158,844 | 14.5 % | |
| Classification Layers | 939,520 | 85.5 % | |
| **Total Parameters** | **1,098,364** | **100 %** | |
| Trainable Parameters | 1,095,836 | 99.8 % | |
| Non-Trainable Parameters | 2528 | 0.2 % | |

increase computational time compared to a single-stage approach, this is not necessarily the case in this paper implementation. Both stages use the same model—PLDs-CNN-Ridge-ELM—without introducing additional resource demands. Therefore, the computational overhead is only marginally increased, primarily due to the division of the task into two levels of classification. This minimal increase in computational time does not significantly impact the efficiency of the system in industrial settings where high throughput and rapid classification are essential. Moreover, the use of lightweight feature extractor such as PLDs-CNN and classifier Ridge-ELM helps ensure that the system remains scalable and efficient even in large-scale operations.

### 3.3. Image preprocessing

The image preprocessing phase involves two essential procedures: normalization (Nahiduzzaman et al., 2023) and image down sampling. The operations aim to resize all input images uniformly to 124 × 124 pixels. These techniques are essential for enhancing model efficiency and extracting features.

### 3.4. Deep learning (DL) model

Current literature predominantly focuses on utilizing large-scale models for waste image classification, while comparatively limited attention has been directed toward the design of lightweight architectures and their applicability in practical, real-time waste sorting environments. In response to these challenges, a new PLDs-CNN feature extractor was developed and compared with state-of-the-art TL models. This customized design incorporates a reduction in model parameters, layers, and total size, requiring less computational resources. The following sections contain an in-depth explanation of the PLDs-CNN, along with brief insights into the TL models. Furthermore, the rationale for PLDs-CNN feature extractor has been presented in the results and discussion.

#### 3.4.1. Feature extraction by PLDs-CNN

The primary challenge when building a CNN model is to determine the optimal layer set up. Limited parameters and layers may restrict the model's ability to capture unique features, imposing restrictions on its performance. On the other hand, an excessive number of parameters and layers might cause overfitting, which leads to longer processing times and higher computational requirements. Therefore, it is essential to achieve the optimal balance to ensure effective feature extraction with successful implementation. The main objective of this study was to create a CNN model that can extract key features with the least number of parameters and layers.

Considering all relevant aspects, a lightweight PLDs-CNN model was designed to efficiently extract distinctive features while minimizing resource consumption. The overall structure of the proposed PLDs-CNN is shown in Fig. 3. To enhance the usability and optimize the configuration, a refined trial-and-error strategy was adopted. After conducting several experiments with different layer arrangements, the final architecture comprises nine convolution layers and two fully connected layers, providing a balanced trade-off between high classification accuracy and reduced computational complexity—measured in terms of the number of parameters, layer depth, model size, and inference time. To enable more effective feature extraction, the architecture incorporates four parallel convolution layers rather than a single sequential one, as illustrated in Fig. 4. Although employing four consecutive convolution layers would typically increase the model's structural complexity, this issue was mitigated by executing them in parallel. The selection and configuration of these layers were determined through iterative empirical evaluation. Collectively, these layers employ 256 convolution kernels, with filter sizes of 9 × 9, 7 × 7, 5 × 5, and 3 × 3, respectively. The use of larger kernel sizes, such as 9 × 9, is supported by prior studies (Krizhevsky et al., 2017; Nahiduzzaman et al., 2023a–c), which indicate their effectiveness in enhancing classification performance. Different kernel sizes generate diverse feature representations; thus, integrating a wide range of filter dimensions was crucial for achieving superior performance. To preserve spatial information, especially from the boundary regions of the input images, a consistent padding size was applied to the first five convolution layers. The resulting feature maps from the parallel paths were then carefully aggregated and forwarded to the next convolution layer to ensure accurate and lossless information flow (Nahiduzzaman et al., 2023; Nahiduzzaman et al., 2024; Nahiduzzaman et al., 2021b).

To enhance the performance of the CNN, depthwise separable convolutions (DSC) were utilized. This technique restructures the convolution process by first applying a spatial (depthwise) filter to each input channel individually, followed by a pointwise convolution to combine the outputs across channels. A compressed kernel is utilized on a specific portion of the DSC to process an infusion feature map, resulting in a new feature map output with the same number of channels. During pointwise convolution, a 1 × 1 convolutional kernel is applied independently to each channel to create a new feature map with fewer channels. This emphasizes the utmost significance of DSC. This fine-tuning of the parameters leads directly to a notable decrease in computational complexity. During the final phase, three convolutional layers (CLs) were incorporated, and N and max pooling with a kernel size of 2 × 2 were employed. The CL filters had 128, 64, 32, and 16 values, respectively. Each filter was set up with three 3 × 3 kernels and designed to use

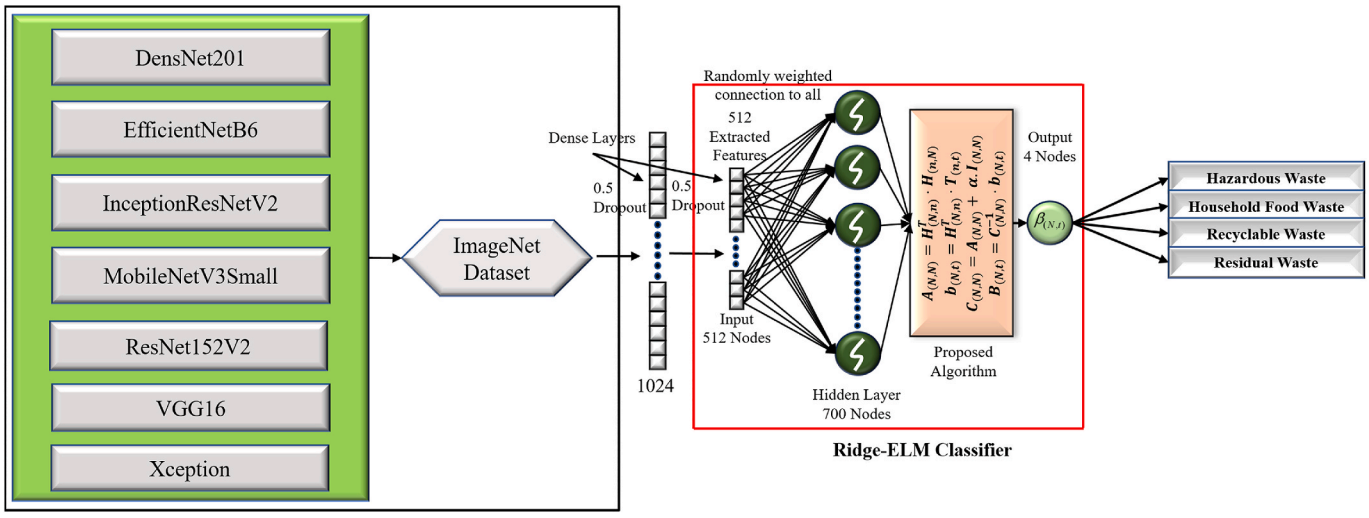


**Fig. 5.** The transfer learning (TL) architectures with Ridge-ELM classifier for classifying waste images.

valid padding. Batch Normalization (BN) is included to enhance the model's effectiveness. This technique efficiently restores the mean and standard deviation of the inputs for each layer, resulting in improved speed and stability during model execution. All convolutional layers (CLs) employed the Rectified Linear Unit (ReLU) activation function. The separation in CL and Parallelization characteristics of the PLDs-CNN model reduce the computational cost and the number of parameters compared to traditional CL in a normal CNN. A reduction in the number of parameters can result in faster training times and lower memory requirements.

To mitigate the risk of overfitting and enhance training efficiency, dropout regularization was applied in conjunction with two fully connected (FC) layers. During each training cycle, 50 % of the neurons were randomly disabled, encouraging generalization and facilitating faster convergence. In the final FC layer, 512 highly informative features were extracted, which contributed to boosting overall classification accuracy. Instead of the conventional SoftMax function, the model employed a Ridge-ELM classifier, aiming to further improve discriminative capability. The training process was guided by a loss function derived from the sparse categorical cross-entropy formulation, ensuring robustness in multi-class scenarios. An ADAM optimizer with a batch size of 32 was selected to update the model parameters. The learning rate, experimentally determined through iterative tuning, was fixed at 0.001, and the training was conducted over 200 epochs. A comprehensive summary of the model's architecture and parameters is presented in Table 3.

### 3.4.2. *Feature extraction by transfer learning models*

Transfer learning (TL) models such as DenseNet201 (Zhao et al., 2021), EfficientNetB6 (Tan and Le, 2019), InceptionResNetV2 (Bhatia et al., 2019), MobileNetV3Small (Maheta and Manisha, 2023), ResNet152V2 (He et al., 2015), VGG16 ((Sudha and Ganeshbabu, 2020), and Xception (Chollet, 2017) have the ability to improve the classification of trash images across many categories. Due to their extensive pretraining on large datasets, these models exhibit high effectiveness in extracting significant features from images. These models can efficiently capture intricate patterns and precise information associated with trash images by being fine-tuned on a limited amount of data. The pretrained models were trained using over 14 million classifications from the ImageNet dataset, spanning approximately 1000 categories. The selection of these specific TL models is grounded in prior literature, where these architectures have been repeatedly demonstrated as state-of-the-art or highly competitive for waste or garbage classification tasks (see Table 1). These models represent a diverse set of architectures, ranging from lightweight to deep networks, providing a comprehensive benchmark. Their consistent strong performance, balanced complexity, and widespread adoption make them suitable and relevant baselines. Furthermore, their extensive pretraining on large-scale datasets such as ImageNet equips them with robust feature extraction capabilities, which is crucial for effective transfer learning in the domain of trash image classification. In this study, the pretrained TL models were retrained by fine-tuning their pretrained weights on the Garbage Classification dataset. To achieve improved classification performance, integration of the proposed Ridge-ELM classifier into the TL model training pipeline was implemented. Subsequently, the models underwent evaluation using the same dataset to determine their predictive capabilities. On the other hand, the novel PLDs-CNN-Ridge-ELM model was entirely developed and trained from the ground up, without relying on any pretrained weights, and was subjected to validation and performance testing using the garbage classification dataset. A detailed comparison between the proposed approach and the TL-based models was carried out, emphasizing classification accuracy and efficiency in computational resource usage. This evaluation considered multiple factors, including predictive performance, number of model parameters, architectural depth, and the durations required for both training and testing phases. Once these models were initialized, their final layers were adjusted by adding two fully connected (FC) layers with 1024 and 512 neurons, respectively. Fig. 5 provides a detailed illustration of the transfer learning architecture combined with the Ridge-ELM classifier.

Several state-of-the-art CNN architectures have demonstrated effectiveness across diverse computer vision tasks. DenseNet (Zhao et al., 2021) introduces dense connectivity, where each layer receives input from all preceding layers, promoting feature reuse and efficient gradient flow; its variants, such as DenseNet-121, -169, and −201, vary in depth. EfficientNetB6 (Tan and Le, 2019), part of the EfficientNet family, employs compound scaling to uniformly balance network depth, width, and resolution, achieving strong performance with approximately 87 million parameters. InceptionResNetV2 (Bhatia et al., 2019) integrates inception modules and residual connections to enable efficient and robust feature extraction. In contrast, MobileNetV3Small (Maheta and Manisha, 2023) is optimized for resource-constrained environments, offering a favorable trade-off between model size, speed, and accuracy. The VGG network (Sudha and Ganeshbabu, 2020) follows a straightforward architecture comprising stacked convolutional layers with ReLU activation and max pooling, culminating in fully connected layers and a SoftMax classifier. ResNet152V2 (He et al., 2015) leverages residual learning through shortcut connections to facilitate deep network training and consists of roughly 60 million parameters. Lastly, Xception (Chollet, 2017) enhances the Inception architecture by adopting

depthwise separable convolutions, significantly reducing computational cost while maintaining high accuracy.

### 3.5. Ridge Extreme Learning Machine (Ridge-ELM) classifier

The Ridge Extreme Learning Machine (ELM) is utilized to categorize waste materials into various classes, collecting features from the dense layer of the PLDs-CNN model. The ELM, developed by Huang et al. (Huang et al., 2017), represents a significant change in feature classification methodology. The employed approach is a feed forward network based on supervised learning, which is a pioneering innovation. By applying the strength of neural networks (NN), the ELM eliminates the necessity of backpropagation, resulting in a remarkable thousand-fold improvement in training speed. This innovative approach has completely transformed the field of feature classification (Nahiduzzaman et al., 2023d).

Recent advancements have provided models with remarkable abilities in classification and generalization. More precisely, the pseudo-ELM has demonstrated exceptional competence in handling large-scale multiclass classification tasks and has outperformed the most recent machine learning (ML) models (Kibria et al., 2022; Maheta and Manisha, 2023; Nahiduzzaman et al., 2021, 2023; Nahiduzzaman et al., 2023; Nahiduzzaman et al., 2021). The Extreme Learning Machine (ELM) is notable for its efficient and flexible parameter initialization scheme, which involves a single hidden layer between the input and output. In traditional ELMs, the output weights are calculated using a pseudoinverse operation. In our proposed approach, this is enhanced by replacing the pseudoinverse with ridge regression, enabling improved regularization and learning capability. This adaptation significantly strengthens the model's ability to extract and generalize meaningful features, thereby contributing to higher classification accuracy. The Ridge-ELM architecture in our framework includes 512 neurons in the input layer, followed by a hidden layer with 700 neurons, forming a robust structure for feature transformation. The output layer comprises four neurons, corresponding to the four primary waste categories. An illustration of the integrated PLDs-CNN and Ridge-ELM framework is provided in Fig. 3, and the detailed steps of the Ridge-ELM algorithm are outlined in **Algorithm 1**.

**Algorithm 1:** Ridge-ELM Multiclass Classification Procedure

1. The input data matrix is represented as $I_{(x,y)}$, while the corresponding target output matrix is denoted as $O_{(x,t)}$. The hidden layer output is expressed as $H_{(x,X)}$, with the input weight matrix defined as $W_{(y,X)}$ and the bias vector as $Bm_{(1,X)}$.
2. The next step is to find the output $H_{(x,N)}$ of the hidden layer.
$H_{(x,X)} = G\left(I_{(x,y)} \cdot W_{(y,X)} + Bm_{(1,X)}\right)$
Here, $G$ is an activation function.
3. Calculate the output weight matrix $\beta_{(X,t)}$ through the use of pseudo inverse method.
$\beta_{(X,t)} = H^{\dagger}_{(X,x)} \times T_{(x,t)}$
4. In this proposed hybrid Ridge regression, the pseudoinverse has been replaced by these equations:
$A_{(X,X)} = H^{T}_{(X,x)} \cdot H_{(x,X)}\ b_{(X,t)} = H^{T}_{(X,x)} \cdot T_{(x,t)}$
$C_{(X,X)} = A_{(X,X)} + \alpha . I_{(X,X)}$
$B_{(X,t)} = C^{-1}_{(X,X)} \cdot b_{(x,t)}$
Where, α denotes regularization parameters.
5. Generate prediction $B_{(X,t)}$

The model displayed a strong sense of assurance in its ability to produce precise and accurate final predictions. Ridge-ELM, a method that seamlessly integrates ridge regression into the ELM framework, achieves a perfect balance between effective feature learning and regularization. Consequently, the predictive capability of the model was enhanced by its greater ability to generalize and interpret intricate patterns in the data.

### 3.6. Explainable Artificial Intelligence (XAI)

Explainable Artificial Intelligence (XAI) is important for enhancing the transparency and interpretability of the PLDs-CNN model. Shapley Additive Explanation (SHAP) was employed to address the "black box" nature of DL models, which usually makes them less understandable. Through the integration of the PLDs-CNN model and SHAP, automatic garbage classification systems are now capable of making smarter and more efficient decisions when classifying waste items into four main classes and twelve different subclasses (Lundberg and Lee, 2017). This technique presents new opportunities for smarter waste management and more efficient categorization of waste materials across multiple categories.

In this study, SHAP was specifically chosen over other XAI methods like Grad-CAM due to the hybrid nature of the proposed architecture, which combines CNN feature extraction with a Ridge Extreme Learning Machine (Ridge-ELM) classifier. Grad-CAM relies on gradient information and is best suited for pure CNN models; however, because Ridge-ELM is a non-gradient-based classifier, Grad-CAM cannot provide meaningful explanations for the full model. In contrast, SHAP is model-agnostic and calculates Shapley values based on feature contributions, allowing it to explain both the CNN's learned representations and the subsequent Ridge-ELM decision process. This makes SHAP more suitable for providing a unified interpretability framework for our hybrid model. Furthermore, recent studies have validated SHAP's effectiveness in explaining complex models in similar classification contexts, supporting our choice. Additionally, SHAP is a relatively recent and advanced XAI method whose application in waste classification and related fields is still emerging.

There was a clear pattern in the Shapley values that were employed in the study to measure the significance of individual pixels. The presence of red pixels enhances the accuracy of class identification, while the presence of blue pixels diminishes the probability of proper categorization (Bhandari et al., 2022). The Shapley values were computed using Equation (1).

$$\varnothing_r = \sum_{V \subseteq N \setminus r} \frac{|V|!(C - |V| - 1)!}{C!} [f_x(V \cup r) - f_x(V)] \tag{1}$$

$$f_x\,(V) = P[f(x)|x_V] \tag{2}$$

$$l\,(b') = \varnothing_0 + \sum_{r=1}^{C} \varnothing_r b'_r \tag{3}$$

The variable $f_x$ quantifies the influence of a specific feature $r$ on the model's output, as interpreted through its corresponding Shapley value. The subset $V$ consists of all features in the set $N$, excluding the feature $r$. The term $\frac{|V|!(C-|V|-1)!}{C!}$ represents the weighting coefficient associated with the number of possible permutations involving subset $V$. Equation (2) expresses the model's prediction for a given subset of features as $f_x\,(V)$. In the SHAP framework, each original feature $x_r$ is substituted with a binary indicator $b'_r$, which indicates the inclusion (1) or exclusion (0) of feature $x_r$, as demonstrated in Equation (3). Within the proposed model $f(x)$, the bias component is denoted as $\varnothing_0$, and the specific contribution of feature $r$ is represented by $\varnothing_r b'_r$. The function $l\,(b')$ serves as a simplified surrogate model that approximates the behavior of the original predictive function. The term $\varnothing_r$ provides insight into the degree to which feature $r$ contributes to the overall prediction, thereby enhancing interpretability and facilitating model transparency.

### 3.7. Hyperparameter and architectural selection

The hyperparameters and architectural choices of the proposed PLDs-CNN-Ridge-ELM model were determined through iterative experimentation and performance-based evaluation. Rather than relying on automated optimization techniques, a trial-and-error approach was

**Table 4**
Summary of selected hyperparameters and architectural choices used in the proposed framework.

| Parameter Name | Combination Applied | Selected | Justification |
|---|---|---|---|
| Feature Extractor | PL-CNN, TL (VGG16, ResNet50, MobileNet), PLDs-CNN | **PLDs-CNN** | Outperformed other architectures in accuracy and generalization with reduced overfitting. |
| Classifier | Softmax, ELM, Ridge-ELM | Ridge-ELM | Outperformed softmax and traditional ELM in both speed and classification accuracy. |
| Image Size | (128 × 128), (124 × 124), (224 × 224) | 124 × 124 | Chosen for standardization and to reduce computational cost while preserving image details. |
| Data Split | 80/10/10, 70/15/15 | 80 % train, 10 % test, 10 % val | Ensures sufficient data for training while reserving samples for robust testing and validation. |
| Convolutional Layers | 5–12 | 9 | Empirically determined through iterative testing to balance performance and model complexity. |
| Kernel Sizes | 3 × 3, 5 × 5, 7 × 7, 9 × 9 | 3 × 3, 5 × 5, 7 × 7, 9 × 9 | Multi-scale kernel design improves feature extraction across various spatial resolutions. |
| Activation Function | ReLU, Leaky ReLU | ReLU | Chosen for non-linearity and computational efficiency in all convolutional layers. |
| Pooling Type | Max Pooling, Avg Pooling | Max Pooling (2 × 2) | Reduces spatial dimensions while retaining dominant features; empirically more effective. |
| Batch Normalization | Yes/No | Yes | Stabilizes training and accelerates convergence. |
| Dropout Rate | 0.3, 0.5, 0.6 | 0.5 | Helps prevent overfitting and encourages generalization. |
| Optimizer | Adam, SGD, RMSprop | Adam | Adaptive learning rate control; proven effective during empirical testing. |
| Learning Rate | 0.001, 0.0005, 0.0001 | 0.001 | Empirically derived through trial and error for optimal convergence. |
| Loss Function | Categorical Cross Entropy, Sparse Categorical Cross Entropy | Sparse Categorical Cross Entropy | Suitable for multi-class classification with integer labels. |
| Batch Size | 16, 32, 64 | 32 | Balanced performance and memory efficiency. |
| Epochs | 100, 200, 300 | 200 | Achieved best convergence without overfitting. |
| Hidden Layer Nodes | 512, 700, 1024 | 700 | Best trade-off between complexity and performance for Ridge-ELM. |
| Explainability Tool | SHAP | SHAP | Provides robust and model-agnostic interpretation of classification decisions. |

adopted to achieve a balance between model accuracy and computational efficiency, particularly in real-time waste classification contexts. The input image size was fixed at 124 × 124 pixels after down-sampling, which preserved essential spatial features while reducing memory overhead. The training/validation/testing split of 80/10/10 was chosen to maximize data available for training without compromising evaluation reliability. For convolutional layers, a parallel architecture incorporating kernel sizes of 9 × 9, 7 × 7, 5 × 5, and 3 × 3 was used to capture features at multiple receptive fields, a strategy shown to improve classification performance in prior literature.

To enhance computational efficiency, depthwise separable convolutions (DSC) were integrated into the later stages of the network architecture. This approach significantly reduced the total number of trainable parameters without compromising model accuracy. Each convolutional layer employed the ReLU activation function, followed by batch normalization to improve training stability and convergence speed. To prevent overfitting, a dropout rate of 0.5 was applied, randomly disabling half of the neurons during training iterations. The model was trained over 200 epochs using the Adam optimizer with a learning rate of 0.001, which was empirically found to ensure consistent convergence. A batch size of 32 was selected to strike an optimal balance between computational resource usage and gradient estimation stability. For handling the multi-class classification task, the sparse categorical cross-entropy loss function was utilized due to its compatibility with integer-labeled targets. Instead of the conventional SoftMax classifier, the Ridge-ELM was employed at the classification stage, offering faster inference and improved generalization. The Ridge-ELM model, composed of 512 input nodes, 700 hidden neurons, and 4 output nodes, demonstrated superior classification performance due to its efficient parameter initialization and regularization capabilities. Overall, the selected hyperparameters and architectural components were empirically justified to ensure a lightweight yet effective classification model suitable for real-time waste sorting applications. Table 4 summarizes the chosen parameters, and their corresponding values used in the proposed model.

### 3.8. Classification experiments and performance matrices

The deep learning models and explainable AI techniques were implemented using the Keras framework, supported by the TensorFlow backend, within the PyCharm IDE (Community Edition, v2021.2.3). Model training and evaluation were conducted on a workstation powered by an 11th generation Intel® Core™ i9-11900 CPU running at 2.50 GHz, supported by 128 GB of RAM and an NVIDIA GeForce RTX 3090 GPU featuring 24 GB of dedicated memory. The system operated on a 64-bit Windows 10 Pro platform.

A confusion matrix (CM) was used to measure the performance of the PLDs-CNN-Ridge-ELM model. The accuracy, precision, recall, F1-score, and area under the curve (AUC) from the CM were determined using the following formulas. In the published literature related to waste classification, these parameters were also employed as performance indicators.

$$Accuracy = \frac{True\ Positive + True\ Negative}{True\ Positive + True\ Negative + False\ Positive + False\ Negative} \tag{4}$$

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \tag{5}$$

$$Recall = \frac{True\ Positive}{True\ Negative + False\ Positive} \tag{6}$$

$$F1 - Score = \frac{2 \times (Precision \times Recall)}{Precision + Recall} \tag{7}$$

$$AUC = \frac{1}{2}\left(\frac{True\ Positive}{True\ Positive + False\ Negative} + \frac{True\ Negative}{True\ Negative + False\ Positive}\right) \tag{8}$$

## 4. Results and discussion

The study organizes its outcomes into two stages. Initially, the dataset is divided into four classes (first stage), and the findings of this

**Table 5**
Four-class performances by using PL-CNN-ELM and PL-CNN-Ridge-ELM architectures.

| Class Name | PL-CNN-ELM | | | PL-CNN-Ridge-ELM | | |
|---|---|---|---|---|---|---|
| | Precision | Recall | F1 | Precision | Recall | F1 |
| Hazardous Waste (0) | 0.99 | 0.86 | 0.92 | 0.99 | 0.94 | 0.96 |
| Household Food Waste (1) | 0.98 | 0.89 | 0.93 | 0.98 | 0.93 | 0.96 |
| Recyclable Waste (2) | 0.98 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 |
| Residual Waste (3) | 0.92 | 0.88 | 0.90 | 0.94 | 0.92 | 0.93 |
| **Average (μ) ± SD (σ) (%)** | 96.75 ± 0.03 | 90.5 ± 0.05 | 93.5 ± 0.03 | **97.5 ± 0.02** | **94.5 ± 0.03** | **96 ± 0.02** |
| **Accuracy (%)** | 97.0 | | | **98.0** | | |
| **AUC (%)** | **99.29** | | | 99.14 | | |

Note: The best results are highlighted in bold. 0–3 indicates the class number.

stage are presented in Section 4.1. This study analyzed the effectiveness of the PL-CNN (Section 4.1.1) and PLDs-CNN (Section 4.1.2) feature extractors across four classes and compared their results with those of other state-of-the-art TL-based feature extractors (Sections 4.1.3 and 4.1.4). Afterwards, the dataset is divided into twelve subclasses, and the results of this stage are outlined in Section 4.2. Twelve-class classification (second stage) was conducted using the PL-CNN (Section 4.2.1) and PLDs-CNN (Section 4.2.2), as these feature extractors demonstrated superior performance in four-class classification compared to other models. The interpretability of SHAP is demonstrated in Section 4.3, and the outcomes of hardware implementation are presented in Section 4.4 in detail.

### 4.1. First-stage classification: analysis of four-class performance

#### 4.1.1. PL-CNN-ELM and PL-CNN-ridge-ELM

Initially, the PL-CNN feature extractor was employed to conduct training using a dataset including 15150 images that encompassed 4 distinct waste categories. An independent assessment of the PL-CNN (excluding DSC) was conducted on a dataset comprising 1552 test images. Both ELM and Ridge-ELM were utilized for evaluating the class-specific performance of the PL-CNN. The findings of these evaluations are displayed in Table 5. The PL-CNN-ELM achieved an average test precision of 96.75 ± 0.03 %, a recall of 90.5 ± 0.05 %, and an F1-score of 93.5 ± 0.03 %. The accuracy and area under the curve (AUC) were 97 % and 99.29 %, respectively. The Ridge-ELM classifier on this model achieved an average precision of 97.5 ± 0.02 % (an improvement of 0.75 %), a recall of 94.5 ± 0.03 % (an improvement of 4 %), and a f1-score of 96 ± 0.02 % (an improvement of 2.5 %). The average accuracy and AUC were 98 % (with a 1 % improvement) and 99.14 %, respectively. The concordance index (C-index) was found to be 0.93 for both classifiers, suggesting similar general discriminative abilities.

Although Ridge-ELM consistently outperformed PL-CNN-ELM across most evaluation metrics, the AUC of PL-CNN-ELM was marginally higher (99.29 % vs. 99.14 %). This subtle difference can be attributed to minor variations in the decision boundary formation between the two models. AUC reflects the model's ability to distinguish between classes across all possible threshold values, and a slightly higher AUC for PL-CNN-ELM suggests that it maintained slightly better discrimination across varying decision thresholds, even though Ridge-ELM performed better at fixed operating points typically selected for classification tasks (e.g., thresholds based on maximum F1-score or balanced accuracy). However, the difference in AUC values is extremely small (0.15 %) and statistically negligible, indicating that both classifiers exhibit excellent and comparable discriminatory performance. These findings demonstrate that while Ridge-ELM offers an advantage in practical classification metrics, both models are highly competitive in terms of overall classification reliability.

The ROC and precision-recall (PR) curves for each class are shown in Figs. 6 and 7, respectively, further illustrating the robustness of the classification performance.

#### 4.1.2. PLDs-CNN-ELM and PLDs-CNN-ridge-ELM (proposed method)

Table 6 highlights the key performance differences between the ELM and Ridge-ELM classifiers when integrated with the proposed PLDs-CNN feature extraction method. To ensure robustness and reliability, evaluations were conducted using a benchmark dataset comprising images from twelve distinct waste material categories. Fig. 8 illustrates the confusion matrices (CMs) for both classifiers in the initial classification phase, involving four categories, offering valuable insights into how each model differentiates among the classes. The Ridge-ELM model notably reduced the rate of incorrect classifications, especially in the cases of Hazardous Waste (class 0), Household Food Waste (class 1), and Residual Waste (class 3). Precision, recall, and F1-scores were computed

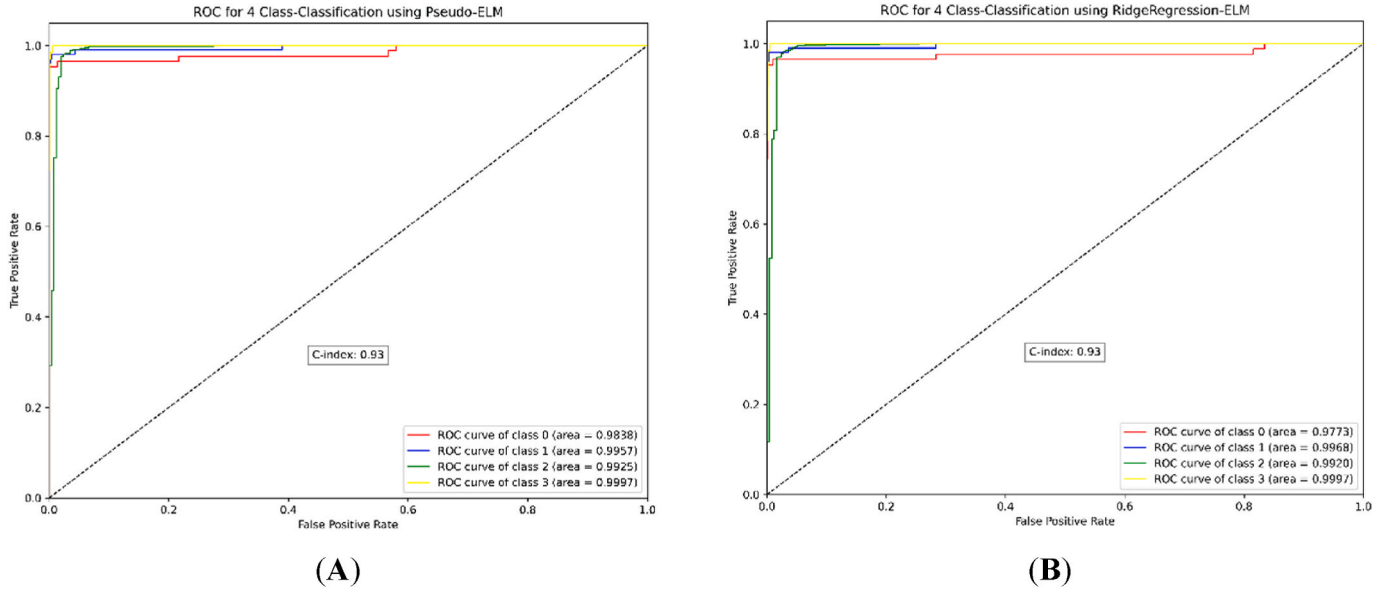


**Fig. 6.** Class-specific ROC Plots for (A) PL-CNN-ELM and (E) PL-CNN-Ridge-ELM models in four-class classifications.

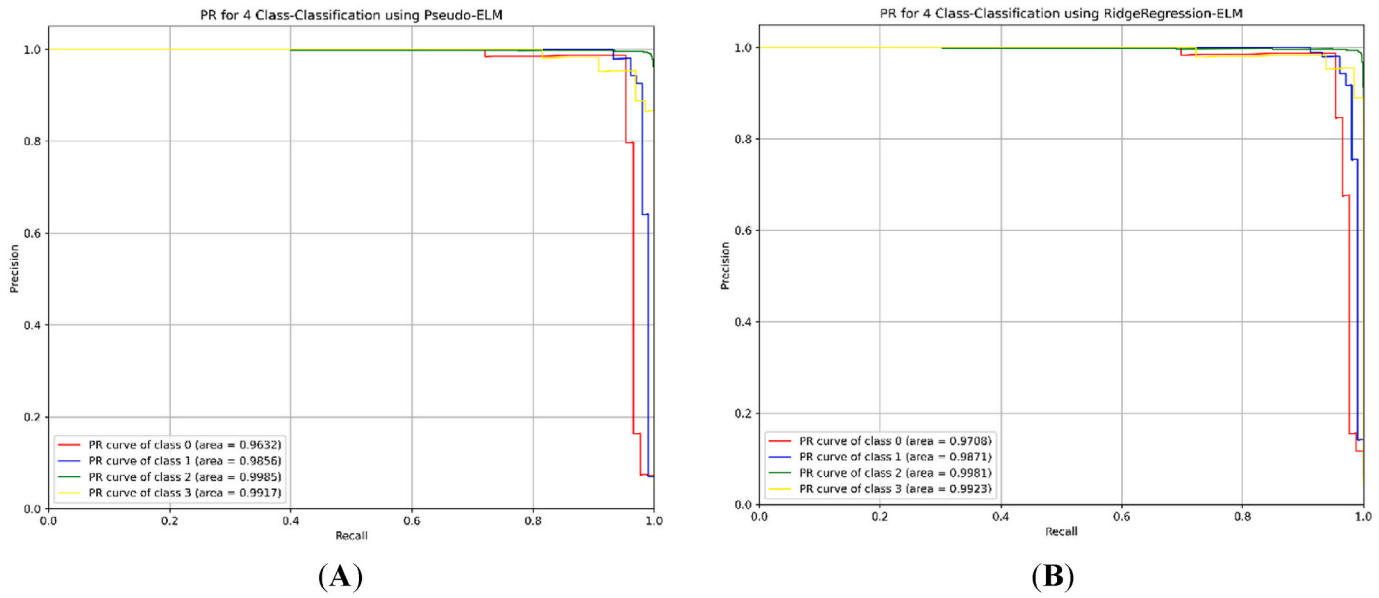


**Fig. 7.** Class-specific PR curves of (A) PL-CNN-ELM and (E) PL-CNN-Ridge-ELM for four-class classification.

**Table 6**
Four-class performances by using PLDs-CNN-ELM and PLDs-CNN-Ridge-ELM architectures.

| Class Name | PLDs-CNN-ELM | | | PLDs-CNN-Ridge-ELM | | |
|---|---|---|---|---|---|---|
| | Precision | Recall | F1 | Precision | Recall | F1 |
| Hazardous Waste (0) | 0.99 | 0.86 | 0.92 | 0.98 | 0.97 | 0.97 |
| Household Food Waste (1) | 0.97 | 0.85 | 0.91 | 0.99 | 0.91 | 0.95 |
| Recyclable Waste (2) | 0.97 | 1.00 | 0.98 | 0.99 | 0.99 | 0.99 |
| Residual Waste (3) | 0.98 | 0.86 | 0.92 | 0.93 | 0.97 | 0.95 |
| **Average (μ) ± SD (σ) (%)** | **97.75 ± 0.009** | 89.25 ± 0.07 | 93.25 ± 0.03 | 97.25 ± 0.02 | **96 ± 0.03** | **96.5 ± 0.01** |
| **Accuracy (%)** | 97.0 | | | **99.0** | | |
| **AUC (%)** | **99.45** | | | 99.28 | | |

Note: The best results are highlighted in bold. 0–3 indicates the class number.

for each category. Ridge-ELM attained an average precision of 97.25 ± 0.02 %, a recall value of 96 ± 0.03 %, and an F1-score of 96.5 ± 0.01 %, reflecting gains of 6.75 % and 3.25 % in recall and F1-score respectively over the baseline ELM. The overall accuracy also rose from 97 % to 99 %. These findings demonstrate that Ridge-ELM is particularly effective at distinguishing between the various waste types. While the Ridge-ELM's area under the ROC curve (AUC) was recorded at 99.28 %, the ELM's performance in this metric remained close at 99.45 %. As for the concordance index (C-index), Ridge-ELM achieved a value of 0.95, marginally higher than the 0.94 score of the ELM. Collectively, these results provide compelling evidence of the Ridge-ELM classifier's efficiency. The enhanced performance and precision affirm the Ridge-ELM's potential as a reliable and high-performing alternative to conventional classification approaches.

Fig. 9 provides a thorough evaluation of the performances of the PLDs-CNN-ELM and PLDs-CNN-Ridge-ELM models in accurately identifying four different waste categories. The ROC analysis for each model confirmed that both classifiers delivered excellent results, with ROC values exceeding 99 % across all classes. The average ROC score of 99.28 % achieved by the proposed Ridge-ELM integrated with the PLDs-CNN framework highlights the model's remarkable precision and robustness in waste image classification. The findings reveal that the PLDs-CNN-Ridge-ELM model maintains strong dependability and classification capability. Notably, the proposed model reached 99.99 % accuracy in identifying Residual Waste (3). The study confirmed the method's effectiveness in reliably recognizing this waste category. Furthermore, the ROC curve indicated a noticeable enhancement in the detection of Recyclable Waste (2). The detection rate improved from

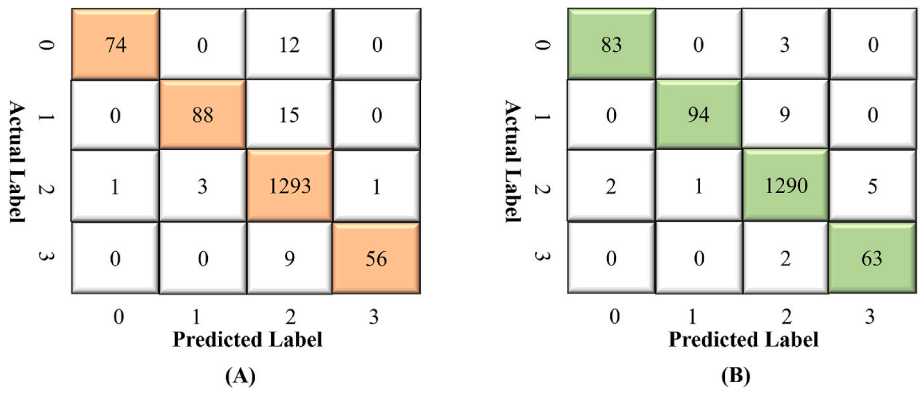


**Fig. 8.** Confusion matrices of PLDs-CNN model with (A) ELM and (B) Ridge-ELM for four-class classification.

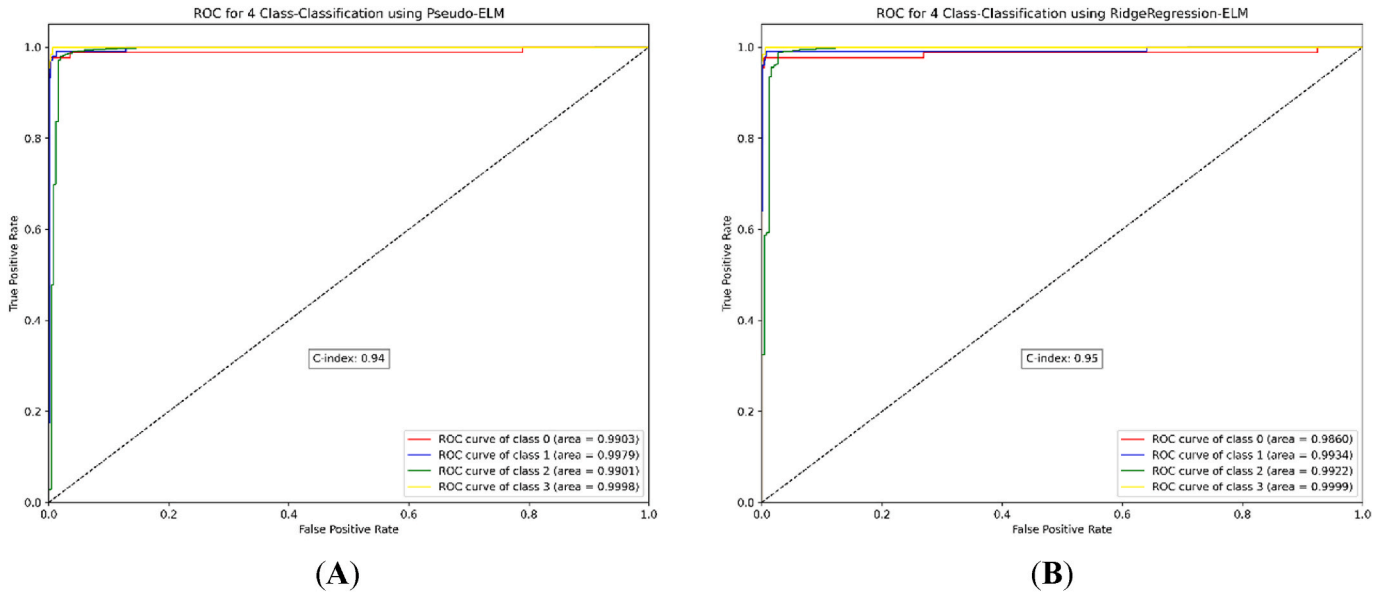


**Fig. 9.** Class-specific ROC Plots for (A) PLDs-CNN-PELM and (E) PLDs-CNN-RELM in four-class classification.

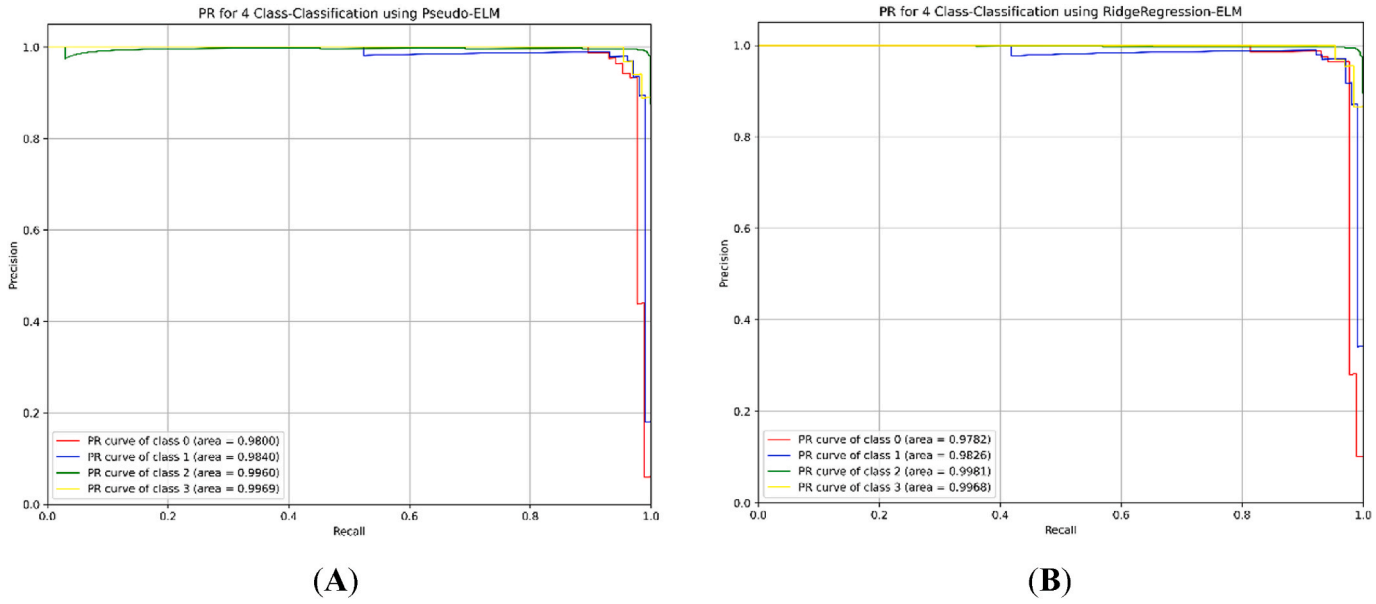


**Fig. 10.** Classwise PR curves of (A) PLDs-CNN-PELM and (E) PLDs-CNN-RELM for four-class classification.

**Table 7**
Four-class classification performance of the TL models and the proposed model.

| Model Name | Average Precision ±SD (%) | Average Recall ±SD (%) | Average F1-score ±SD (%) | Average Accuracy (%) | AUC (%) |
|---|---|---|---|---|---|
| DenseNet201-Ridge-ELM | **98 ± 0.018** | 85.25 ± 0.119 | 90.5 ± 0.066 | 97 | 96.85 |
| EfficientNetB6-Ridge-ELM | 35 ± 0.42 | 26.5 ± 0.49 | 25.5 ± 0.43 | 84 | 74.93 |
| InceptionResNetV2-Ridge-ELM | 94.75 ± 0.02 | 86.25 ± 0.08 | 90.25 ± 0.05 | 96 | 97.88 |
| MobileNetV3Small-Ridge-ELM | 80 ± 0.11 | 61.25 ± 0.26 | 68.25 ± 0.19 | 89 | 91.06 |
| ResNet152V2-Ridge-ELM | 94.75 ± 0.03 | 81.5 ± 0.13 | 87 ± 0.07 | 95 | 97.19 |
| VGG16-Ridge-ELM | 95 ± 0.01 | 90 ± 0.08 | 92.25 ± 0.04 | 97 | 97.47 |
| Xception-Ridge-ELM | 96 ± 0.02 | 82.75 ± 0.11 | 88.5 ± 0.06 | 96 | 96.72 |
| PL-CNN-ELM | 96.75 ± 0.03 | 90.5 ± 0.05 | 93.5 ± 0.03 | 97 | 99.29 |
| PLDs-CNN-ELM | 97.75 ± 0.009 | 89.25 ± 0.07 | 93.25 ± 0.03 | 97 | **99.45** |
| PL-CNN-Ridge-ELM | 97.50 ± 0.02 | 94.5 ± 0.03 | 96 ± 0.02 | 98 | 99.14 |
| PLDs-CNN-Ridge-ELM | 97.25 ± 0.02 | **96 ± 0.03** | **96.5 ± 0.01** | **99.00** | 99.28 |

Note: The best results are highlighted in bold.

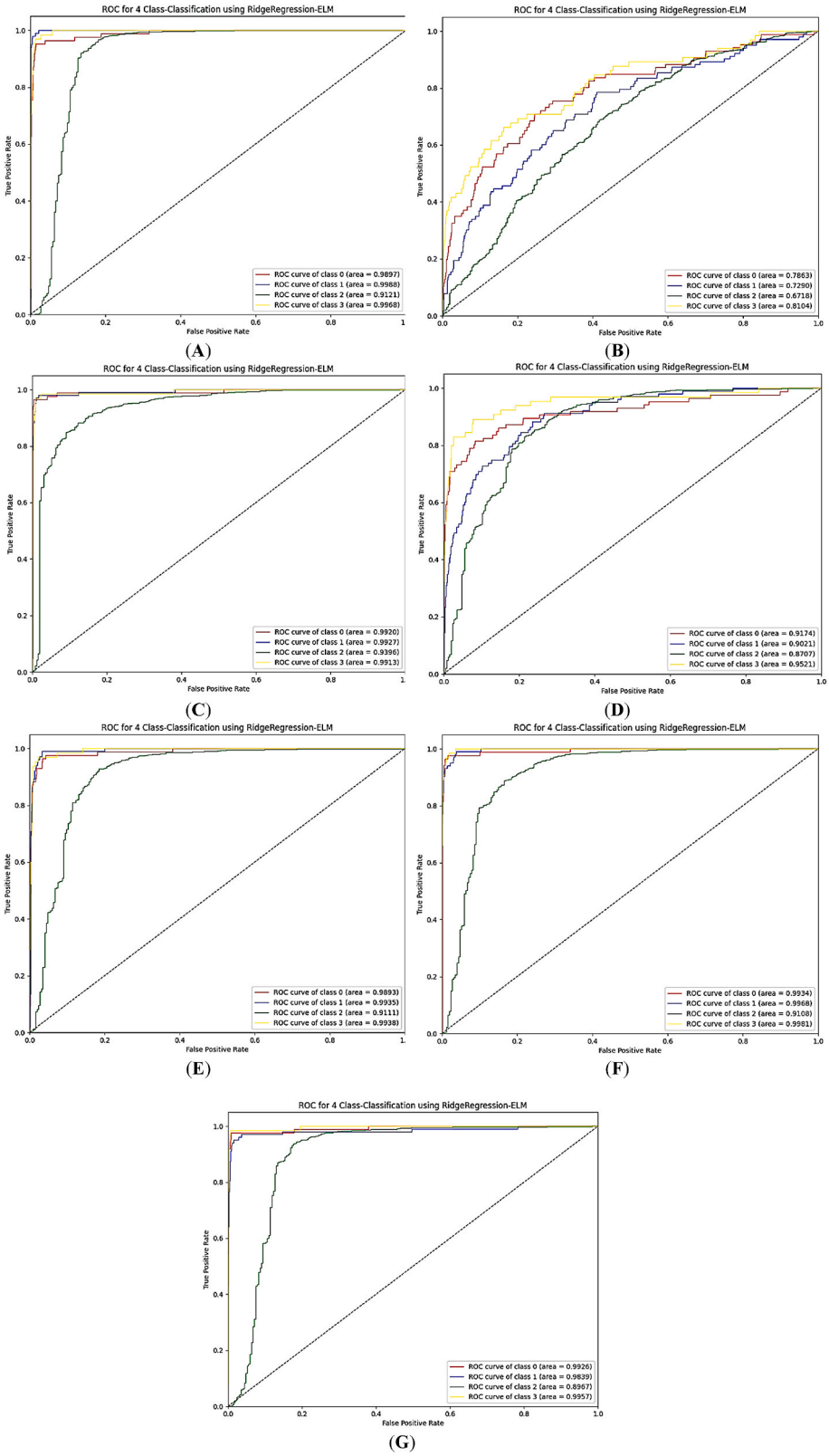


**Fig. 11.** Class-specific ROC Plots for (A) DensNet201, (B) EfficientNetB6, (C) InceptionResNetV2, (D) MobileNetV3Small, (E) ResNet152V2, (F) VGG16, and (G) Xception with Ridge-ELM in four-class classification.

99.01 % to 99.22 %. The proposed PLDs-CNN-Ridge-ELM method outperformed all compared techniques in terms of class-specific identification. There were not enough images for each group in the Garbage Classification dataset, so the dataset was not balanced. One of the primary goals of model development was to efficiently address this data imbalance. An adequate weighting system was implemented to guarantee that each class had an equal influence on the final result. The AUC values ranged from 0.9922 to 0.9999 for PLDs-CNN-Ridge-ELM in all

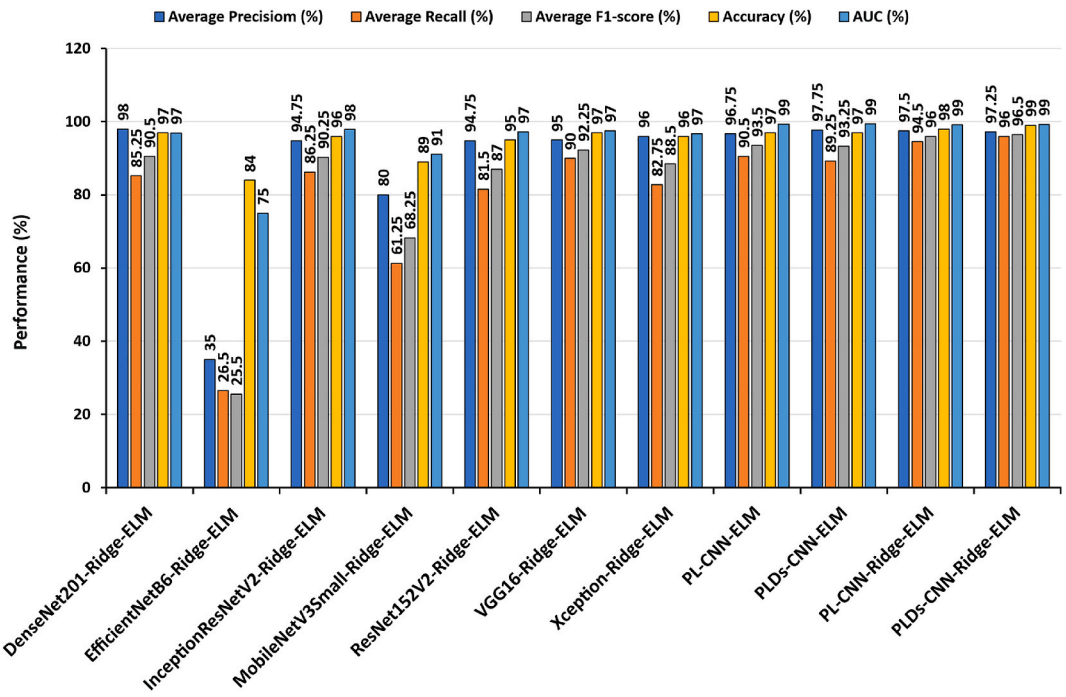


**Fig. 12.** Performance of the proposed and TL models for four class classifications.

classes, demonstrating that superior AUC outcomes were achieved even when dealing with imbalanced datasets. The model's recall rate for Recyclable Waste (2) was 99 %, demonstrating that the suggested framework can consistently differentiate every type of waste material. The proposed framework has the capacity to greatly diminish the hurdles faced by municipal authorities and improve waste management systems through our creative contributions. Fig. 10 displays classwise PR curves for both classifiers.

The proposed PLDs-CNN-Ridge-ELM outperforms PLDs-CNN-ELM across almost all evaluation metrics, confirming its superiority for practical deployment. PLDs-CNN-Ridge-ELM achieved comparable average precision (97.25 % vs. 97.75 %), recall (96 % vs. 89.25 %), F1-score (96.5 % vs. 93.25 %), and accuracy (99 % vs. 97 %) compared to PLDs-CNN-ELM. While the AUC values are very close (99.28 % for Ridge-ELM and 99.45 % for ELM), practical classification performance (precision, recall, F1-score, and accuracy) is more critical in real-world waste sorting systems, where correct class assignments at specific thresholds are prioritized. Additionally, PLDs-CNN-Ridge-ELM demonstrated better robustness in handling imbalanced datasets, as evidenced by its improved recall for minority classes. Therefore, PLDs-CNN -Ridge-ELM is the better approach, striking an effective balance between high discrimination capability and consistent practical classification performance.

#### *4.1.3. Comparative performance analysis between the proposed model and other TL models*

Table 7 presents a comparison of the performance metrics between the proposed PLDs-CNN and several state-of-the-art transfer learning (TL) models, all evaluated using the Ridge-ELM classifier for four-class classification. Among the seven TL models assessed, VGG16 yielded the strongest results overall, with an average precision of 95 ± 0.01 %, recall of 90 ± 0.08 %, F1-score of 92.25 ± 0.04 %, accuracy of 97 %, and an AUC of 97.47 %. In contrast, the weakest performance was observed for the EfficientNetB6 model. Notably, DenseNet201 attained the highest average precision (98 ± 0.018 %) among the TL models. The proposed PLDs-CNN-Ridge-ELM model achieved 99 % accuracy in four-class classification, marking an approximate 2 % improvement over VGG16. Additionally, it recorded a superior AUC of 99.28 %, surpassing the 97.47 % AUC of VGG16. Fig. 11 displays the ROC curves for all four classes across the transfer learning models, whereas Fig. 12 provides a bar chart summarizing the overall performance of these models. These results clearly demonstrate that the PLDs-CNN-Ridge-ELM consistently outperforms the other evaluated models.

#### *4.1.4. Computational time and resource comparison*

Table 8 presents a comparative analysis of PLDs-CNN, PL-CNN, and several transfer learning (TL) feature extractors in terms of model parameters, layer count, model size, and both training and testing time. Considering accuracy, computational efficiency, and model complexity, the PLDs-CNN surpasses all other evaluated models, making it the most resource-effective method. Fig. 13 visualizes the overall computational effort and resource requirements. The findings confirm that the PLDs-CNN-Ridge-ELM method is dependable and effectively handles the classification of different waste categories.

Among all compared models, ResNet152V2 has the highest parameter count, totaling 92.41 million, with 193 layers and a model size of 567.614 MB. In contrast, InceptionResNetV2 has the greatest model size, and includes the highest number of layers at 783. Conversely, PLDs-CNN demonstrates superior efficiency, having only 1.09 million parameters, a compact size of 12.7 MB, and 9 convolutional layers (CL). The proposed model requires approximately 84.78 times fewer parameters than ResNet152V2 and 2.15 times fewer than PL-CNN, which has 2.344 million parameters. Furthermore, Ridge-ELM exhibits optimized training and inference durations, taking 0.1006 s for training and 0.0079 s for testing. Although some TL models might offer marginally faster computational times due to their architecture, they typically demand higher resources. The PLDs-CNN-Ridge-ELM model successfully balances classification performance with minimal resource usage, making it highly suitable for real-world waste management applications due to its compact architecture and reliable performance.

### *4.2. Second-stage classification: analysis of twelve-class performance*

#### *4.2.1. PL-CNN-ELM and PL-CNN Ridge-ELM*

The same multiclass categorization technique was employed for twelve-class classification. The PL-CNN model performed training,

**Table 8**
Computational resources (model parameters and size) and time comparison for multiclass classifications.

| Criteria | PL-CNN-Ridge-ELM | PLDs-CNN-Ridge-ELM | DenseNet201-Ridge-ELM | EfficientNetB6-Ridge-ELM | InceptionResNetV2-Ridge-ELM | MobileNetV3Small-Ridge-ELM | ResNet152V2-Ridge-ELM | VGG16-Ridge-ELM | Xception-Ridge-ELM |
|---|---|---|---|---|---|---|---|---|---|
| Total Parameters (Million) | 2.344 | **1.09** | 36.54 | 79.23 | 61.15 | 18.83 | 92.41 | 19.95 | 27.42 |
| Trainable Parameters (Million) | 2.341 | **1.09** | 18.22 | 38.27 | 6.81 | 17.30 | 34.08 | 5.24 | 6.55 |
| Number of Layers | **9** | **9** | 598 | 710 | 783 | 246 | 193 | 25 | 36 |
| Size (Megabyte) | 27 | **12.7** | 710280 | 669596 | 783287 | 246204 | 567614 | 22116 | 135470 |
| Training Time (Ridge-ELM) (4-class) (seconds) | 0.0657 | 0.1006 | 0.0239 | **0.01591** | 0.03024 | 0.02217 | 0.03450 | 0.01942 | 0.02722 |
| Testing Time (Ridge-ELM) (4-class) (millisecond) | 6 | 7.9 | 9.569 | 15.62 | 5 | 9.5 | 5 | **3** | 4.7 |

Note: The best results are highlighted in bold.

testing, and validation on 12566, 1552, and 1397 images, respectively. The classification performances of the models are presented in Table 9. The PL-CNN-ELM model achieved precision, recall, and F1-scores of 94.25 ± 0.04 %, 93.75 ± 0.04 %, and 94 ± 0.84 %, respectively. The model's effectiveness was improved by utilizing Ridge-ELM. The test accuracy for both the PL-CNN-ELM and the PL-CNN-Ridge-ELM is 96 %, indicating that they perform equally well. The PL-CNN-Ridge-ELM model improved the AUC by approximately 0.05 %, increasing it from 98.79 % to 98.84 %. The precision, recall, and f1-score achieved the highest values of 94.41 ± 0.04 % (with a 0.16 % improvement), 93.75 ± 0.04 %, and 94.08 ± 0.04 % (with a 0.08 % improvement), respectively. In Figs. 14 and 15, the classwise ROC curves and PR curves are displayed for both classifiers on the PL-CNN model.

#### 4.2.2. PLDs-CNN-ELM and PLDs-CNN-ridge-ELM

Fig. 16 illustrates essential information about the performance of both classifiers. Clearly, both classifiers performed very well. The precision, recall, and f1-scores demonstrated a high level of balance and surpassed 90 % for the maximum category. The results provide solid evidence for the efficacy of the Ridge-ELM approach in successfully classifying multiclass waste images. The proposed method exhibits a substantial performance advantage over conventional models, thus demonstrating its superiority. Upon further examination, a novel and less heavy model is discovered that incorporates a Ridge-ELM classifier, outperforming existing models in terms of both performance and accuracy. These findings establish a strong foundation for the application of the Ridge-ELM approach in waste image categorization. ELM showed an outstanding average precision of 94.66 ± 0.034 % across all 12 categories. ELM also achieved a recall of 93.66 ± 0.036 % and a F1-score of 94.5 ± 0.033 %. On the other hand, Ridge-ELM achieved improvements of 0.34 %, 0.67 %, and 0.16 % in comparison to ELM. The accuracy score improved by 1.0 %, increasing from 95 % to 96 %. Fig. 17 and Table 10 thoroughly examine the ability of the PLDs-CNN-ELM and PLDs-CNN-Ridge-ELM models to distinguish between 12 various waste types. The ROC curves for each category exhibit exceptional performance from both models, with ROC values surpassing 98 % for the majority of classes. Every class identification record is considered satisfactory for the proposed PLDs-CNN-Ridge-ELM framework. This emphasizes the importance of the proposed system for practical implementation in real-world scenarios. Fig. 18 displays the classwise PR curves.

### 4.3. Interpretability of PLDs-CNN-ridge-ELM using SHAP

Through a systematic examination of every conceivable combination of wastage attributes, Shapley values were formulated, giving rise to representations characterized by pixels. A distinct pattern manifested during the investigation, wherein red pixels demonstrated robust efficacy in identifying class distinctions. In the first stage of the testing phase, SHAP results were provided with explanation images for different classifications. These classifications included four classes: hazardous, household, recyclable, and residual waste. The explanation images showed that red pixels correspond to higher relevance scores for the target class. In contrast, blue pixels represent regions less associated with the predicted category. It is important to note that the SHAP visualization overlays are rendered on semi-transparent grayscale backgrounds blended into the input images, as shown in Fig. 19 (A). The top row in the SHAP visualization highlights red pixels indicating the identification of hazardous waste. In contrast, minimal presence of blue and a reduced number of red activations suggest the exclusion of alternate classes. The model assigns the hazardous class label with high certainty where dense red regions are observed. The red regions confirm the model's strongest prediction for that specific class. The second row of SHAP visualizations revealed a different structure: red activations correspond to household food waste. In that particular case, a predominance of blue pixels appeared over recyclable waste regions, suggesting reduced model confidence in that category compared to others.

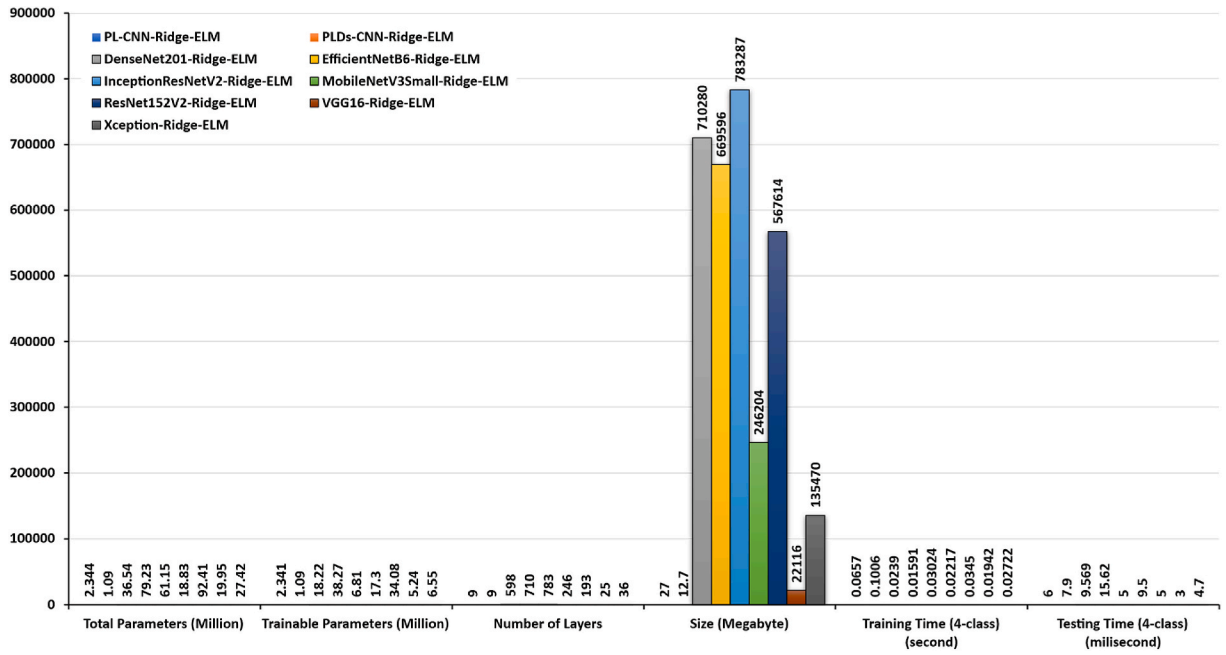


**Fig. 13.** Comparison of computational resources—including model size, number of parameters, and processing time—between the proposed PLDs-CNN-Ridge-ELM and other TL models.

**Table 9**
Performances of PL-CNN-ELM and PL-CNN-Ridge-ELM models for twelve-class waste classifications.

| Class Name | PL-CNN-ELM | | | PL-CNN-Ridge-ELM | | |
|---|---|---|---|---|---|---|
| | Precision | Recall | F1 | Precision | Recall | F1 |
| Battery (0) | 0.94 | 0.93 | 0.94 | 0.97 | 0.90 | 0.93 |
| Expired Food (1) | 0.99 | 0.97 | 0.98 | 0.96 | 0.98 | 0.97 |
| Brown Glass (2) | 0.97 | 0.97 | 0.97 | 0.97 | 0.98 | 0.98 |
| Cardboard (3) | 0.97 | 0.95 | 0.96 | 0.97 | 0.96 | 0.96 |
| Clothes (4) | 0.98 | 1.00 | 0.99 | 0.98 | 1.00 | 0.99 |
| Green Glass (5) | 0.97 | 0.97 | 0.97 | 0.98 | 0.97 | 0.98 |
| Metal (6) | 0.88 | 0.88 | 0.88 | 0.89 | 0.89 | 0.89 |
| Paper (7) | 0.94 | 0.96 | 0.95 | 0.94 | 0.95 | 0.95 |
| Plastic (8) | 0.91 | 0.87 | 0.89 | 0.90 | 0.88 | 0.89 |
| Shoes (9) | 0.96 | 0.96 | 0.96 | 0.96 | 0.96 | 0.96 |
| White Glass (10) | 0.96 | 0.95 | 0.95 | 0.96 | 0.93 | 0.94 |
| Trash (11) | 0.84 | 0.84 | 0.84 | 0.85 | 0.85 | 0.85 |
| **Average (μ) ± SD (σ) (%)** | 94.25 ± 0.04 | 93.75 ± 0.04 | 94 ± 0.84 | **94.41 ± 0.04** | **93.75 ± 0.04** | **94.08 ± 0.04** |
| **Accuracy (%)** | 96.00 | | | **96.00** | | |
| **AUC (%)** | 98.79 | | | **98.84** | | |

Note: The best results are highlighted in bold. 0–11 indicates the class number.

The proposed model was tested in the second stage, which included 12 subclasses, making it more complex to classify. Fig. 19 (B) shows how the model extracted the actual class during testing. The proposed model could accurately predict the outcome even when faced with increased data complexity. Visual SHAP explanations were used to confirm our model's results, providing a more comprehensive understanding of different classifications of waste. These explanations helped to improve the system's understanding of the various forms of waste.

### 4.4. Software and hardware development

#### 4.4.1. Graphical user interface (GUI)

For better testing and real-time implementation flexibility in the waste management industry, a graphical user interface (GUI) was designed based on PYQT5 used for the QT application framework. The GUI was programmed for three individual tasks for four-class and twelve-class classifications of the three proposed models and their real-time classifications (Fig. 20). The interface is generally designed to support conveyor belt-assisted waste sorting for four classes. For user flexibility, the tasks that need to be performed are made easily accessible by simply clicking buttons without any need for loading files for every new runtime.

The app can test any of the three proposed models from the drop-down menu. After clicking the "Classification_4_Class" or "Classification_12_Class" button, the app will open a file dialog for selecting an image for classification. Upon selecting the image, it will undergo several preprocessing tasks, such as resizing it to 124 × 124 to match the model's requirements. This is followed by rescaling the image within 0–1 for faster convergence and reduced computational load. Finally, before prediction, the dimension of the numpy image array is increased to align it to the requirements of the model. The time taken for prediction is saved along with some other important classification results, such as the top 4 class names with corresponding confidence scores in Fig. 21.

To understand the model's decision-making process for classifying images, the most influential parts are highlighted using SHAP. This approach provides a better understanding of based on which features the images are being predicted and can also assist in debugging the model (Fig. 22). The process starts by initializing the Deep Explainer using one of the models to identify the garbage object. The explanation process was based on this model, which acquired the ability to correlate certain pixel patterns with distinct classifications. In addition to the model, a background dataset was selected, which was a smaller portion of the training data. This dataset was essential since it embodied the standard input space of the model, functioning as a benchmark for comprehending the degree of novelty or conformity of a new input in relation to what the model encountered throughout its training. The subsequent step involved the calculation of SHAP values for a particular picture,

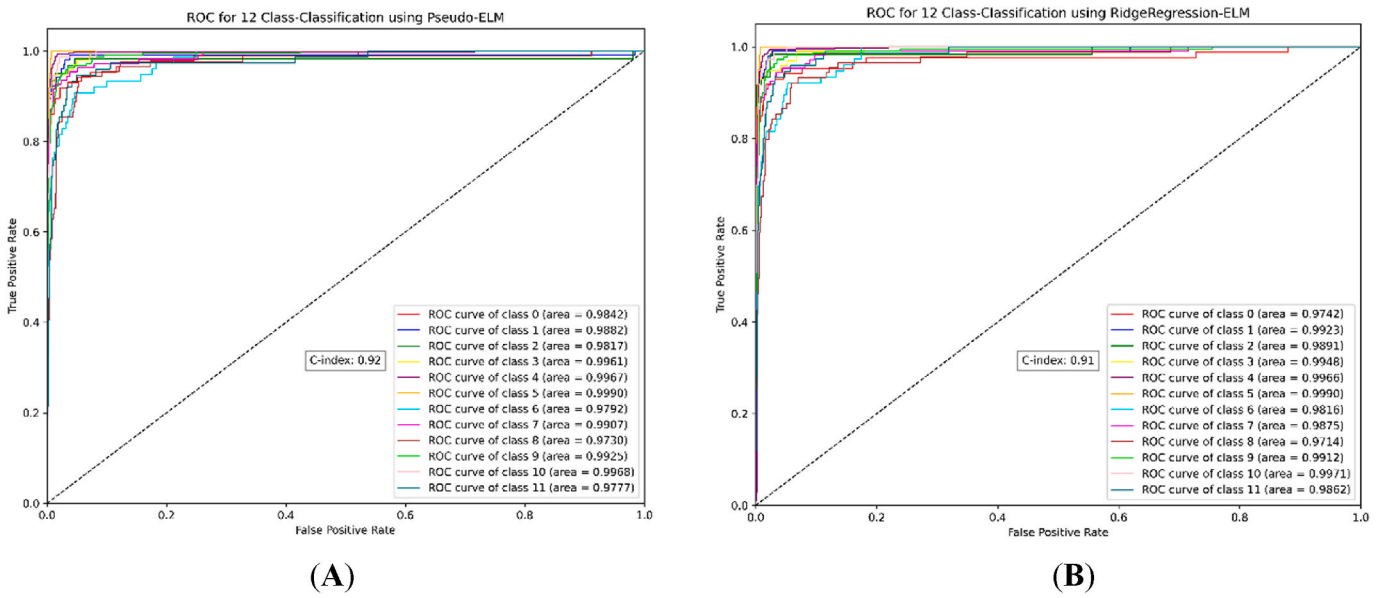


**Fig. 14.** Classwise ROCs on (A) PL-CNN-ELM and (E) PL-CNN-Ridge-ELM for twelve-class classification.

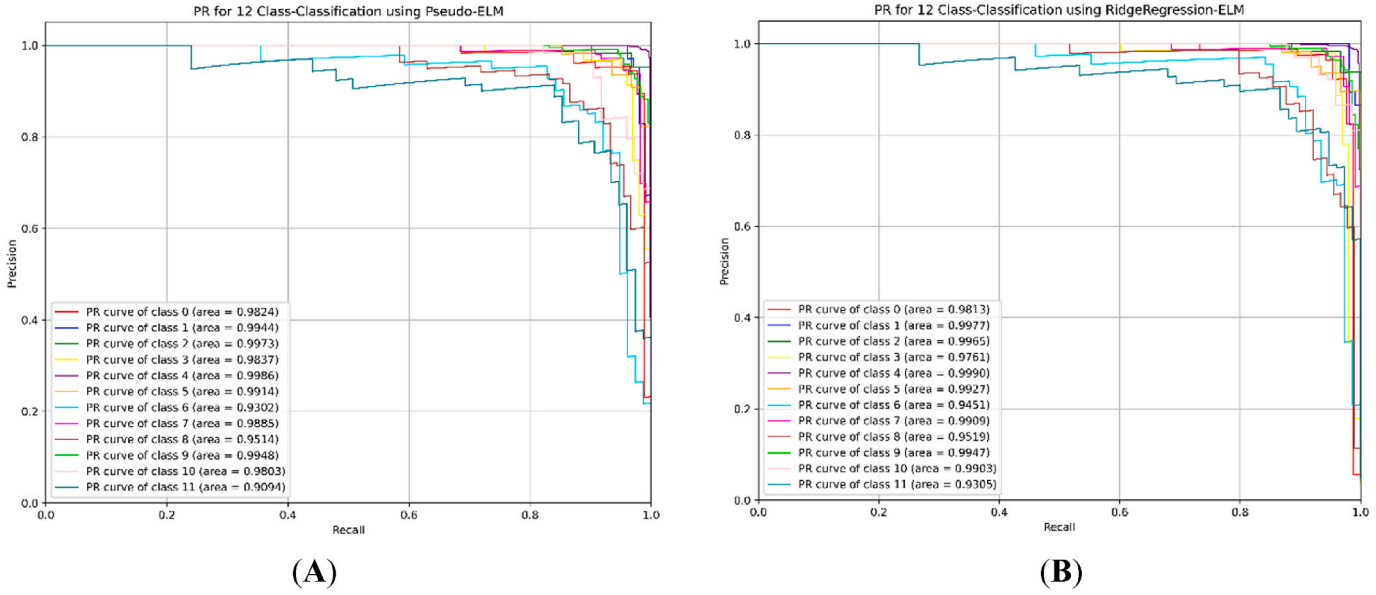


**Fig. 15.** Classwise PR curves on (A) PL-CNN-ELM and (E) PL-CNN-Ridge-ELM for classification of twelve-class.

which was the fundamental aspect of this procedure. The SHAP method measures the influence of each pixel on the model's conclusion by systematically altering the visible pixels (Linardatos et al., 2020). This process effectively determined the contribution of each pixel to the final prediction.

### 4.4.2. *Development of conveyer belt sorting mechanism*

Following accurate waste class prediction by the developed model and app, a real-world smart system can be implemented for automatically sorting waste. The total system (Fig. 23) will require a garbage chute (a) from dropping unsorted garbage items on the conveyer (b) belt. Upon passing over the device, the waste will be detected by a motion-triggered camera (e), which will use the image to predict the class to which the waste belongs in the edge device (c). Once the correct waste is identified, tray (f) will direct it to its suitable bin by moving the pantilt mechanism. The system allows loading of a webcam and takes multiple image frames of waste carried by a conveyer belt in a recycling plant and conducts classification to direct them to suitable bins for recycling or disposal. Based on the output class from the captured image, the computer sends a command to Arduino using seral communication that makes the tray turn accordingly to any of the four sides where the respective bins are placed. The waste then slides to the desired bin due to gravity. The proposed conveyor belt sorting mechanism is just an early concept for a larger and robust waste management system.

A circuit diagram of a conceptual system suitable for real-time usage is presented in Fig. 24. The model was tested on an AMD Ryzen 5 5500U processor with a base clock speed of 2.1 GHz. The processor consisted of six CPU cores with 12 threads, making it an average-performance device for running the model. In addition to the hardware architecture, several individual components, such as servo motors, microcontroller board, and power supply boxes, were used for the hardware part, as discussed further after the circuit diagram section. At the product development stage, the processing device will be replaced by an edge device with an embedded graphics processing unit (GPU), such as Jetson Orin or Nano,

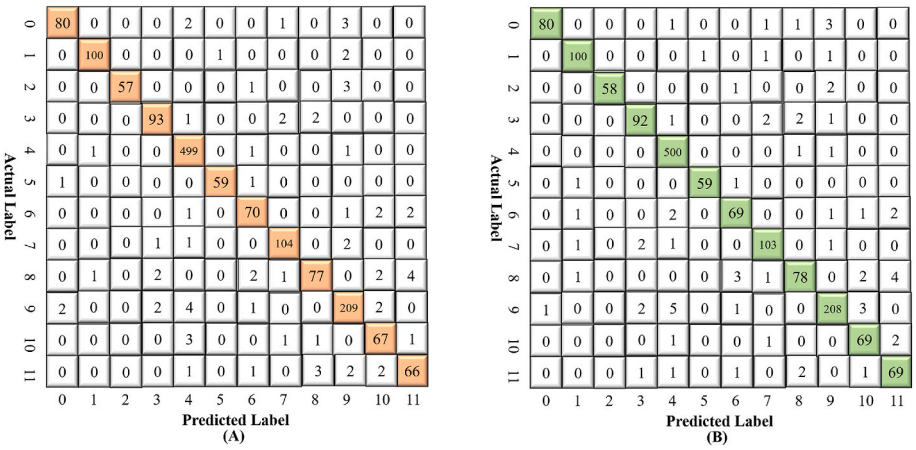


**Fig. 16.** Confusion matrices of PLDs-CNN model with (A) ELM and (B) Ridge-ELM for twelve classes.

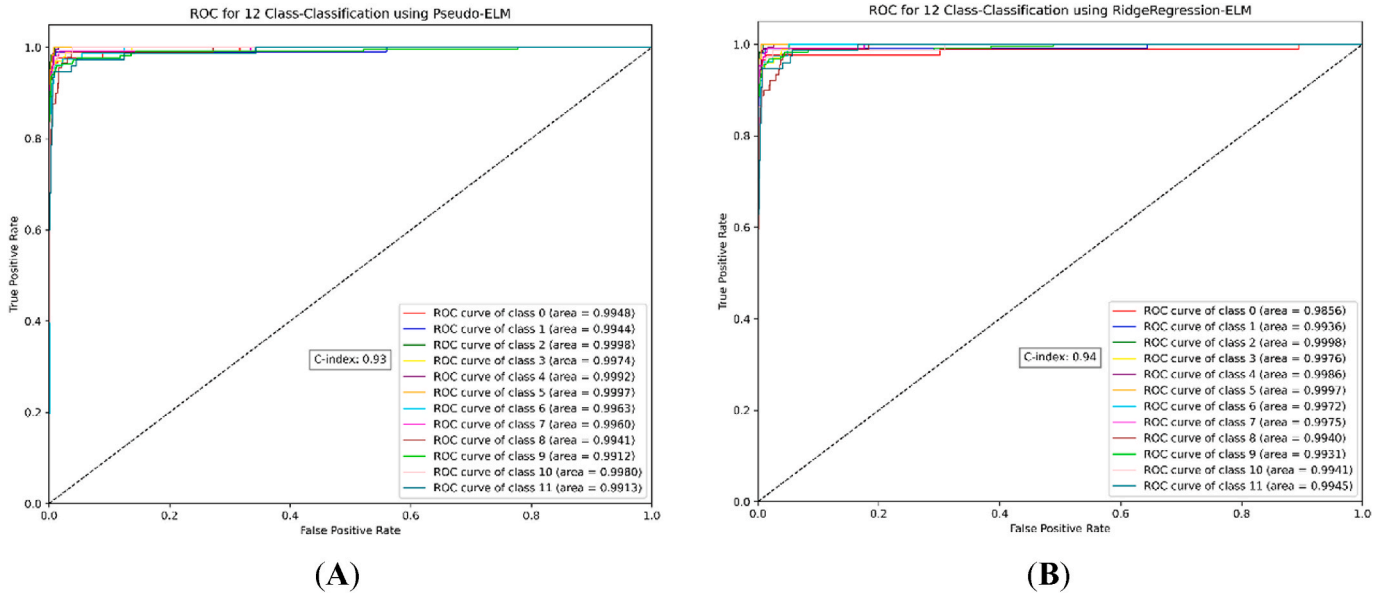


**Fig. 17.** Classwise ROCs of (A) PLDs-CNN-ELM and (E) PLDs-CNN-Ridge-ELM for twelve-class classification.

as they offer superior parallel processing power, enabling faster inference times, which are essential for real-time classification using this model.

Apart from the CPU for image classification, multiple hardware devices were utilized to develop the concept of the waste sorting system, which included two servos, an Arduino mega microcontroller board for performing directional commands, and a power supply. The two Mg996 servos had operating voltages ranging from 4.8 to 6 V with a total rotational capability of 180°. The Arduino micro had a flash memory of 256 KB, which was sufficient to store commands in it from the app GUI if there was any delay for the hardware part during the sorting period. The two power supplies used had a 24 V/10 A configuration, and the latter was converted to 5 V for operation of the servos. The system was designed to visualize the concept of a waste sorting system. However, in actual implementations, more sophisticated equipment is needed, but this work clearly demonstrated the ability of the developed software model to make correct decisions for the proposed hardware manipulations.

The preliminary design of the pan and tilt mechanism and its function were programmed using two servo motors and Arduino as the slave device respectively. The 2-axis pan-tilt mounted servo assembly in Fig. 25 was made using two 5vservos with Fused deposition modeling (FDM), stereolithography (SLA), and selective laser sintering (SLS) (3D)-printed servo brackets. The designs were made with Autodesk Tinkercad software with high precision, allowing the tray to move in any direction. Its design and the coding for tray alignment were conducted for sorting four types of trash. The initial positions are set to 0° for the pan servo and 90° for the tilt servo. After predicting the waste class for recyclable or residual waste, the tilt servo bracket/mount rotates left (Fig. 25(D)) or right (Fig. 25(B)), respectively, from its initial position, while the pan servo bracket/mount remains fixed for both of them. Similarly, for household food waste, the tilt servo bracket/mount turns right (Fig. 25 (A)), and for hazard waste turns left (Fig. 25(C)), but the pan servo bracket/mount rotates 90° counterclockwise, causing the tilt servo itself to rotate along with the tray. In this manner, depending on the predicted class, the tray is shifted to four different sides to divert incoming waste.

### *4.4.3. Demonstration of real-time waste classification*

Multiple tests were conducted to evaluate the model's real-time performance using a webcam and real waste. For this purpose, a Xiaomi Vidlok W91 webcam was used, which was attached with a mount pointing downward where the waste was placed. A white background was made with paper that represented the conveyor platform on which waste was placed, and images were captured by the camera

**Table 10**
Performances of PLDs-CNN-PELM and PLDs-CNN-RELM models for twelve-class waste classification.

| Class Name | PLDs-CNN-ELM | | | PLDs-CNN-Ridge-ELM | | |
|---|---|---|---|---|---|---|
| | Precision | Recall | F1 | Precision | Recall | F1 |
| Battery (0) | 0.96 | 0.93 | 0.95 | 0.99 | 0.93 | 0.96 |
| Expired Food (1) | 0.98 | 0.97 | 0.98 | 0.96 | 0.97 | 0.97 |
| Brown Glass (2) | 1.00 | 0.93 | 0.97 | 1.00 | 0.95 | 0.97 |
| Cardboard (3) | 0.95 | 0.95 | 0.95 | 0.95 | 0.94 | 0.94 |
| Clothes (4) | 0.97 | 0.99 | 0.99 | 0.98 | 1.00 | 0.99 |
| Green Glass (5) | 0.98 | 0.97 | 0.98 | 0.98 | 0.97 | 0.98 |
| Metal (6) | 0.91 | 0.92 | 0.92 | 0.91 | 0.91 | 0.91 |
| Paper (7) | 0.95 | 0.96 | 0.96 | 0.94 | 0.95 | 0.95 |
| Plastic (8) | 0.93 | 0.87 | 0.90 | 0.93 | 0.88 | 0.90 |
| Shoes (9) | 0.94 | 0.95 | 0.94 | 0.95 | 0.95 | 0.95 |
| White Glass (10) | 0.89 | 0.92 | 0.91 | 0.91 | 0.95 | 0.93 |
| Trash (11) | 0.90 | 0.88 | 0.89 | 0.90 | 0.92 | 0.91 |
| **Average (μ) ± SD (σ) (%)** | 94.66 ± 0.034 | 93.66 ± 0.036 | 94.5 ± 0.033 | **95 ± 0.033** | **94.33 ± 0.031** | **94.66 ± 0.02** |
| **Accuracy (%)** | 95.00 | | | **96.00** | | |
| **AUC (%)** | **99.60** | | | 99.54 | | |

Note: The best results are highlighted in bold. 0–11 indicates the class number.

above. Fig. 26 shows two tests conducted under natural light using household food and recyclable waste. In SHAP, accurate visualization was presented for both the household food waste and recyclable waste classes where the model was able to pinpoint the areas of interest accurately.

The analysis of the model showed each of the classification confidence scores and inference times with the prediction of the top four classes, allowing for a better understanding of the classification behavior of the model and its performance time even though it is relatively dependent on the device, as shown in the GUI, and computational efficiency. The computational performance of the model, in floating point operation per second (FLOPS), was found to be 39.4 G, which is acceptable for running on industrial-based computing devices. The results were obtained from CPU-based devices, which could be later replaced with a suitable GPU for faster inference on real-time classification.

Both tests were conducted at different times under natural lighting conditions, resulting in differently shaped shadows on the platform. Thus far, the model has been able to quickly and accurately classify the waste presented below the camera without the need for preprocessing the image other than its size. All the tests conducted using this model showed accurate results in terms of the class index, confidence score, and SHAP visualizations when using a high-quality webcam compared to an average one, as the model was able to extract better image data. Thus, after conducting these practical validation experiments, it was found that the model is suitable for functioning under real-life conditions even under natural light. Ideally, in industry, artificial light might be used for better visualization, which will eventually enhance the accuracy and efficiency of waste classification, ensuring optimal performance and reliability in various operational environments.

For a practical test on the functioning of the concept, a PowerPoint video of random waste images was accumulated and loaded in our app to create the same condition of wastes passing over a conveyor belt. Upon predicting each frame of waste from the video, the app shows the corresponding class (Fig. 27(A) and (C)) of the current trash along with its confidence score and immediately sends a command to Arduino via serial communication to position the tray for the correct bin (Fig. 27(B) and (D)).

In comparison to other SOTA waste classification methods for real-time analytical performance with hardware mechanisms shown in (Fan et al., 2023; Jin et al., 2023; S. Zhang et al., 2021), their highest waste classification accuracy was not more than 94.71 %, while our proposed model reached 99 % for four-class classification and 96 % for 12-class classification, making it superior. Apart from the model performance, the hardware used in these three papers had some limitations (Fan et al., 2023; S. Zhang et al., 2021). used ULN2003 and Nema series stepper motors, which are based on an open-loop system, while this study used closed-loop servo motors. Using a closed-loop system in motors provides the additional benefit of always knowing the reference position, while open-loop system motors are prone to losing this position under overloaded conditions. Additionally, there was no relevant information on the practical inference time or computational complexity of the models or the use of any GUI app for real-time monitoring and control (Fan et al., 2023; Jin et al., 2023; S. Zhang et al., 2021), whereas our proposed model integrates a user-friendly GUI application for

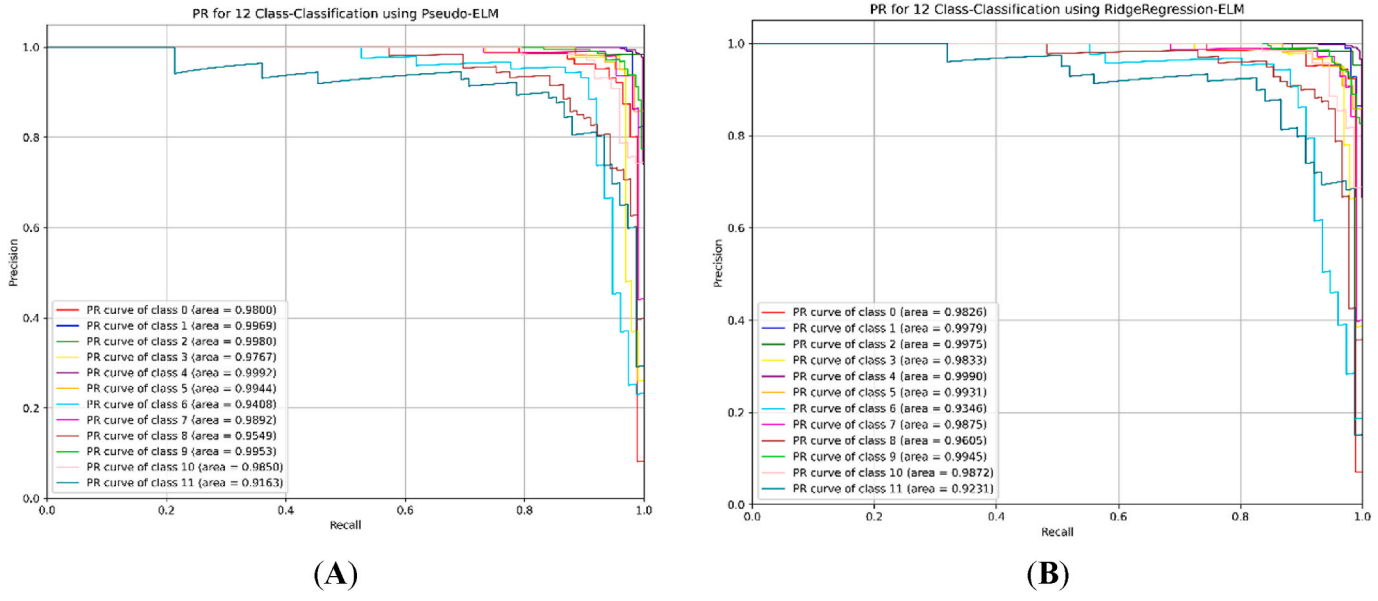


**Fig. 18.** Classwise PR curves on (A) PLDs-CNN-ELM and (E) PLDs-CNN-Ridge-ELM for twelve-class classification.

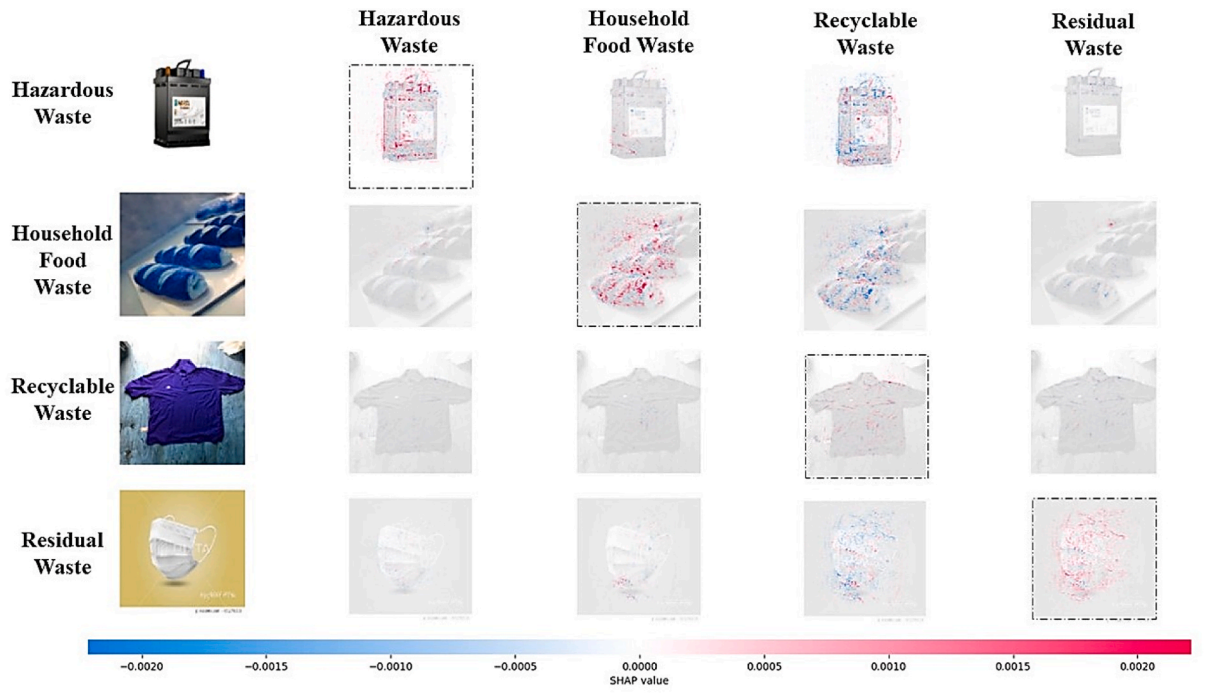


**Fig. 19.** SHAP-based visual explanations for the PLDs-CNN-Ridge-ELM model—(A) for four-class classification and (B) for twelve-class classification.

monitoring inference time of approximately 0.3 s for each frame during video inference due to reduced model complexity, confidence scores, and management of the waste classification process, further enhancing its usability and practicality in industrial applications. The concept of comparing testing times became irrelevant because the device performance configurations did not match each other. Moreover, although the current implementation is at the prototype stage, it demonstrates the potential for a cost-effective industrial solution. By using low-cost components such as Raspberry Pi, Arduino Mega, and servo motors, the system offers a promising alternative to high-cost industrial waste sorting systems. With further refinement and scaling, the approach could enable affordable deployment in small to medium-sized facilities or municipalities where budget and resource constraints are a key consideration.

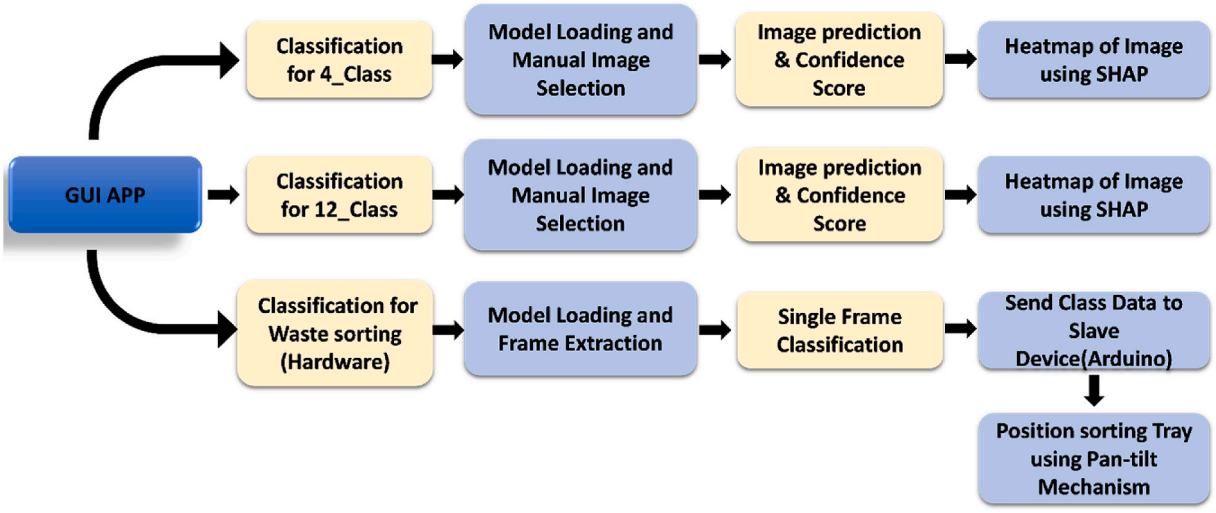


**Fig. 20.** Flowchart of GUI application.

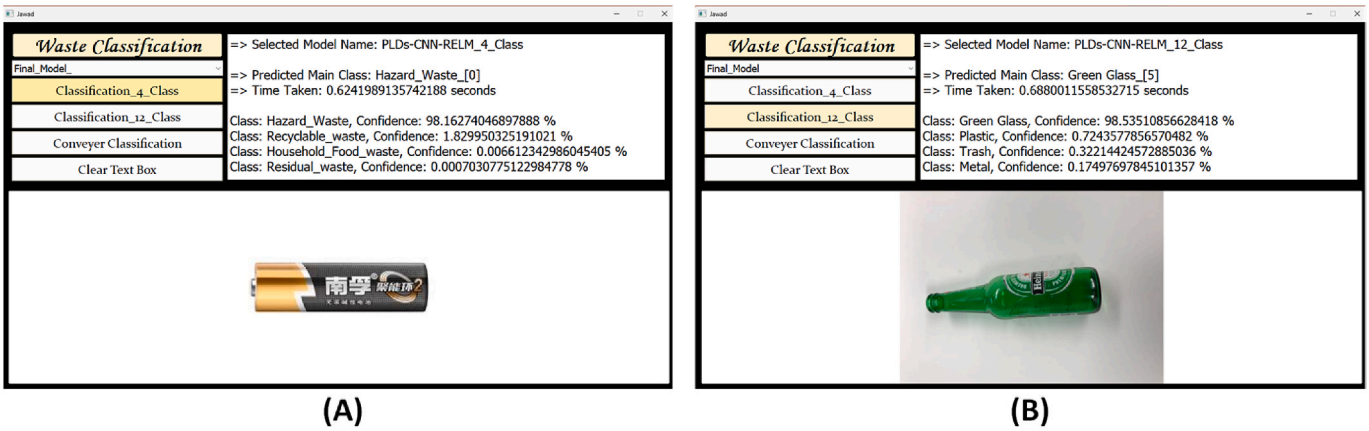


**Fig. 21.** (A) Classification of hazard items with inference time and confidence of all four classes using a four-class model, (B) Classification of green glass with inference time and confidence of the top four classes with a twelve-class model. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

### *4.5. Discussion, limitations and future work*

The design of the lightweight PLDs-CNN-Ridge-ELM model is relatively straightforward, incorporating a total of nine convolutional operations along with three fully connected layers. However, the proposed model executes the initial four convolutional layers in parallel to enhance the extraction of discriminative features, thereby reducing the overall convolutional depth from nine to four. These parallel operations contribute to more efficient feature extraction, leading to improved model performance. Additionally, the Ridge-ELM classifier's regularization term aids in refining weight updates, which further contributes to higher classification precision. As illustrated in Table 5, the suggested model demonstrates superior results compared to the other seven TL models. The central aim of this study was to introduce an architecture that improves prediction accuracy while maintaining a compact structure by reducing layer count and trainable parameters. This goal was accomplished (1.09 million parameters) by adopting depthwise separable convolutions in place of standard convolution layers and optimizing them for deployment in low-resource environments. It should also be emphasized that the dataset used in this investigation includes an inherent class distribution skew, particularly across both the four major categories and their twelve subcategories. In spite of this, the model maintained strong classification performance, confirming its capability to extract effective features from both dominant and underrepresented classes. This resistance to skewed distributions highlights the model's stability and the generalization strength of the Ridge-ELM mechanism. Furthermore, integrating SHAP into the analysis ensures that the model identifies the most informative parts of the input image for feature extraction while ignoring non-contributive areas. This enhances the interpretability of the model, allowing it to perform precise and reliable classification of waste items. Moreover, the deployment of the hardware prototype affirms the feasibility of the model in real-time waste categorization tasks. Table 11 offers a comparative summary of leading models and the developed PLDs-CNN-Ridge-ELM. According to the report by (Al-Mashhadani, 2023), InceptionV3 reached an accuracy peak of 100 %, while (Mao et al., 2021) reported 99.60 % using an optimized DenseNet121. However, in the study by (Al-Mashhadani, 2023), the evaluation was conducted on a limited dataset comprising just 1451 samples spanning four categories. Similarly, the authors in (Mao et al., 2021) utilized the TrashNet collection, which includes six classes and 2527 samples. In contrast, the current model achieved a closely matched accuracy of 99.0 % for the four-class setup using a considerably larger dataset of 15,150 images. Since DenseNet121 served as the baseline in prior work, the parameter count was significantly higher than in the present model. In addition, the proposed model recorded a minimal inference duration of 0.0079 s for the four-class

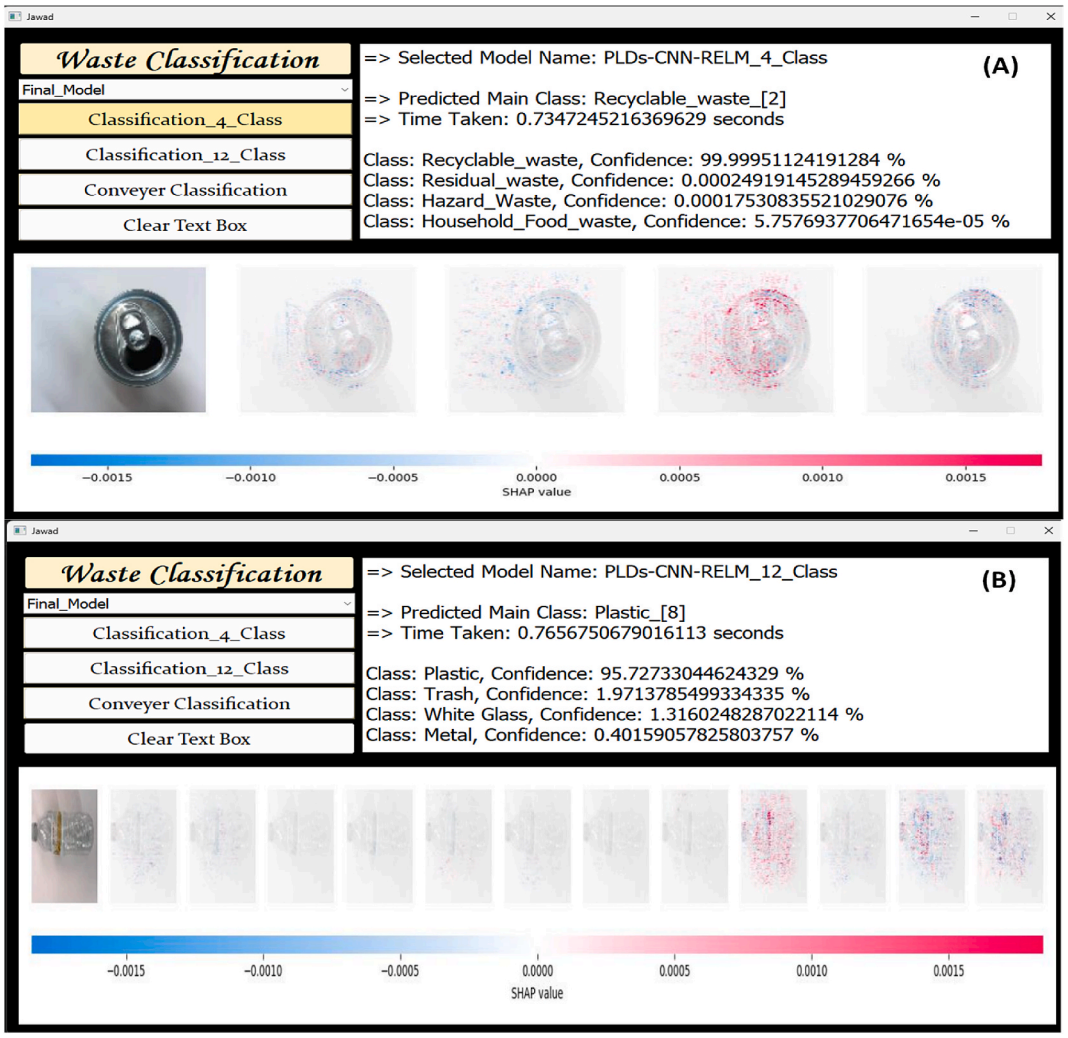


**Fig. 22.** Visualization of significant features using SHAP for (a) 4-class and (b) 12-class models for single image classification.

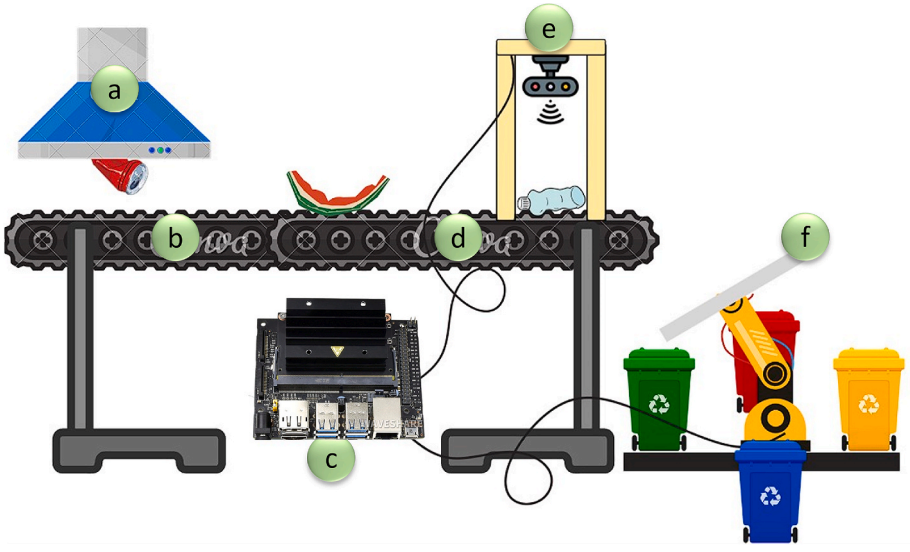


**Fig. 23.** Concept of the total waste sorting mechanism. (a) Garbage Chute, (b) Conveyer Belt, (c) Edge Device, (d) Cable Wire, (d) Motion Sensor with Camera, (f) Pan-tilt joint for trash sorting.

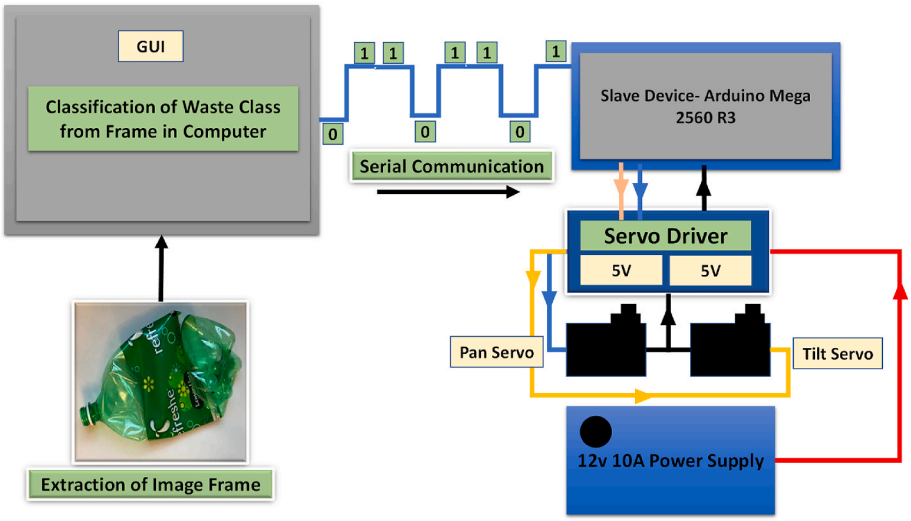


**Fig. 24.** Circuit diagram of the hardware configuration with GUI app.

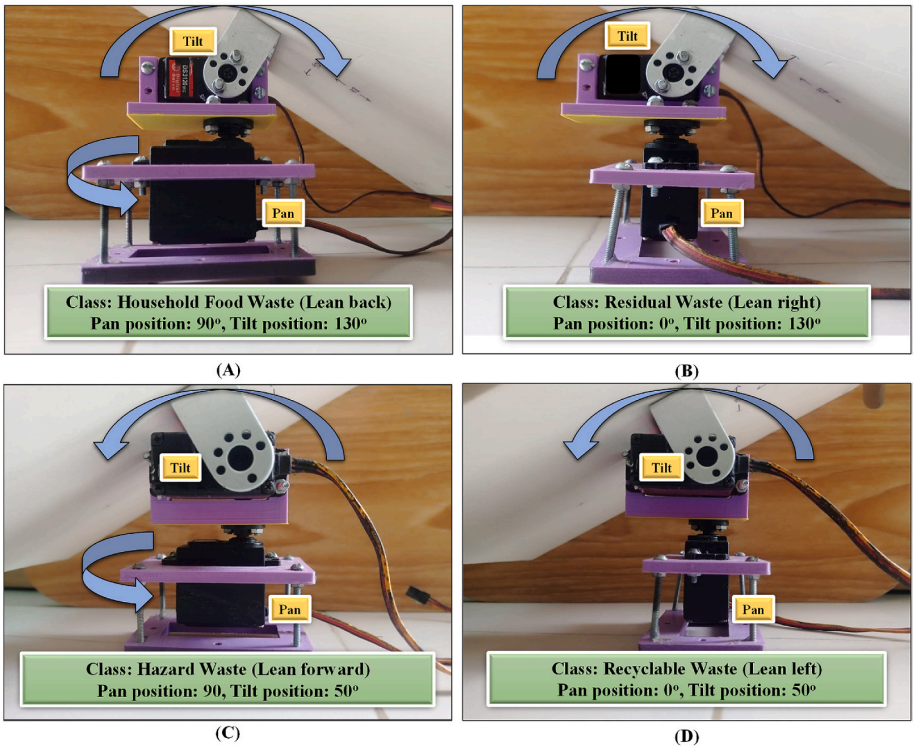


**Fig. 25.** Different servo positions based on class using a pan-tilt mechanism. (A) Servo position for household food waste. (B) Servo position for recyclable waste. (C) Servo position for hazard waste. (D) Servo position for residual waste.

setup.

Yang and Li (Z. Yang and Li, 2020) proposed a lightweight model with 1.5 million parameters. However, this model achieved lower classification accuracy (82.5 %) for larger datasets such as the Huawei garbage classification dataset. In contrast, the proposed model achieved excellent performance, with 96 % accuracy for twelve-class classification on a large dataset with only 1.09 million parameters. Feng et al. (2022) also proposed a lightweight model based on EfficientNet. However, the proposed PLDs-CNN-Ridge-ELM successfully outperformed this model in terms of accuracy, model parameters and number of images in the dataset. Additionally, Chen et al. (2022) achieved a higher accuracy of 97.9 %, exceeding the performance of the proposed model. Nevertheless, it is essential to note that they conducted research on a smaller dataset (4256 images), and additionally, their model had a greater number of parameters than did the proposed model. Moreover, unlike the introduction of SHAP by the proposed model, no studies have demonstrated real-time XAI. Table 11 also demonstrates that the reduced 4-class classification has 99 % accuracy compared to 96 % accuracy on the granular 12 classes, highlighting specialized problems associated with the sub-categories (Li et al., 2021). Reducing the number of classes simplifies decision boundaries, enhancing the learning process and promoting better generalization. Although closely related groups may share similarities, merging them into broader categories introduces a more intricate decision boundary. In a fixed dataset, a lower number of classes can provide adequate features for effective generalization. Conversely, an increase in classes without a proportional rise in data and

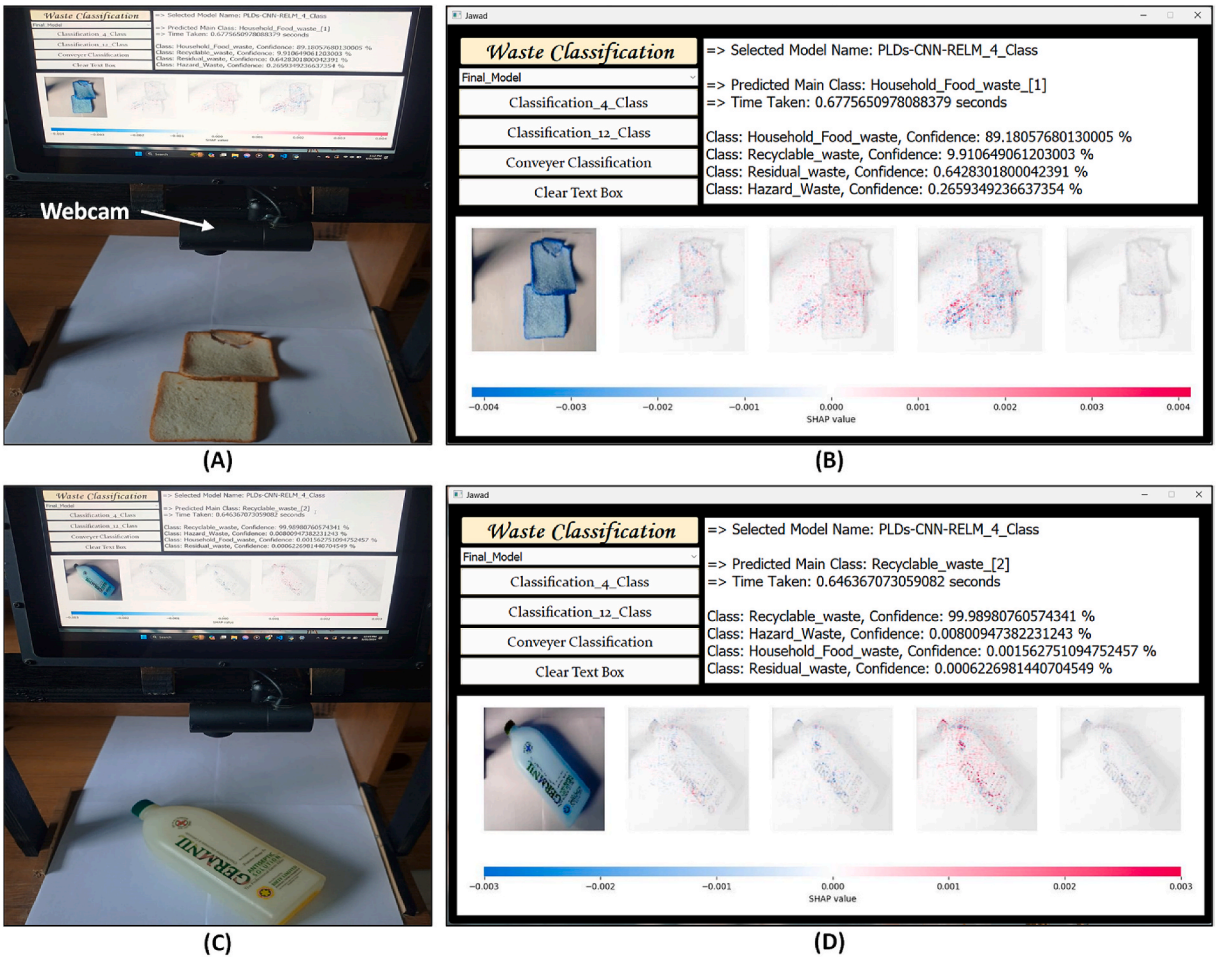


**Fig. 26.** Real-time classification using a webcam and SHAP visualization (A), (B) Household Food Waste, (C), (D) Recyclable Waste.

feature variability may lead to reduced model accuracy. However, the simplicity of decision boundaries in a smaller number of classes can help mitigate these challenges despite the limited feature set.

The proposed model can be implemented to establish an automated waste categorization system in an industrial setting, offering several significant advantages. Firstly, automation drastically reduces labor costs and minimizes human error by replacing manual sorting procedures. Secondly, the system accelerates the sorting process by accurately and efficiently identifying waste, thereby increasing throughput and optimizing workflow. This enhanced sorting capability contributes to improved recycling by accurately distinguishing recyclables from non-recyclables, thereby optimizing resource recovery and accelerating the recycling process. Such capabilities also help advance environmental sustainability by encouraging eco-friendly practices such as recycling. Lastly, by facilitating the repurposing of waste materials and creating value from previously discarded resources, the system supports a circular economy and strengthens the overall waste management ecosystem.

Despite the model's strong performance, there is room for further enhancement. The study utilized an existing dataset from Kaggle, containing 15150 images across twelve waste classes. While this dataset is larger than those used in comparable studies, it may still not fully represent the diversity of waste encountered in real-world scenarios. Additionally, certain categories—such as hazardous, household food, and residual waste—lack detailed subcategories, which may not reflect the practical complexities of waste management where, for instance, hazardous materials require specialized disposal methods. This limitation may impact the model's applicability in real-world industrial systems that demand finer categorization. Additionally, the dataset exhibits class imbalance, with certain categories—such as Recyclable Waste and Cloths—having significantly more samples than others. Despite the absence of explicit class balancing techniques, the proposed model achieved high classification accuracy (99 % in the first-stage four-class task), demonstrating strong generalization capabilities. This performance is attributed to several factors. First, the dataset consists of high-quality, well-structured images in which each image clearly contains a single type of waste object, minimizing intra-class variability and facilitating more effective learning. Second, the model architecture is highly capable of extracting discriminative features, allowing it to perform robustly even in the presence of class imbalance. These strengths enabled the model to maintain consistent performance across both majority and minority classes during evaluation.However, it is important to note that the images used for training and evaluation were relatively clean and well-structured, as shown in Fig. 2. In practical settings, waste is often dirty, occluded, or contaminated, which could negatively affect model performance. Such real-world data was not publicly available at the time of this study. Consequently, future work will focus on collecting more realistic datasets that reflect actual waste conditions in operational environments. The model will also be further fine-tuned and validated on this real-world data to ensure robustness, reliability, and deployment readiness.

Moreover, limitations such as the inability to handle incomplete, partially obscured, or mixed-type waste samples are acknowledged. To address these, future developments will explore noise-tolerant training strategies, enhanced preprocessing pipelines, and robust feature extraction methods to improve model performance under challenging

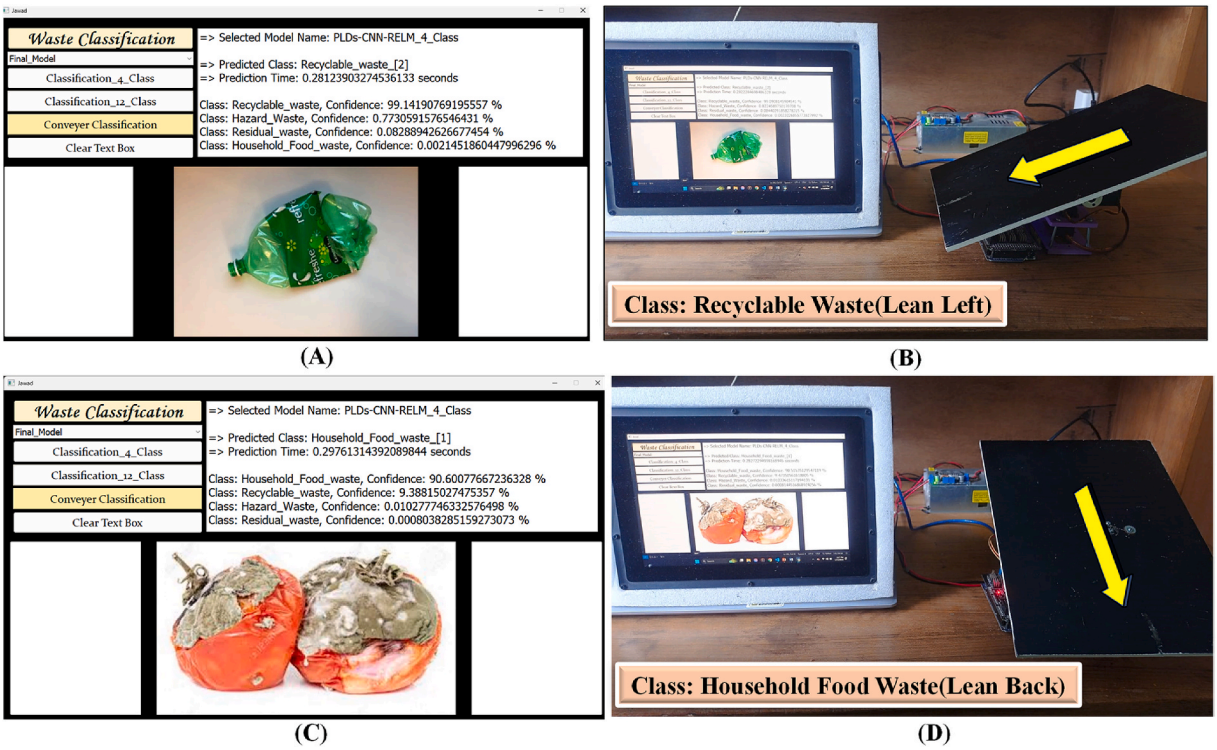


**Fig. 27.** Continuous classification process test for real-time sorting and tray movement for (a), (b) Recyclable Waste, (c), (d) Household Food Waste.

conditions. In addition, efforts will be made to collect a more comprehensive and diverse dataset, incorporating additional waste categories such as electronic, industrial, and medical waste, while also expanding subcategory coverage for all major classes. Although the proposed model is not currently suited for disaster waste management due to the structured nature of its training data and its focus on municipal solid waste, it could be extended to such applications in the future. This would require dataset adaptation using disaster-specific imagery, robust model re-training for mixed-material classification, and integration with mobile platforms such as drones or robotic units. Exploring such extensions could broaden the model's applicability in emergency response and post-disaster recovery efforts. Furthermore, techniques such as data augmentation and oversampling may be employed to mitigate class imbalance. Together, these advancements aim to improve the model's practical applicability, scalability, and overall performance in complex, real-world waste management scenarios.

In the two-stage classification approach, error propagation from the first to the second stage is minimized by ensuring that both stages operate independently. The first stage classifies waste into broader categories, while the second stage refines these categories into more specific subcategories. This separation ensures that misclassification in the first stage does not directly impact the second stage. Although both stages use the same underlying model architecture, their distinct tasks enhance the robustness and resilience of the classification process. Robust feature extraction methods are employed in both stages to capture the most relevant and discriminative features. Additionally, the integration of SHAP (Shapley Additive Explanations) enhances model interpretability by identifying the most influential features in the classification process. This transparency enables a deeper understanding of the system's decision-making process and aids in its fine-tuning. Moreover, the SHAP integration helps minimize errors by highlighting the key features responsible for accurate classification.

Although the hardware system was successfully validated for both real-time and offline scenarios, it was implemented on a small scale using limited functionality hardware, such as a webcam and generated video. To enhance performance, future work will involve upgrading the hardware with components more suitable for industrial applications. These could include motion sensors, high-precision motors, high-resolution cameras, and conveyor belt systems typically used in waste management industries. Additionally, the system could be implemented on an edge device for better portability and maintenance in real-world applications. A modified robotic arm mechanism could be integrated for faster and more accurate waste sorting, along with IoT-based tracking systems for optimized waste management. Improving computational efficiency will also be a key focus, particularly for enabling deployment on edge devices. Techniques such as model pruning, quantization, and knowledge distillation will be explored to reduce model size, memory consumption, and inference time, ensuring that the system remains responsive and resource-efficient without compromising classification performance. Finally, the approach could be expanded to include the identification of additional waste categories, such as electronic waste, industrial waste, and medical waste.

## 5. Conclusions

Precise waste classification facilitates better waste management, hence promoting environmental sustainability and optimizing resource utilization. By correctly identifying and managing hazardous waste materials, automated classification systems can protect public health and safety and increase efficiency across a variety of industries. This study introduces an efficient strategy for waste classification by integrating the PLDs-CNN feature extraction mechanism with the Ridge-ELM classifier. The PLDs-CNN model comprises nine layers and approximately 1.09 million parameters, enabling effective categorization of both four and twelve waste classes with reduced computational burden. The model demonstrated fast inference, achieving processing times of 0.0079 s for four-class and 0.0041 s for twelve-class classification tasks. Replacing the conventional pseudo ridge regression technique with Ridge-ELM significantly enhanced the model's predictive capability. The approach achieved high classification accuracy—99 % for the initial four-class task and 96 % for the twelve-class task. For the twelve-class scenario, strong performance was further confirmed by

**Table 11**
Performance comparison with state-of-the-art models and the proposed model.

| Ref. | Dataset Name | Number of Images in the Dataset | Number of Class | Best Model | Testing Accuracy (%) | Best Model's Parameters (million) | Testing Time (seconds) | Real-time XAI |
|---|---|---|---|---|---|---|---|---|
| Mao et al. (2021) | TrashNet | 2527 | 6 | DenseNet121 (Optimized) | 99.60 | 7.2 | – | None |
| Nowakowski and Pamuła (2020) | Custom | 210 | 3 | Deep CNN and R-CNN | 96.7 | – | – | None |
| Khan et al. (2022) | Kaggle Garbage Categorization Dataset | 750 | 6 | RWC-EPODL | 98.96 | – | – | None |
| Lin et al. (2022) | TrashNet | 2527 | 6 | RWNet | 88.8 | 58.5 | – | None |
| Abdulkareem et al. (2024) | Custom | 1451 | 4 | ResNet50-GoogleNet-Inception | 98 | – | – | None |
| Kumar et al. (2021) | Custom | 2400 | 4 | support vector machine (SVM) | 96.5 | – | – | None |
| Al-Mashhadani (2023) | Custom | 1451 | 4 | InceptionV3 | 100 % | – | – | None |
| Yang and Li (2020) | (i) TrashNet, (ii) Huawei Garbage Classification Dataset | 252 and 18079 | 6 (TrashNet), 4 (Huawei Garbage Classification Dataset) | WasNet | 96.10 (For TrashNet Dataset), 82.5 (For Huawei Garbage Classification Dataset) | 1.5 | – | None |
| Z. Chen et al. (2022) | Custom | 4256 | 4 (14 sub-classes) | GCNet (Improved ShuffleNetv2) | 97.9 (For 14 sub-class) | 1.3 | – | None |
| Fan et al. (2023) | Huawei Cloud Garbage Classification Dataset | 14802 | 4 | EfficientNetB2 with PMAM | 93.38 | 7.8 | 6.756 | None |
| Feng et al. (2022) | Custom | 7361 | 4 (18 sub-classes) | GECM-EfficienNet | 94.54 (For 18 sub-class) | 1.23 | – | None |
| Jin et al. (2023) | Huawei Garbage Classification Challenge Cup Dataset | 14683 | 4 | Improved MobileNetV2 | 90.7 | 3.4 | – | None |
| S. Zhang et al. (2021) | Custom | 1040 | 4 (13 sub-classes) | RevM | 94.71 (For 4 class) | – | – | None |
| **Proposed Model** | Kaggle Garbage Classification Dataset | 15150 | 4 (12 sub-classes) | PLDs-CNN-Ridge-ELM | **99.0** (For 4 class), 96.0 (For 12 sub-class) | 1.09 | **0.0079 (For 4 class), 0.0041 (For 12 sub-class)** | **SHAP** |

precision, recall, and F1-scores of 95 ± 0.033 %, 94.33 ± 0.031 %, and 94.66 ± 0.02 %, respectively, along with an outstanding AUC score of 99.54 %. With a compact model size of just 12.7 MB, this method is highly suitable for deployment in practical waste management solutions, particularly on low-resource edge devices. The integration of real-time SHAP explainability adds value for end-users by offering clear and trustworthy interpretation of the model's decisions, improving the credibility of classification outcomes. These encouraging findings open up opportunities for advancing intelligent and sustainable waste sorting systems. Furthermore, the proposed waste classification model was successfully realized in both hardware and software prototype implementations, validating its feasibility in operational environments. Overall, the PLDs-CNN Ridge-ELM model significantly enhances waste classification accuracy while remaining practical for real-world applications.

## CRediT authorship contribution statement

**Mansura Naznine:** Writing – original draft, Validation, Investigation, Data curation, Writing – review & editing, Visualization, Methodology, Formal analysis, Conceptualization. **Md. Nahiduzzaman:** Writing – review & editing, Visualization, Methodology, Formal analysis, Conceptualization, Writing – original draft, Validation, Investigation, Data curation. **Md. Jawadul Karim:** Writing – original draft, Validation, Investigation, Data curation, Writing – review & editing, Visualization, Methodology, Formal analysis, Conceptualization. **Md. Faysal Ahamed:** Writing – review & editing, Visualization, Methodology, Formal analysis, Conceptualization, Writing – original draft, Validation, Investigation, Data curation. **Abdus Salam:** Writing – original draft, Validation, Formal analysis, Writing – review & editing, Visualization, Methodology, Conceptualization. **Mohamed Arselene Ayari:** Visualization, Supervision, Investigation, Conceptualization, Writing – review & editing, Validation, Methodology, Formal analysis. **Amith Khandakar:** Writing – original draft, Validation, Investigation, Conceptualization, Writing – review & editing, Visualization, Methodology, Formal analysis. **Azad Ashraf:** Writing – original draft, Validation, Investigation, Conceptualization, Writing – review & editing, Visualization, Methodology, Formal analysis. **Mominul Ahsan:** Visualization, Supervision, Investigation, Conceptualization, Writing – review & editing, Validation, Methodology, Formal analysis. **Julfikar Haider:** Writing – review & editing, Validation, Visualization, Supervision, Investigation, Conceptualization, Methodology, Formal analysis.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgement

The authors employed several AI-based tools to improve the clarity and linguistic quality of the manuscript during its preparation. Following this, they thoroughly reviewed and refined the content as

needed. All aspects of the core research, including the study design, analysis, and conclusions, remain the sole responsibility of the authors.

## Data availability

Data will be made available on request.

## References


Abdulkareem, K.H., Subhi, M.A., Mohammed, M.A., Aljibawi, M., Nedoma, J., Martinek, R., et al., 2024. A manifold intelligent decision system for fusion and benchmarking of deep waste-sorting models. Eng. Appl. Artif. Intell. 132, 107926. https://doi.org/10.1016/j.engappai.2024.107926.

Abuga, D., Raghava, N.S., 2021. Real-time smart garbage bin mechanism for solid waste management in smart cities. Sustain. Cities Soc. 75, 103347. https://doi.org/10.1016/j.scs.2021.103347.

Ahsan, A., Alamgir, M., El-Sergany, M.M., Shams, S., Rowshon, M.K., Daud, N.N.N., 2014. Assessment of municipal solid waste management system in a developing country. Chin. J. Eng. 1–11. https://doi.org/10.1155/2014/561935.

Al-Mashhadani, I.B., 2023. Waste material classification using performance evaluation of deep learning models. J. Intell. Syst. 32 (1). https://doi.org/10.1515/jisys-2023-0064.

Ashikuzzaman, Md, Howlader, Md H., 2020. Sustainable solid waste management in Bangladesh. https://doi.org/10.4018/978-1-7998-0198-6.ch002.

Bhandari, M., Shahi, T.B., Siku, B., Neupane, A., 2022. Explanatory classification of CXR images into COVID-19, Pneumonia and Tuberculosis using deep learning and XAI. Comput. Biol. Med. 150, 106156. https://doi.org/10.1016/j.compbiomed.2022.106156.

Bhatia, Y., Bajpayee, A., Raghuvanshi, D., Mittal, H., 2019. Image captioning using google's inception-resnet-v2 and recurrent neural network. In: 2019 Twelfth International Conference on Contemporary Computing (IC3). IEEE, pp. 1–6. https://doi.org/10.1109/IC3.2019.8844921.

Carrera, B., Piñol, V.L., Mata, J.B., Kim, K., 2022. A machine learning based classification models for plastic recycling using different wavelength range spectrums. J. Clean. Prod. 374, 133883. https://doi.org/10.1016/j.jclepro.2022.133883.

Cheah, C.G., Chia, W.Y., Lai, S.F., Chew, K.W., Chia, S.R., Show, P.L., 2022. Innovation designs of industry 4.0 based solid waste management: machinery and digital circular economy. Environ. Res. 213, 113619. https://doi.org/10.1016/j.envres.2022.113619.

Chen, Y., Luo, A., Cheng, M., Wu, Y., Zhu, J., Meng, Y., Tan, W., 2023. Classification and recycling of recyclable garbage based on deep learning. J. Clean. Prod. 414, 137558. https://doi.org/10.1016/j.jclepro.2023.137558.

Chen, Z., Yang, J., Chen, L., Jiao, H., 2022. Garbage classification system based on improved ShuffleNet v2. Resour. Conserv. Recycl. 178, 106090. https://doi.org/10.1016/j.resconrec.2021.106090.

Chollet, F., 2017. Xception: deep learning with depthwise separable convolutions. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, pp. 1800–1807. https://doi.org/10.1109/CVPR.2017.195.

Fan, H., Dong, Q., Guo, N., Xue, J., Zhang, R., Wang, H., Shi, M., 2023. Raspberry Pi-based design of intelligent household classified garbage bin. Internet of Things 24, 100987. https://doi.org/10.1016/j.iot.2023.100987.

Feng, Z., Yang, J., Chen, L., Chen, Z., Li, L., 2022. An intelligent waste-sorting and recycling device based on improved EfficientNet. Int. J. Environ. Res. Publ. Health 19 (23), 15987. https://doi.org/10.3390/ijerph192315987.

Gaba, S., Budhiraja, I., Kumar, V., Garg, S., Kaddoum, G., Hassan, M.M., 2022. A federated calibration scheme for convolutional neural networks: models, applications and challenges. Comput. Commun. 192, 144–162. https://doi.org/10.1016/j.comcom.2022.05.035.

Habib, Md A., Ahmed, M.M., Aziz, M., Beg, Mohd R.A., Hoque, Md E., 2021. Municipal solid waste management and waste-to-energy potential from rajshahi city corporation in Bangladesh. Appl. Sci. 11 (9), 3744. https://doi.org/10.3390/app11093744.

Haque, A.K.M., Razy, S., 2021. Practices of 3Rs (Reduce, Reuse and Recycle) Strategy in Urban Solid Waste Management in Rajshahi City Corporation of Bangladesh, vol. 23, pp. 75–87.

Harris, N.L., Gibbs, D.A., Baccini, A., Birdsey, R.A., de Bruin, S., Farina, M., et al., 2021. Global maps of twenty-first century forest carbon fluxes. Nat. Clim. Change 11 (3), 234–240. https://doi.org/10.1038/s41558-020-00976-6.

He, K., Zhang, X., Ren, S., Sun, J., 2015. Deep Residual Learning for Image Recognition.

Hoornweg, D., Bhada-Tata, P., 2012. What a waste: a global review of solid waste management. Urban Dev. Ser. Knowl. Pap. 15, 87–88.

Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q., 2017. Densely connected convolutional networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, pp. 2261–2269. https://doi.org/10.1109/CVPR.2017.243.

Jaunich, M.K., Levis, J.W., DeCarolis, J.F., Barlaz, M.A., Ranjithan, S.R., 2019. Solid waste management policy implications on waste process choices and systemwide cost and greenhouse gas performance. Environ. Sci. Technol. 53 (4), 1766–1775. https://doi.org/10.1021/acs.est.8b04589.

Jerin, D.T., Sara, H.H., Radia, M.A., Hema, P.S., Hasan, S., Urme, S.A., et al., 2022. An overview of progress towards implementation of solid waste management policies in Dhaka, Bangladesh. Heliyon 8 (2), e08918. https://doi.org/10.1016/j.heliyon.2022.e08918.

Jiang, P., Zhang, L., You, S., Fan, Y. Van, Tan, R.R., Klemeš, J.J., You, F., 2023. Blockchain technology applications in waste management: overview, challenges and opportunities. J. Clean. Prod. 421, 138466. https://doi.org/10.1016/j.jclepro.2023.138466.

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

Kang, K.D., Kang, H., Ilankoon, I.M.S.K., Chong, C.Y., 2020. Electronic waste collection systems using Internet of Things (IoT): household electronic waste management in Malaysia. J. Clean. Prod. 252, 119801. https://doi.org/10.1016/j.jclepro.2019.119801.

Kaza, S., Bhada-Tata, P., 2018. Decision Maker's Guides for Solid Waste Management Technologies. World Bank, Washington, DC. https://doi.org/10.1596/31694.

Kaza, S., Yao, L.C., Bhada-Tata, P., Van Woerden, F., 2018. What a Waste 2.0: A Global Snapshot of Solid Waste Management to 2050. World Bank, Washington, DC. https://doi.org/10.1596/978-1-4648-1329-0.

Khan, A.I., Almalaise Alghamdi, A.S., Abushark, Y.B., Alsolami, F., Almalawi, A., Marish Ali, A., 2022. Recycling waste classification using emperor penguin optimizer with deep learning model for bioenergy production. Chemosphere 307, 136044. https://doi.org/10.1016/j.chemosphere.2022.136044.

Kibria, H.B., Nahiduzzaman, M., Goni, Md O.F., Ahsan, M., Haider, J., 2022. An ensemble approach for the prediction of diabetes mellitus using a soft voting classifier with an explainable AI. Sensors 22 (19), 7268. https://doi.org/10.3390/s22197268.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2017. ImageNet classification with deep convolutional neural networks. Commun. ACM 60 (6), 84–90. https://doi.org/10.1145/3065386.

Kumar, N.M., Mohammed, M.A., Abdulkareem, K.H., Damasevicius, R., Mostafa, S.A., Maashi, M.S., Chopra, S.S., 2021. Artificial intelligence-based solution for sorting COVID related medical waste streams and supporting data-driven decisions for smart circular economy practice. Process Saf. Environ. Prot. 152, 482–494. https://doi.org/10.1016/j.psep.2021.06.026.

Li, Y., Zhang, X., 2024. Intelligent X-ray waste detection and classification via X-ray characteristic enhancement and deep learning. J. Clean. Prod. 435, 140573. https://doi.org/10.1016/j.jclepro.2024.140573.

Li, Z., Kamnitsas, K., Glocker, B., 2021. Analyzing overfitting under class imbalance in neural networks for image segmentation. IEEE Trans. Med. Imag. 40 (3), 1065–1077. https://doi.org/10.1109/TMI.2020.3046692.

Lin, K., Zhao, Y., Gao, X., Zhang, M., Zhao, C., Peng, L., et al., 2022a. Applying a deep residual network coupling with transfer learning for recyclable waste sorting. Environ. Sci. Pollut. Control Ser. 29 (60), 91081–91095. https://doi.org/10.1007/s11356-022-22167-w.

Lin, K., Zhao, Y., Kuo, J.-H., Deng, H., Cui, F., Zhang, Z., et al., 2022b. Toward smarter management and recovery of municipal solid waste: a critical review on deep learning approaches. J. Clean. Prod. 346, 130943. https://doi.org/10.1016/j.jclepro.2022.130943.

Linardatos, P., Papastefanopoulos, V., Kotsiantis, S., 2020. Explainable AI: a review of machine learning interpretability methods. Entropy 23 (1), 18. https://doi.org/10.3390/e23010018.

Lu, X., Pu, X., Han, X., 2020. Sustainable smart waste classification and collection system: a bi-objective modeling and optimization approach. J. Clean. Prod. 276, 124183. https://doi.org/10.1016/j.jclepro.2020.124183.

Lundberg, S., Lee, S.-I., 2017. A Unified Approach to Interpreting Model Predictions.

Maheta, S., Manisha, 2023. Deep Learning-Based Cancelable Biometric Recognition Using MobileNetV3Small Model. https://doi.org/10.1007/978-981-99-1203-2_29.

Mao, W.-L., Chen, W.-C., Fathurrahman, H.I.K., Lin, Y.-H., 2022. Deep learning networks for real-time regional domestic waste detection. J. Clean. Prod. 344, 131096. https://doi.org/10.1016/j.jclepro.2022.131096.

Mao, W.-L., Chen, W.-C., Wang, C.-T., Lin, Y.-H., 2021. Recycling waste classification using optimized convolutional neural network. Resour. Conserv. Recycl. 164, 105132. https://doi.org/10.1016/j.resconrec.2020.105132.

Mohammed, M.A., Abdulhasan, M.J., Kumar, N.M., Abdulkareem, K.H., Mostafa, S.A., Maashi, M.S., et al., 2023. Automated waste-sorting and recycling classification using artificial neural network and features fusion: a digital-enabled circular economy vision for smart cities. Multimed. Tool. Appl. 82 (25), 39617–39632. https://doi.org/10.1007/s11042-021-11537-0.

MOSTAFA MOHAMED, 2021. Garbage Classification (12 Classes). Retrieved from. https://www.kaggle.com/datasets/mostafaabla/garbage-classification?fbclid=IwAR3OjKZepc8ML8AmwJek_Iwb2JGL7VtCeRDhsIn1GALeLfi5CUptBVQVwmY.

Nahiduzzaman, Md, Chowdhury, M.E.H., Salam, A., Nahid, E., Ahmed, F., Al-Emadi, N., et al., 2023a. Explainable deep learning model for automatic mulberry leaf disease classification. Front. Plant Sci. 14. https://doi.org/10.3389/fpls.2023.1175515.

Nahiduzzaman, Md, Faisal Abdulrazak, L., Arselene Ayari, M., Khandakar, A., Islam, S.M. R., 2024. A novel framework for lung cancer classification using lightweight convolutional neural networks and ridge extreme learning machine model with SHapley Additive exPlanations (SHAP). Expert Syst. Appl. 248, 123392. https://doi.org/10.1016/j.eswa.2024.123392.

Nahiduzzaman, Md, Goni, Md O.F., Anower, Md S., Islam, Md R., Ahsan, M., Haider, J., et al., 2021a. A novel method for multivariant pneumonia classification based on hybrid CNN-PCA based feature extraction using extreme learning machine with CXR images. IEEE Access 9, 147512–147526. https://doi.org/10.1109/ACCESS.2021.3123782.

Nahiduzzaman, Md, Goni, Md O.F., Hassan, R., Islam, Md R., Syfullah, M.K., Shahriar, S. M., et al., 2023b. Parallel CNN-ELM: a multiclass classification of chest X-ray images

to identify seventeen lung diseases including COVID-19. Expert Syst. Appl. 229, 120528. https://doi.org/10.1016/j.eswa.2023.120528.
Nahiduzzaman, Md, Islam, Md R., Hassan, R., 2023c. ChestX-Ray6: prediction of multiple diseases including COVID-19 from chest X-ray images using convolutional neural network. Expert Syst. Appl. 211, 118576. https://doi.org/10.1016/j.eswa.2022.118576.
Nahiduzzaman, Md, Islam, Md R., Islam, S.M.R., Goni, Md O.F., Anower, Md S., Kwak, K.-S., 2021b. Hybrid CNN-SVD based prominent feature extraction and selection for grading diabetic retinopathy using extreme learning machine algorithm. IEEE Access 9, 152261–152274. https://doi.org/10.1109/ACCESS.2021.3125791.
Nahiduzzaman, Md, Robiul Islam, Md, Omaer Faruq Goni, Md, Shamim Anower, Md, Ahsan, M., Haider, J., Kowalski, M., 2023d. Diabetic retinopathy identification using parallel convolutional neural network based feature extractor and ELM classifier. Expert Syst. Appl. 217, 119557. https://doi.org/10.1016/j.eswa.2023.119557.
Nowakowski, P., Pamuła, T., 2020. Application of deep learning object classifier to improve e-waste collection planning. Waste Manag. 109, 1–9. https://doi.org/10.1016/j.wasman.2020.04.041.
Özkan, K., Ergin, S., Işık, Ş., Işıklı, İ., 2015. A new classification scheme of plastic wastes based upon recycling labels. Waste Manag. 35, 29–35. https://doi.org/10.1016/j.wasman.2014.09.030.
Qiao, Y., Zhang, Q., Qi, Y., Wan, T., Yang, L., Yu, X., 2023. A waste classification model in low-illumination scenes based on ConvNeXt. Resour. Conserv. Recycl. 199, 107274. https://doi.org/10.1016/j.resconrec.2023.107274.
Rahman, A.U., Saeed, M., Mohammed, M.A., Abdulkareem, K.H., Nedoma, J., Martinek, R., 2023. Fppsv-NHSS: fuzzy parameterized possibility single valued neutrosophic hypersoft set to site selection for solid waste management. Appl. Soft Comput. 140, 110273. https://doi.org/10.1016/j.asoc.2023.110273.
Shams, S., Sahu, J.N., Rahman, S.M.S., Ahsan, A., 2017. Sustainable waste management policy in Bangladesh for reduction of greenhouse gases. Sustain. Cities Soc. 33, 18–26. https://doi.org/10.1016/j.scs.2017.05.008.
Sudha, V., Ganeshbabu, T.R., 2020. A convolutional neural network classifier VGG-19 architecture for lesion detection and grading in diabetic retinopathy based on deep learning. Comput. Mater. Continua (CMC) 66 (1), 827–842. https://doi.org/10.32604/cmc.2020.012008.
Tan, M., Le, Q.V., 2019. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks.
Wang, S., Wang, J., Yang, S., Li, J., Zhou, K., 2020. From intention to behavior: comprehending residents' waste sorting intention and behavior formation process. Waste Manag. 113, 41–50. https://doi.org/10.1016/j.wasman.2020.05.031.
Yang, J., Xu, Y.-P., Chen, P., Li, J.-Y., Liu, D., Chu, X.-L., 2023. Combining spectroscopy and machine learning for rapid identification of plastic waste: recent developments and future prospects. J. Clean. Prod. 431, 139771. https://doi.org/10.1016/j.jclepro.2023.139771.
Yang, Z., Li, D., 2020. WasNet: a neural network-based garbage collection management system. IEEE Access 8, 103984–103993. https://doi.org/10.1109/ACCESS.2020.2999678.
Zhang, Q., Yang, Q., Zhang, X., Bao, Q., Su, J., Liu, X., 2021. Waste image classification based on transfer learning and convolutional neural network. Waste Manag. 135, 150–157. https://doi.org/10.1016/j.wasman.2021.08.038.
Zhang, Q., Yang, Q., Zhang, X., Wei, W., Bao, Q., Su, J., Liu, X., 2022. A multi-label waste detection model based on transfer learning. Resour. Conserv. Recycl. 181, 106235. https://doi.org/10.1016/j.resconrec.2022.106235.
Zhang, Q., Zhang, X., Mu, X., Wang, Z., Tian, R., Wang, X., Liu, X., 2021. Recyclable waste image recognition based on deep learning. Resour. Conserv. Recycl. 171, 105636. https://doi.org/10.1016/j.resconrec.2021.105636.
Zhang, S., Chen, Y., Yang, Z., Gong, H., 2021. Computer vision based two-stage waste recognition-retrieval algorithm for waste classification. Resour. Conserv. Recycl. 169, 105543. https://doi.org/10.1016/j.resconrec.2021.105543.
Zhang, Y.-L., Kim, Y.-C., Cha, G.-W., 2023. Assessment of deep learning-based image analysis for disaster waste identification. J. Clean. Prod. 428, 139351. https://doi.org/10.1016/j.jclepro.2023.139351.
Zhao, C., Shuai, R., Ma, L., Liu, W., Hu, D., Wu, M., 2021. Dermoscopy image classification based on StyleGAN and DenseNet201. IEEE Access 9, 8659–8679. https://doi.org/10.1109/ACCESS.2021.3049600.