KSÜ Mühendislik Bilimleri Dergisi, 28(2), 2025
Araştırma Makalesi

KSU J Eng Sci, 28(2), 2025
Research Article

Kahramanmaras Sutcu Imam University
Journal of Engineering Sciences

Geliş Tarihi : 18.02.2025
Kabul Tarihi :21.04.2025

Received Date : 18.02.2025
Accepted Date : 21.04.2025

INNOVATIVE APPROACH TO ENVIRONMENTAL WASTE CLASSIFICATION WITH ENSEMBLE LEARNING MODELS

TOPLULUK ÖĞRENME MODELLERİ İLE ÇEVRESEL ATIK SINIFLANDIRMADA YENİLİKÇİ YAKLAŞIM

Mehmet Umut SALUR[1*] (ORCID: 0000-0003-0296-6266)
Nermin ELMAS[1] (ORCID: 0009-0004-5617-6344)

[1] Gaziantep Islam Science and Technology University, Department of Computer Engineering, Gaziantep, Turkey

*Sorumlu Yazar / Corresponding Author: Mehmet Umut SALUR, mehmetumut.salur@gibtu.edu.tr

ABSTRACT

Increasing waste production and inadequate waste management have further complicated global environmental problems. The limited natural resources and the damage caused by waste to the environment necessitate the improvement of waste management systems. Accurate and effective classification of waste provides both economic benefits and reduces environmental impacts. In this study, a hybrid approach is presented by combining deep learning, machine learning, and ensemble learning techniques to classify environmental waste. ResNet50, InceptionResNet-V2, and DenseNet169 models were used, and these models were fine-tuned using pre-trained weights. We created an ensemble model by combining the feature maps obtained from each model. Among the features extracted by the ensemble deep learning model, the most effective features were determined with ANOVA, Variance Threshold, Mutual Information, Random Forests, Lasso, RFE, PCA, and Ridge Regression feature selection methods. The selected features were classified with SVM, MLP and Random Forest, XGBoost, hard voting, and soft voting methods. This study presents the contributions of both individual and ensemble models for environmental waste classification. The effectiveness of the proposed method was tested on two different datasets, and its effectiveness was verified. The results show that the proposed method can make a significant contribution to waste management and recycling processes.

Keywords: Waste classification, ensemble learning, machine learning, deep learning, feature selection.

ÖZET

Artan atık üretimi ve yetersiz atık yönetimi, küresel çevre sorunlarını daha da karmaşık hale gelmiştir. Doğal kaynakların sınırlılığı ve atıkların çevreye verdiği zararlar atık yönetim sistemlerinin iyileştirilmesini zorunlu kılmaktadır. Atıkların doğru ve etkili sınıflandırılması hem ekonomik fayda sağlamakta hem de çevresel etkileri azaltmaktadır. Bu çalışmada, çevresel atıkların sınıflandırmak için derin öğrenme, makine öğrenmesi ve topluluk öğrenme knikleri birleştirilerek hibrit bir yaklaşım sunulmuştur. ResNet50, InceptionResNet-V2 ve DenseNet169 modelleri kullanılmış ve bu modeller önceden eğitilmiş ağırlıklar kullanılarak fine-tuning yapılmıştır. Her bir modelden elde edilen özellik haritaları birleştirilerek ensemble bir model oluşturulmuştur. Ensemble deep learning modeli tarafından çıkarılan öznitelikler arasından ANOVA, Variance Threshold, Mutual Information, Random Forests, Lasso, RFE, PCA ve Ridge Regresyon özellik seçim yöntemleri ile en etkili öznitelikler belirlenmiştir. Seçilen öznitelikler SVM, MLP ve Random Forest, XGBoost, çoğunluk oylama ve yumuşak oylama yöntemleriyle sınıflandırılmıştır. Bu çalışma çevresel atık sınıflandırmak için hem bireysel modellerin hem de ensemble modellerin sağladığı katkıları ortaya koymaktadır. Önerilen yöntemin etkinliği iki farklı veri kümesinde test edilmiş ve etkinliği doğrulanmıştır. Elde edilen sonuçlar, önerilen yöntemin atık yönetimi ve geri dönüşüm süreçlerine anlamlı bir katkı sağlayabileceğini göstermiştir.

ToCite: SALUR, M. U., & ELMAS, N., (2025). INNOVATIVE APPROACH TO ENVIRONMENTAL WASTE CLASSIFICATION WITH ENSEMBLE LEARNING MODELS. *Kahramanmaraş Sütçü İmam Üniversitesi Mühendislik Bilimleri Dergisi*, 28(2), 933-956.

Anahtar Kelimeler: Atık sınıflandırma, topluluk öğrenimi, makine öğrenmesi, derin öğrenme, öznitelik seçimi.

INTRODUCTION

In everyday life, a range of environmental waste is produced from the consumption of food and beverages by individuals. Annually, over 2.01 billion tons of solid waste are generated globally, with merely 19% of this quantity being recycled (Kaza et al., 2018). Providing waste management and developing waste recycling systems are important for environmental sustainability. One of the important obstacles to the recycling of environmental waste is the separation of waste according to its type. This process is often costly in terms of both time and labour. However, today, developments in machine learning and deep learning, especially their high capabilities in image processing, offer significant opportunities in classifying environmental waste. The utilization of deep learning and machine learning techniques in environmental waste management is progressively rising.

Recent improvements have been achieved in the investigation of artificial intelligence techniques for waste classification. Deep learning algorithms, including Convolutional Neural Networks (CNN), have been extensively used for visual data classification (Poudel & Poudyal, 2022). However, individual models often exhibit high performance on specific types of data but are generally insufficient for overarching classification tasks (Meyen et al., 2021). To address this issue, ensemble learning methods combine the outputs of various models, providing greater accuracy and precise predictions (Şirin, E. (2017). The challenges encountered in the classification of environmental arise from various factors, including the lack of diversity and size in datasets, imbalances in data distribution, and the complexity of waste types (Shahab et al., 2022)(Abdu & Noor, 2022). DL models offer significant potential in overcoming these challenges by extracting meaningful features from waste images (Ma et al., 2023). Moreover, the use of transfer learning and data augmentation techniques has mitigated the problem of data scarcity (Hutchinson et al., 2017). However, dataset imbalance issues are more effectively addressed through ensemble learning techniques (W. Chen et al., 2024). In addition to ensemble learning, hyperparameter optimization plays a critical role in enhancing the performance of DL models. Zheng and Gu developed the EnCNN-UPMWS model, which implemented a different strategy for classifying household solid waste and attained 93% accuracy on the TrashNet dataset, outperforming other individual models (Zheng & Gu, 2021). Similarly, Single et al. evaluated the performance of DL models in real-world waste classification challenges and achieved 89% accuracy using the Inception V3 model (Single et al., 2023).

Ensemble DL and ensemble machine learning (ML) approaches achieve better classification performance compared to individual models. However, the absence of a model in the literature that integrates both approaches has served as the motivation for this study. In this research, a hybrid model based on ensemble DL and ML is proposed for the classification of environmental waste. In the proposed model, ResNet50, InceptionResNet-V2, and DenseNet169 DL models were fine-tuned and ensembled to extract features from environmental waste images. Subsequently, various feature selection methods, including ANOVA, Variance Threshold (VT), Mutual Information (MI), Random Forests (RF), Lasso, Recursive Feature Elimination (RFE), Principal Component Analysis (PCA), and Ridge Regression (RR), were applied to the features obtained from the ensemble model to identify high-quality features. The selected features were then used as input for Support Vector Machines (SVM), Multi-Layer Perceptron (MLP), Random Forest (RF), XGBoost, hard voting, and soft voting models. In the proposed approach, ensemble methods were employed in both the feature extraction and classification stages. The performance of the proposed model was tested on two different datasets. The key contributions of this study are presented as follows:

- A novel end-to-end hybrid classification model is proposed using an ensemble approach for environmental trash classification, integrating the advantages of both ML and DL.
- Fine-tuning was performed on ResNet50, InceptionResNet-V2, and DenseNet169 models, and an ensemble DL model was constructed by combining these three models.
- Multiple feature selection techniques, including as ANOVA, VT, MI, RF, Lasso, RFE, PCA, and RR, were used to choose the most important features. The effectiveness of these feature selection methods was evaluated.

- Hyperparameter optimization was performed for SVM, MLP, and RF ML algorithms. Environmental waste classification was conducted using XGBoost, hard voting, and soft voting ensemble ML models.

The second section of the paper presents the literature review, the third section details the methodology, the fourth section discusses the experimental results and analysis, and the final section provides the overall conclusions of the study.

LITERATURE REVIEW

The classification of environmental waste and the improvement of recycling processes are critically important for environmental sustainability. Ouedraogo et al. combined transfer learning and ensemble learning to process collected waste images and classify landfill waste into nine categories using InceptionResNet-V2, EfficientNetB3, and DenseNet201 models. While the InceptionResNet-V2, EfficientNetB3, and DenseNet201 models achieved accuracy rates of 86%, 87%, and 88%, respectively, the ensemble model outperformed them with the highest accuracy of 90% (Ouedraogo et al., 2023).

Salur et al. designed a smart waste container prototype using Raspberry Pi. The proposed system performs the tasks of detecting the discarded waste, capturing the image of the discarded waste, classifying the image with DL, and moving the waste to the category to which it belongs via a stepper motor. In the system, DL models. ResNet50, DenseNet201, DenseNet169, Inception-V3, and VGG16 are used to determine the category of the waste. The highest classification accuracy was achieved with the DenseNet169 model at 95.22%. The proposed system has the potential to reduce manual labor and optimize recycling processes (Salur et al., 2024).

Al Duhayyim et al. implemented hazardous and non-hazardous waste detection and classification using the HWDC-EL (Hybrid Weighted Deep Classification - Ensemble Learning) technique. Additionally, the Flower Pollination Algorithm (FPA) was utilized to optimize the hyperparameters of EfficientNet and DenseNet121 models. A weighted voting-based ensemble classifier was constructed using three ML algorithms: Support Vector Machines (SVM), Extreme Learning Machine (ELM), and Gradient Boosting Tree (GBT). The HWDC-EL classifier achieved the highest classification accuracy of 98% (Al Duhayyim et al., 2023).

Zheng and Gu proposed a novel ensemble learning model, EnCNN-UPMWS, based on CNNs and an Unequal Precision Measurement Weighting Strategy (UPMWS). In the proposed method, various deep learning models, including ResNet50, MobileNetV2, and GoogLeNet, are employed as classifiers. The UPMWS strategy was applied to determine the weight coefficients of the ensemble models. The proposed model is tested on the TrashNet dataset. The EnCNN-UPMWS model achieved a 93% accuracy, outperforming other DL models (Zheng & Gu, 2021).

Yıldız et al. utilized AlexNet, GoogLeNet, ResNet50, DenseNet201, ShuffleNet, and SqueezeNet architectures for environmental waste classification. Feature maps were extracted using the DenseNet201 architecture, and these features were fed into Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Decision Trees (DT) algorithms. The combination of the DenseNet201 architecture with SVM, DT, and KNN classifiers resulted in classification accuracies of 89%, 66%, and 83%, respectively (Yıldız et al., 2023).

Single et al. aimed to provide a more accurate and scalable alternative to traditional waste classification methods (such as visual inspection, weight measurement, and manual sorting) by utilizing various CNN models, including VGG-16, InceptionResNetV2, DenseNet121, Inception V3, and MobileNetV2. The study demonstrated the practical applications of DL in waste classification by using real-world examples in the newly developed RealWaste dataset. Among the tested models, Inception V3 achieved the highest classification performance with an accuracy of 89.19% (Single et al., 2023). Similarly, Younis and Obaid conducted a performance comparison of five commonly used CNN models (EfficientNet, VGG-19, GoogLeNet, ResNet152, and ShuffleNet) on the RealWaste dataset. Their results showed that EfficientNet achieved the highest test accuracy of 96.31%, significantly outperforming the other models (Younis & Obaid, 2024).

Toğaçar et al. reconstructed the waste classification dataset using an AutoEncoder. Features were extracted from the dataset using AlexNet, GoogLeNet, and ResNet50 models, and these feature sets were combined. The Ridge Regression (RR) method was applied to the merged feature set to reduce the number of features. After feature selection, the highest accuracy of 99% was achieved using the SVM classifier (Toğaçar et al., 2020).

Chatterjee et al. investigated the effectiveness of transfer learning and ensemble learning approaches in the automated classification of transparent plastic bottles into well-mannered and poorly-mannered categories. InceptionV3, Xception, ResNet152, and DenseNet169 architectures were used as baseline models for classification. They developed IncepX-Ensemble, an ensemble model based on InceptionV3 and Xception. To address the dataset imbalance problem, data augmentation techniques were applied. As a result, the proposed method achieved 99% accuracy in classifying transparent plastic bottles (Chatterjee et al., 2022).

Santoso et al. have proposed a hybrid approach that combines transfer learning and dimensionality reduction methods to achieve high accuracy in the classification of solid waste. They used CNN and EfficientNet models for feature extraction and PCA for dimensionality reduction. The features extracted from CNN and EfficientNet were combined and sent to the classifier layer. The joint evaluation of both local and global visual features has enhanced the model's success (Santoso et al., 2024)

Yulita et al. performed feature extraction using the Inception V3 model for waste classification. The extracted features were input into various ensemble learning models (XGBoost, AdaBoost, and RF) and classified. XGBoost has shown higher performance than all other algorithms, and it has performed well in cases of class imbalance.

Wu et al. proposed a two-level ensemble CNN model that takes images with different colors and features as input to classify environmental waste. The outputs of two different CNN models operating at the first level are transferred to a third CNN at the second level, not only as class probabilities but also along with the feature maps obtained from the intermediate layers. The CNN at the second level learns multidimensional information to understand which model is more reliable in which situations and classifies complex patterns more accurately.

In the literature on environmental waste classification, studies have primarily focused on transfer learning, ensemble learning, feature extraction and optimization techniques, data augmentation, addressing data imbalance issues, and the combination of ML and DL approaches. Compared to existing studies, the novelty of the proposed model in this research lies in its end-to-end integration of multiple techniques, including fine-tuning DL models, constructing an ensemble model by combining different architectures, applying various feature extraction methods, performing hyperparameter optimization for ML algorithms, and implementing ensemble ML within a single framework.

METHODOLOGY

This study proposes an end-to-end model that utilizes ensemble learning for environmental waste classification. The proposed model has been applied to two different datasets. It consists of three main stages: feature extraction using ensemble DL, feature selection, and classification using ensemble ML. The overall diagram of the proposed model is presented in Figure 1. We propose an end-to-end hybrid model based on ensemble DL and ensemble ML for the classification of environmental waste. The core idea of this integrated approach is the assumed synergy between robust feature extraction and robust classification. Ensemble learning models provide higher performance than individual classifiers since they combine the power of various classifiers. In this direction, in our proposed method, we adopt the ensemble learning approach both in the feature extraction phase with DL models and in the ML classification phase. However, the outputs of the deep learning models have formed high-dimensional feature sets. In this case, we made improvements through feature selection to retain only the most discriminative information. Then, we reasoned that applying ensemble ML techniques (such as soft voting using SVM, MLP, and RF as base models) to these selected features would provide a more reliable and accurate final classification than relying on a single ML model or the direct output of the ensemble DL stage. In fact, this two-stage ensemble strategy aims to leverage the complementary strengths of deep feature learning and ML classification paradigms.

In the proposed method for environmental waste classification, the datasets were initially randomly split into 80% train and 20% test. During the transfer learning, ResNet50, InceptionResNet-V2, and DenseNet169 models, pre-trained on ImageNet, were utilized. Global Average Pooling, Dense, and Dropout layers were added to each model for fine-tuning. The outputs of these models were then combined to generate the main model feature vector. This combined feature vector was further processed through batch normalization, dense, and dropout layers to construct the feature set. Various feature selection methods, including ANOVA, VT, MI, RF, Lasso, RFE, PCA, and RR, were applied to select between 5 and 200 features. After feature selection, the selected features were fed into SVM, MLP, and RF algorithms, optimized using the GridSearchCV method, and their performances were recorded. In addition to ML models, the performance of ensemble methods, including XGBoost, hard voting, and soft voting, was evaluated on the selected feature sets. The proposed model integrates ensemble DL for feature extraction and ensemble ML for classification. The performance of the proposed model was validated on two different datasets.

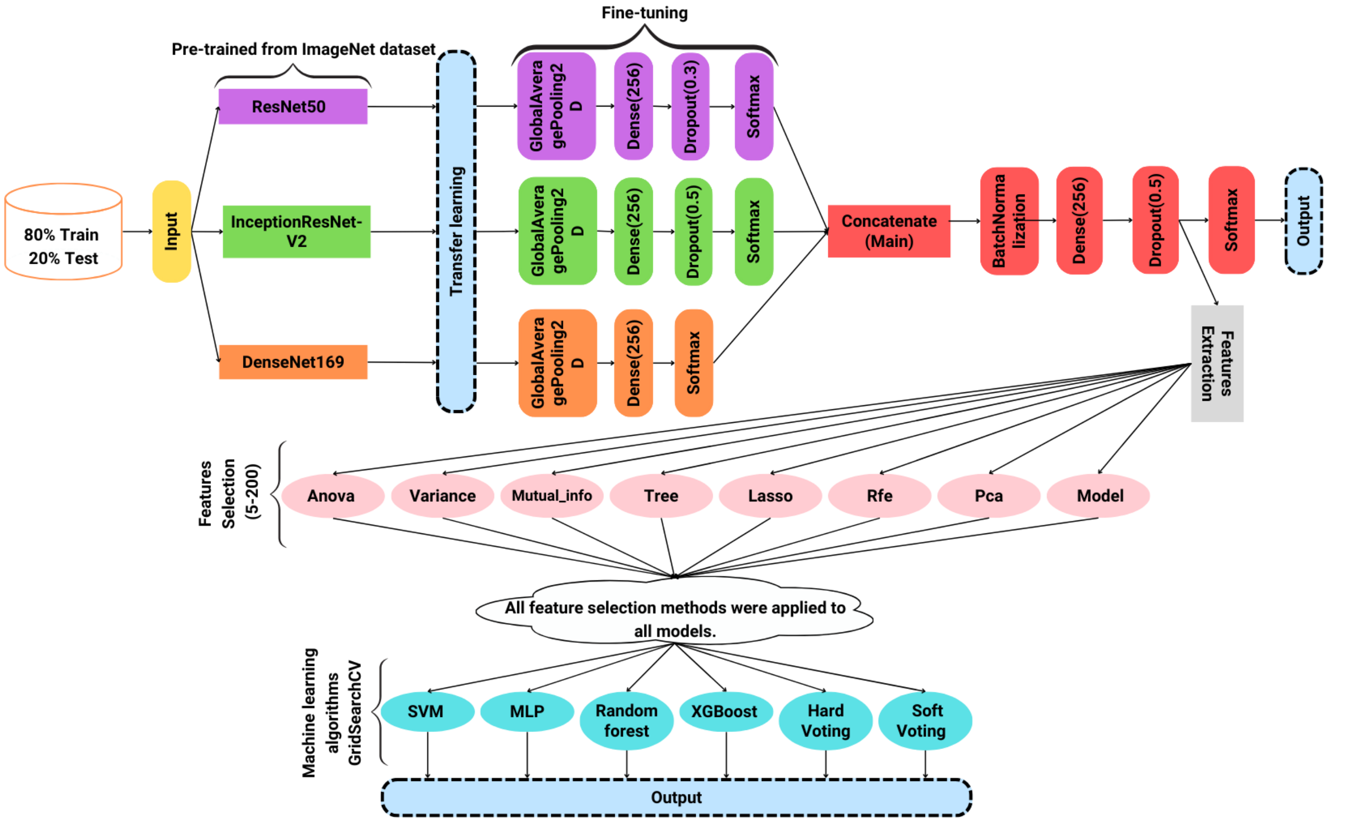

Figure 1. Diagram of the Proposed End-to-End Ensemble Model

Datasets

There are some open access shared datasets for the classification of environmental waste. In order to test the classification performance of the proposed model, two datasets frequently used in the literature were used. This section presents the characteristics of these datasets and their sample distributions.

Dataset-1

The Garbage Classification dataset was utilized to test the performance and analyze the generalization potential of the proposed model. The Garbage Classification dataset is widely preferred in the literature and was published in 2021 (*Garbage Classification*, 2021). The reliability of the dataset is ensured by its availability on the Kaggle platform and the assumption that it is correctly labeled. The dataset consists of a total of 2527 images and includes six classes: cardboard, glass, metal, paper, plastic, and trash. The distribution of images used for training and testing in each class is presented in Figure 2. The dataset is imbalanced in terms of sample distribution. Since ensemble learning can effectively handle class imbalance problems, the performance of the proposed model on this dataset is expected to be more significant compared to a dataset with a balanced distribution.

Dataset-2

To further evaluate the efficacy of the proposed model, the RealWaste dataset was utilized as the secondary dataset (Single et al., 2023). The RealWaste dataset is relatively new in the literature and comprises real-world waste samples. It consists of a total of 4752 images categorized into nine classes: cardboard, food organics, glass, metal, miscellaneous trash, paper, plastic, textile trash, and vegetation. The number of images used for training and testing in each class is presented in Figure 3. Similar to Dataset-1, the RealWaste dataset also exhibits an imbalanced distribution.

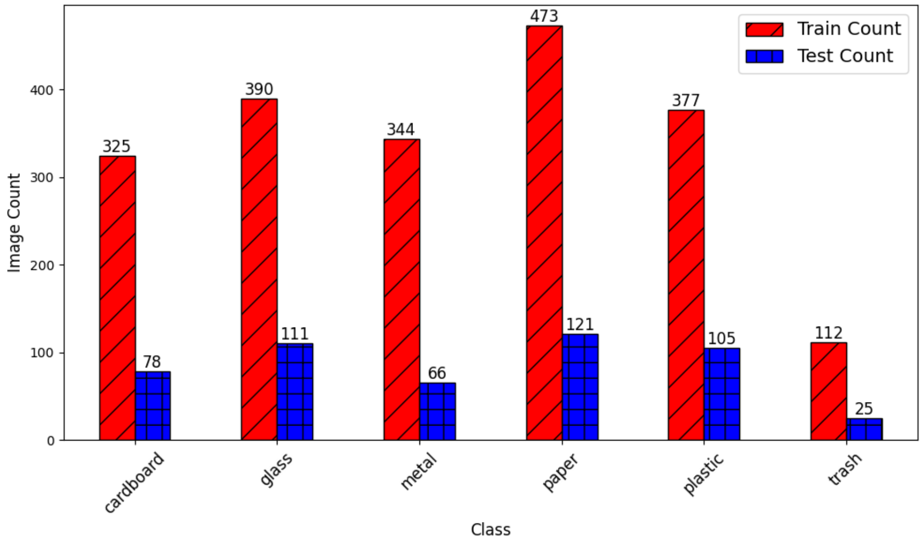

Figure 2. Class Distribution of Dataset-1

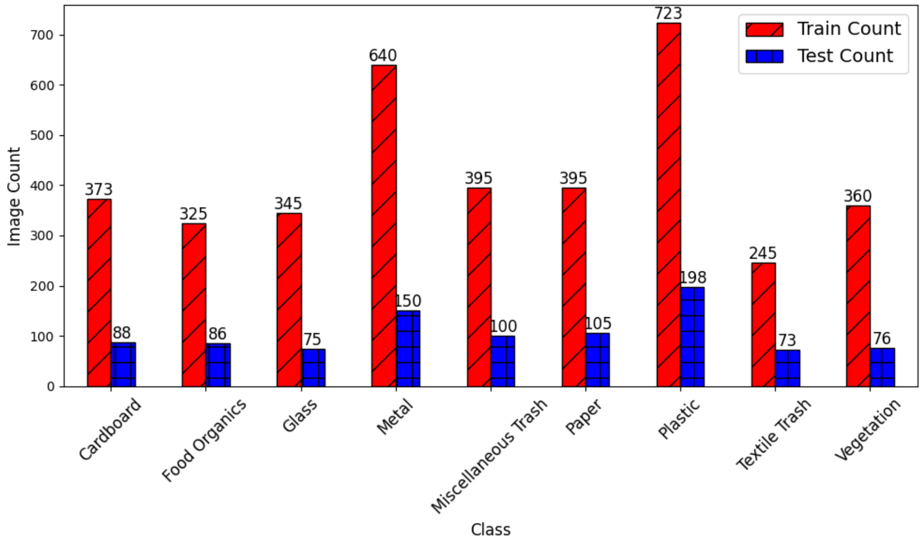

Figure 3. Class Distribution of Dataset-2

Deep Learning Classification Models

For the classification of environmental waste, fine-tuning and parameter optimization were performed to adapt the pre-trained ResNet50, InceptionResNet-V2, and DenseNet169 models to the datasets. Various experiments were conducted by modifying the number of trainable layers and the number of neurons in the dense layers added at the end of the models. Through fine-tuning, an effort was made to optimize computational cost efficiency. Additionally, ResNet50, InceptionResNet-V2, and DenseNet169 models were combined to construct an ensemble model. The details of the ResNet50, InceptionResNet-V2, DenseNet169, and the ensemble model used in this study are presented in Table 1.

Despite the availability of more recent models, we preferred ResNet50, InceptionResNet-V2, and DenseNet169 models owing to their established efficacy in image classification tasks, including waste classification in existing literature. For our objective of creating an ensemble model, these models have represented distinct and successful architectural paradigms (namely, residual connections, inception modules combined with residual connections, and dense connections). This variety in design has been intentionally chosen to pick up different useful features from the input images, which will improve the effectiveness of the next deep learning stage.

Table 1. Fine-tuning Parameters of Deep Learning Models

Model	Reference	Highlight	Layer freeze	Added layers
ResNet50	(He et al., 2016)	ResNet50 is a variant of the ResNet architecture consisting of 50 layers. It utilizes residual connections to address the vanishing gradient problem.	The first 143 layers were frozen, while the remaining layers were trained.	{GlobalAveragePooling2D, Dense (256 unit), Dropout (0.3), Dense (Softmax)}
InceptionResNet-V2	(Szegedy et al., 2017)	InceptionResNet-V2 integrates Inception modules with residual connections, reducing computational cost while achieving higher accuracy.	The last 30 layers were unfrozen, while the preceding layers were kept frozen.	{GlobalAveragePooling2D, Dense (256 unit), Dropout (0.5), Dense (Softmax)}
DenseNet169	(Huang et al., 2017)	By transferring the output of each layer to subsequent layers, it ensures efficient utilization of parameters. DenseNet169 is a 169-layer variant of DenseNet, allowing deep features to be effectively used for classification tasks.	The first 120 layers were frozen, while the remaining layers were trained.	GlobalAveragePooling2D, Dense (256 unit), Dense (Softmax)

The feature vectors (output tensors) extracted from the three models (ResNet50, InceptionResNet-V2, and DenseNet169) were merged using a concatenation layer. The goal was to leverage the diverse feature extraction capabilities of each architecture to build a more robust model. Features obtained from the intermediate or final layers of the models were combined using the Concatenation method, and these combined features were fed into the ensemble classifier. The effectiveness of ensemble methods is explained by their ability to balance individual model errors and enhance overall performance (Dietterich, 2000). The concatenated feature vectors were passed through BatchNormalization, Dense(256), Dropout(0.5), and Dense (Softmax) layers for classification. During the training phase of the ensemble model, the Adam optimizer, a low learning rate (0.00001), the categorical cross-entropy loss function, and accuracy as the performance metric were utilized. Additionally, EarlyStopping was applied during training to prevent the model from overfitting. The validation loss metric was monitored as the EarlyStopping criterion. It has been determined that training will be stopped if there is no improvement in the validation loss value for 5 consecutive times. Additionally, by using the 'restore_best_weights=True' parameter in the model, it was ensured that the model returns to the weights with the lowest validation loss. Thus, the model was ensured to stop at the optimum point and exhibit the best overall performance.

Ensemble models have higher computational costs than individual models. The individual DL models used in this study and the trainable parameter numbers of the ensemble DL model proposed for feature extraction from environmental wastes are presented in Table 2. The proposed ensemble model has approximately 82 M trainable parameters.

Table 2. Trainable Parameters of Deep Learning Models

Model	Trainable params
ResNet50	15.502.857
InceptionResNet-V2	54.671.977
DenseNet169	11.847.241
Main Ensemble DL	82.222.788

Feature vectors were extracted for each of the ResNet50, InceptionResNet-V2, DenseNet169, and Ensemble models. Separate feature vectors were generated for both train and test data for each model and saved in DataFrame format. These feature vectors were then stored as CSV files for use in feature selection and ML algorithms. The combination of different features learned from the three models enhanced the model's generalization capability and improved classification performance.

Feature selection

Feature selection (FS) is a critical preprocessing step in ML processes, aiming to enhance model performance by selecting the most important features from the dataset. This approach helps prevent overfitting and improves the interpretability of the model. In the proposed model, we performed classification experiments on 8 different feature selection methods. We were to choose this set of methods to cover a variety of established techniques representing different FS approaches (filter, wrapper, embedding, and dimensionality reduction). Our motivation for applying feature selection methods with different characteristics comes from the ensemble use of deep learning models. Because in ensemble deep learning models, different models generate different levels of semantic representations of the same input image. In this case, some features may be more effective than others in the process of combining features. In this study, we observe this situation and analyze the performance of multiple feature selection methods. The methods used for FS in this study are as follows:

- **ANOVA**: ANOVA performs feature selection by statistically testing the relationships between independent variables and target class information. In this method, the variance ratio between groups is calculated using the F-statistic, and then it determines the variables that represent the differences between classes at a significant level (*Filtreleme Yöntemleri · Miuul*, 2021). This method uses techniques such as SelectKBest, which are frequently used in variable selection processes. SelectKBest evaluates each variable based on a specific statistical test and selects the k variable with the highest score. As a result, the selection of the most distinctive features that can contribute to the performance of the model is provided.
- **Variance Threshold (VT)**: This method relates feature selection to the variance of the feature. This method ensures that features with low variance are removed from the dataset. By setting a variance threshold, irrelevant or redundant features are removed from the dataset (Guyon & Elisseeff, 2003).
- **Mutual Information (MI):** This method is used in decision tree algorithms to measure the dependency between variables in determining the nodes. It is also used to determine the relationship between features and the target class variable in FS for a similar purpose. Features with higher mutual information scores are selected as they provide more information about the target variable (*Vergara al., 2014)*.
- **Random Forest (RF):** RF is a powerful ensemble learning model that uses decision trees. Calculates feature importance scores by evaluating each node split using methods such as the Gini index or entropy. Features that reduce impurity the most are considered more important (Genuer et al., 2020).
- **Lasso:** Performs direct feature selection by shrinking the coefficients of less important features towards zero. This approach, which applies L1 regularization, is an effective solution for datasets with excessive irrelevant or redundant features (Tibshirani, 1996).
- **Recursive Feature Elimination (RFE):** RFE is a computationally expensive method that tries to find the most valuable features by starting from the entire feature set and eliminating the less important features from the dataset after each training for feature selection. (Guyon et al., 2002).
- **Principal Component Analysis (PCA):** PCA is a dimensionality reduction technique that transforms data into a new coordinate system, where each axis represents a principal component. Since the features extracted from the DL model are generally in high dimensions, this method is effective in dimension reduction in DL models. PCA selects the components that explain the maximum variance in the dataset (Greenacre et al., 2022).
- **Ridge Regression (RR)**: RR ranks each feature's influence on the target variable and chooses the most significant ones. In order to prevent overfitting during feature selection and train the model appropriately, ridge regression is applied. The model ranks the value of every feature, chooses the top k most significant ones, and reorganizes the train and test sets based on the features that were selected (McDonald, 2009).

Machine learning models

In this study, Support Vector Machines (SVM), Multi-Layer Perceptron (MLP), and Random Forest (RF) were used as individual models for environmental waste classification, while XGBoost, hard voting, and soft voting were employed as ensemble methods. For the hard voting and soft voting ensemble classifiers, SVM, MLP, and RF were used as base models.

- **SVM:** A powerful method used for both linear and nonlinear classification problems. It constructs hyperplanes to classify data and aims to maximize the margin between two classes. With the use of kernel functions (such as RBF and polynomial), it can also achieve effective results on nonlinear data (Kanevski et al., 2002).

- **MLP:** A feedforward artificial neural network commonly used to learn complex relationships between data. It consists of an input layer, hidden layers, and an output layer. Learning is performed through the backpropagation algorithm (Pinkus, 1999).
- **RF:** An ensemble method formed by combining multiple decision trees. The final prediction is made by voting or averaging the outputs of each tree. It is particularly robust against noisy data and helps reduce the risk of overfitting (Genuer et al., 2020).
- **XGBoost:** An improved version of the gradient boosting algorithm. It is known for its fast execution and ability to achieve effective results in large datasets due to its regularization techniques. It is widely used in classification and regression problems (T. Chen & Guestrin, 2016).
- **Soft Voting:** Combines the probability values of predictions from multiple models and selects the class with the highest average probability. This method integrates the strengths of different models, leading to more balanced results (Dietterich, 2000).
- **Hard Voting:** Uses the class labels predicted by each model and determines the final class based on majority voting (Dietterich, 2000). An odd number of models is used, and the class predicted by the majority is considered the true class.

ML algorithms, while powerful for solving complex classification problems such as environmental waste classification, each have their individual advantages and limitations. In this study, GridSearchCV was used to select the parameters that best fit the algorithms to the datasets to obtain maximum performance from each algorithm. The highest performance of each algorithm for Dataset-1 and Dataset-2 was achieved with the parameters given in Table 3.

Table 3. Machine Learning Models Discovered Using GridSearchCV

	Hyperparameters	
Machine learning models	*Dataset-1 (Gc)*	*Datset-2 (Real weste)*
SVM	{'C': 10, 'gamma': 0.01, 'kernel': 'rbf'}	{'C': 10, 'gamma': 0.01, 'kernel': 'rbf'}
RF	{'criterion': 'gini', 'max_depth': None, 'n_estimators': 50}	{'criterion': 'entropy', 'max_depth': 10, 'n_estimators': 100}
MLP	{'hidden_layer_sizes': (100, 50), 'learning_rate_init': 0.001, 'max_iter': 500}	{'hidden_layer_sizes': (50, 50), 'learning_rate_init': 0.001, 'max_iter': 500}
XGBoost	{'learning_rate': 0.2, 'max_depth': 7, 'n_estimators': 100}	{'learning_rate': 0.2, 'max_depth': 7, 'n_estimators': 100}

Implementation Details

During the model development process, the TensorFlow library was used, and models were implemented using the Keras API. Pre-trained DL models such as ResNet50, DenseNet169, and InceptionResNet-V2 were fine-tuned using ImageNet weights. Image augmentation was applied using ImageDataGenerator to prepare the image data. During the training process, the Adam optimization algorithm and EarlyStopping were employed to control the learning rate and prevent overfitting. For feature extraction, intermediate layer outputs of ResNet50, DenseNet169, and InceptionResNet-V2 models were used to generate separate feature vectors, which were later combined. In this process, libraries such as NumPy and Pandas were utilized for data manipulation.

The performance of different ML models, including SVM, KNN, and RF, was analyzed. Additionally, VotingClassifier was used with both hard voting and soft voting to combine the outputs of multiple models. At the end of the training process, the final models and extracted feature vectors were executed in the Google Colab environment, and the data was stored on Google Drive.

Performance Metrics

The performance evaluation of the classification methods for environmental waste images in both datasets was conducted using well-established metrics from the literature. The confusion matrix is presented in Figure 4. The performance of the models was assessed using Accuracy (Equation 1), Precision (Equation 2), Recall (Equation 3), F-score (Equation 4), and ROC-AUC (Equations 5 and 6).

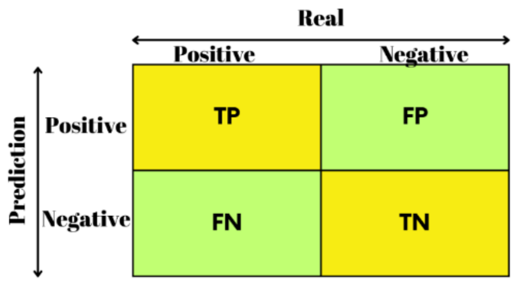

Figure 4. Confusion Matrix

$$\text{Accuracy} = (\text{TP} + \text{TN})/(\text{TP} + \text{TN} + \text{FP} + \text{FN}) \quad (1)$$

$$\text{Recall} = \text{TP}/(\text{TP} + \text{FN}) \quad (2)$$

$$\text{Precision} = \text{TP}/(\text{TP} + \text{FP}) \quad (3)$$

$$\text{F} - \text{score} = 2 * ((\text{Precision} * \text{Recall})/(\text{Precision} + \text{Recall})) \quad (4)$$

$$\text{False Positive Rate (FPR)} = \text{FP}/\text{FP} + \text{TN} \quad (5)$$

$$\{\text{ROC} - \text{AUC}\} = \int_0^1 \{\text{Recall(FPR)}\}, \text{d}(\{\text{FPR}\}) \quad (6)$$

EXPERIMENTAL RESULTS AND DISCUSSION

This section presents the classification performance of the proposed end-to-end ensemble model on Dataset-1 and Dataset-2.

Classification Results of the Dataset-1

The confusion matrix for ResNet50, InceptionResNet-V2, DenseNet169, and the ensemble DL models for Dataset-1 is presented in Figure 5. A detailed analysis of the matrices in Figure 5 reveals that the proposed ensemble DL method achieves higher accuracy than the other individual DL models. The ensemble DL model misclassifies only 28 out of 506 images. Moreover, InceptionResNet-V2 achieved the second-best results following the ensemble DL method. On the other hand, InceptionResNet-V2 correctly predicted 464 waste images, while 42 images were misclassified. When the confusion matrices for each waste class are examined, it is seen that the ensemble DL model exhibits better classification performance than individual models for cardboard, metal, paper, and plastic classes. However, it is seen that the InceptionResNet-V2 model is better than the ensemble DL model for the glass class.

The test results of ResNet50, InceptionResNet-V2, DenseNet169, and the ensemble DL models in terms of Accuracy, Precision, Recall, F-score, and ROC-AUC metrics are presented in Figure 6. While the performance metrics for ResNet50, InceptionResNet-V2, and DenseNet169 models exhibit similar trends, the ensemble DL model shows a 5.05% increase in accuracy.

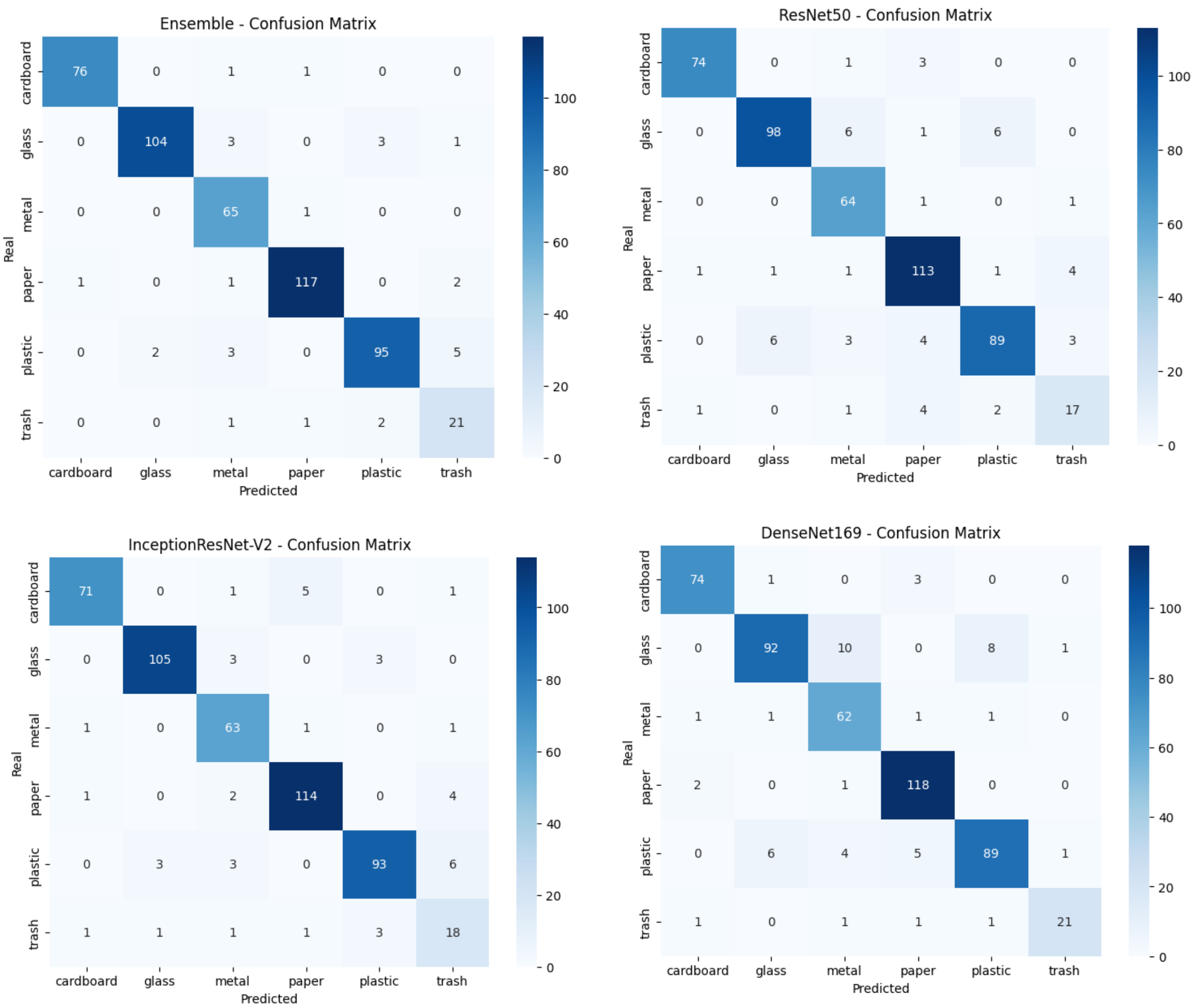

Figure 5. Confusion Matrices of ResNet50, InceptionResNet-V2, DenseNet169 and Ensemble DL Models in Dataset-1

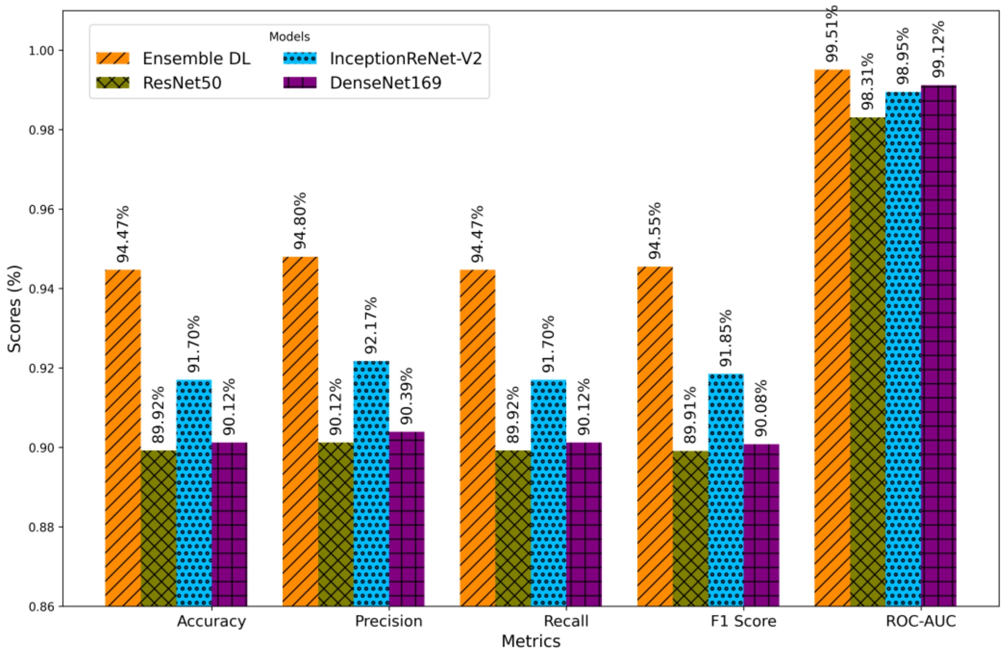

Figure 6. Performance Metrics of ResNet50, InceptionResNet-V2, DenseNet169, and the Ensemble DL Models in Dataset-1

Figure 7 illustrates the accuracy and loss of the models throughout the training and validation. Upon analyzing the curves, the train accuracy curve reveals that the accuracy rates of all models progressively rise, attaining elevated accuracy after 50 epochs. In the validation accuracy curve, the models' validation performance follows a trend parallel to their training performance. Observing the loss curve, both train loss and validation loss decrease as the number of epochs increases, demonstrating that the model optimization process progresses effectively. Notably, the ensemble DL model outperforms the other models by achieving lower validation loss and higher accuracy, indicating the best performance among them. These results confirm that the models were successfully trained and did not encounter issues such as overfitting during the validation phase. Furthermore, the superior performance of the ensemble DL model compared to the other models suggests that it generalizes better to the dataset.

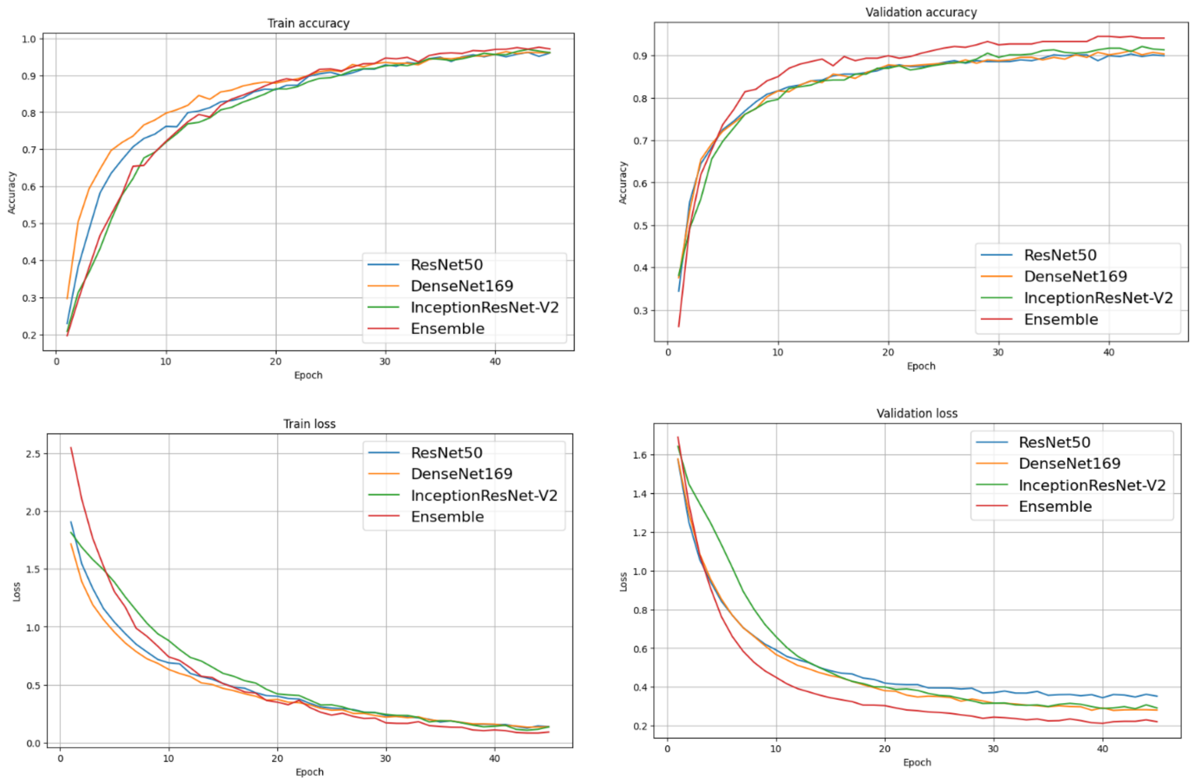

Figure 7. Accuracy and Loss of DL and Ensemble Models During the Train and Test in Dataset-1

The ML approach involved the installation of requisite libraries and the utilization of features derived from the ensemble DL. The labels in the dataset were converted into numerical values for classification algorithms, and the data was scaled using StandardScaler. During the feature selection phase, different feature subsets were created by applying ANOVA, VT, MI, RF, Lasso, RFE, PCA, and RR methods, increasing the number of selected features in increments of five between 5 and 200. Each of these feature subsets was classified using SVM, MLP, RF, and other classification algorithms. For ensemble ML models, SVM, MLP, and RF were used as base classifiers. Additionally, ensemble ML models were developed using XGBoost, hard voting, and soft voting methods. The classification results of the models were comprehensively evaluated using performance metrics such as accuracy, precision, recall, f-score, and ROC-AUC, along with confusion matrices. The classification results were analyzed to compare the performance of feature selection methods and algorithms and to determine the most suitable feature selection method and feature count combinations. Figure 8 illustrates that the Lasso method, utilizing 145 selected features alongside the soft voting ensemble model, attained the highest performance, evidenced by an accuracy of 95.26%. Among the reasons for the superior performance of the Lasso method with 145 features in the Soft Voting classifier is the L1 property, which performs feature selection by forcing the coefficients of less informative features to zero. Therefore, it is understood that Lasso's tendency to create sparse models with L1 regularization is better suited to the feature space of Dataset-1 compared to other methods.

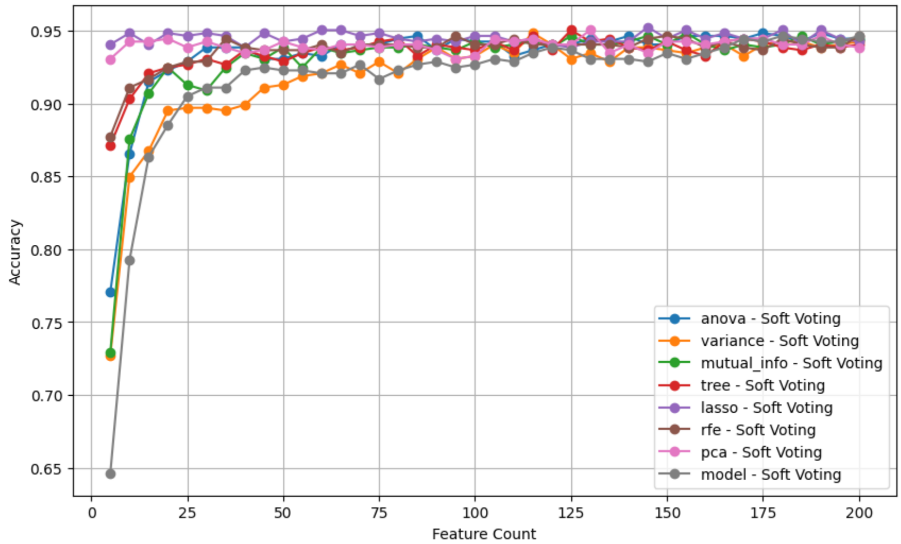

Figure 8. Accuracy Value of the Soft Voting Classifier in Dataset-1 Based on the Feature Selection Method and the Number of Selected Features

The results of accuracy, precision, recall, F1-score, and ROC-AUC metrics for ML and ensemble ML models are presented in Figure 9. The soft voting model demonstrated superior or comparable performance across all metrics, attaining a precision of 95.35%, recall of 95.26%, F1-score of 95.28%, and ROC-AUC of 99.21%. In the classification accuracy metric, both soft voting and hard voting outperformed the other individual models. Among the individual models, the RF model exhibited remarkable performance in the ROC-AUC metric, while the SVM and MLP models provided generally balanced results. However, the XGBoost model showed relatively lower classification performance metrics compared to the others. Considering that Dataset-1 has an unbalanced class distribution, it is obvious that the F1-measure will be important in model performance. The highest F1 score among the six models is 95.28% and 94.72% in soft voting and hard voting models, respectively.

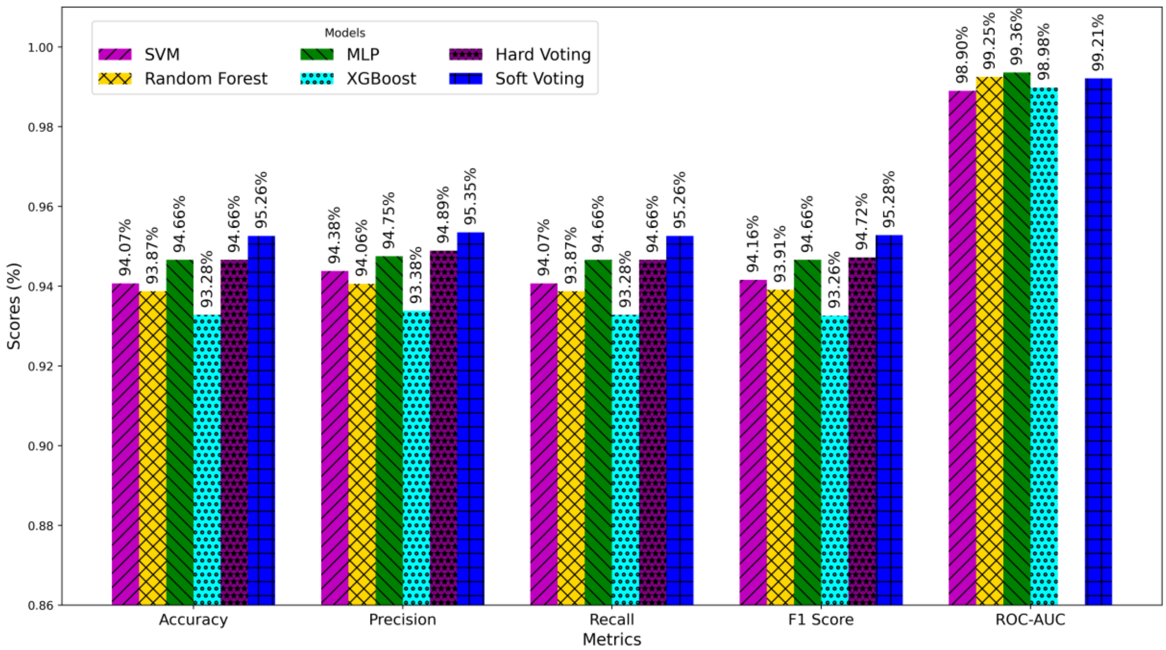

Figure 9. Performance Metrics of ML Models and Ensemble ML Models in Dataset-1

The confusion matrix of the soft voting ensemble ML model (Lasso-145 features) is presented in Figure 10. The examination of the matrix reveals that the proposed ensemble method achieved high classification accuracy. While 24 of the 506 images in Dataset-1 were misclassified, the remaining 492 images were correctly classified. When compared to the DL ensemble method, it was observed that the soft voting ensemble method correctly classified 4 more images. Class-wise analysis indicates that the best classification performance was achieved in the paper class, while the highest misclassification occurred in the trash class.

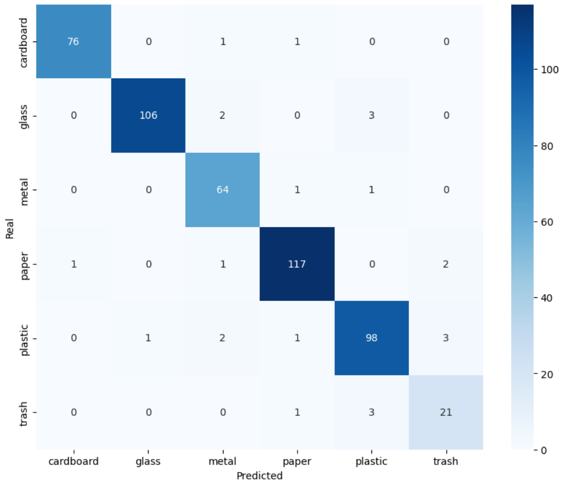

Figure 10. Confusion Matrix of the Soft Voting Model with Lasso-145 Features in Dataset-1

A heatmap summarizing the precision, recall, and F1-score performance metrics for different classes in the soft voting ensemble ML model is presented in Figure 11. The cardboard, glass, metal, paper, and plastic classes have metric values above 90%, indicating that the model has demonstrated a strong ability to distinguish these classes. This indicated that the model has effectively learned the characteristics of these waste types and can accurately classify them. However, the trash class stands out as the lowest-performing class, with precision of 81%, recall of 84%, and F1-score of 82%. These results indicate that the model makes more misclassifications in terms of false positives or false negatives for the trash class. The feature similarity of this class with other classes leads to increased misclassification.

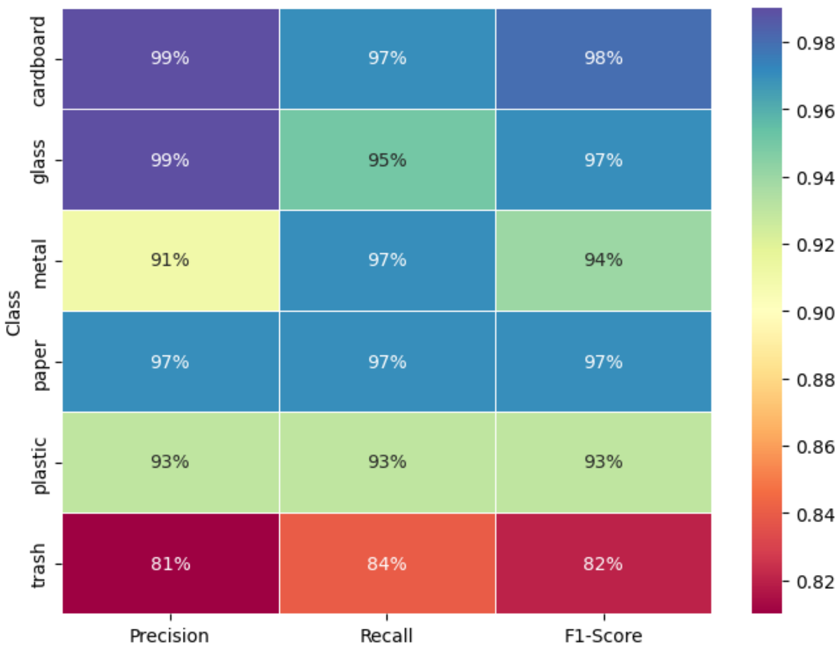

Figure 11. Heatmap of the Soft Voting Model with Lasso-145 Features in Dataset-1

Figure 12 demonstrates the performance metrics of the ensemble DL and ensemble soft voting ML models. The results in Figure 12 indicate that the soft voting model outperforms the ensemble DL model in terms of all classification metrics. The soft voting model achieved 95.26% accuracy, 95.35% precision, 95.26% recall, 95.28% F1-score, and 99.21% ROC-AUC, demonstrating its superior performance. On the other hand, the ensemble DL model attained 94.47% accuracy, 94.80% precision, 94.47% recall, 94.55% F1-score, and 99.51% ROC-AUC. The

ROC-AUC value of 99.51% in the ensemble DL model was found to be slightly higher than that of the ensemble ML soft voting model. However, the fact that soft voting achieved better results in accuracy, precision, recall, and F1-score highlights its potential to accurately predict positive classes and enhance overall classification performance. These findings indicate that soft voting is a effective approach than ensemble DL in terms of classification. Moreover, the fact that the classification metrics (approximately 94% for ensemble DL and 95% for ensemble ML) are close to each other in both ensemble DL and ensemble ML models indicates that the classifiers are stable and have high generalization capacities. However, since the proposed method uses ensemble methods in both feature extraction and feature classification stages, the high computational load is a disadvantage. However, we believe that performance will be more valuable than computational cost for critical tasks such as environmental waste classification.

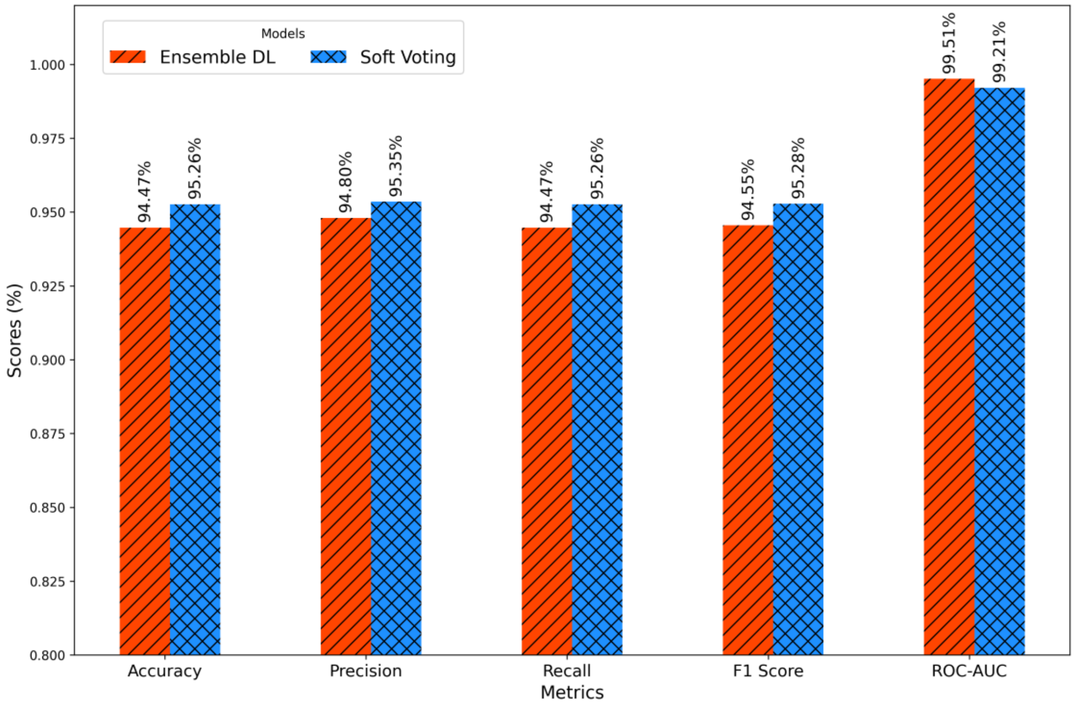

Figure 12. Comparison of the Performance Metrics of the Ensemble DL Model and the Soft Voting Model in Dataset-1

Classification results of the Dataset-2

Dataset-2 is a more comprehensive dataset than Dataset-1, both in terms of classes and instances. Datasets with different characteristics were specifically selected to test the performance of the proposed model. Figure 13 presents the confusion matrices of the ResNet50, InceptionResNet-V2, DenseNet169, and ensemble DL models. The results indicate that the proposed ensemble DL method achieved higher accuracy compared to the other individual models. Additionally, ResNet50 was observed to provide the second-best results after the ensemble DL method. While the Ensemble DL, ResNet50, and DenseNet169 models made the most misclassifications in the plastic class, it was observed that the misclassification examples of the InceptionResNet-V2 model were distributed across all classes. When the confusion matrices for each waste class are examined, it is seen that the ensemble DL model exhibits better classification performance than individual models for cardboard, glass, metal, paper, and textile trash classes. However, it is seen that the InceptionResNet-V2 model is better than the ensemble DL model for the miscellaneous class.

The test accuracy, precision, recall, F-score, and ROC-AUC metrics for ResNet50, InceptionResNet-V2, DenseNet169, and the ensemble DL models are presented in Figure 14. The performance metrics for ResNet50 and InceptionResNet-V2 models exhibit similar levels, while the DenseNet169 model performs relatively weaker. Furthermore, the ensemble DL model showed a 10% increase in accuracy over DenseNet169. The Ensemble DL model outperformed all individual models in all metrics. The reason why the Ensemble DL model is more stable and successful than individual models is that it feeds from different representations of the input image.

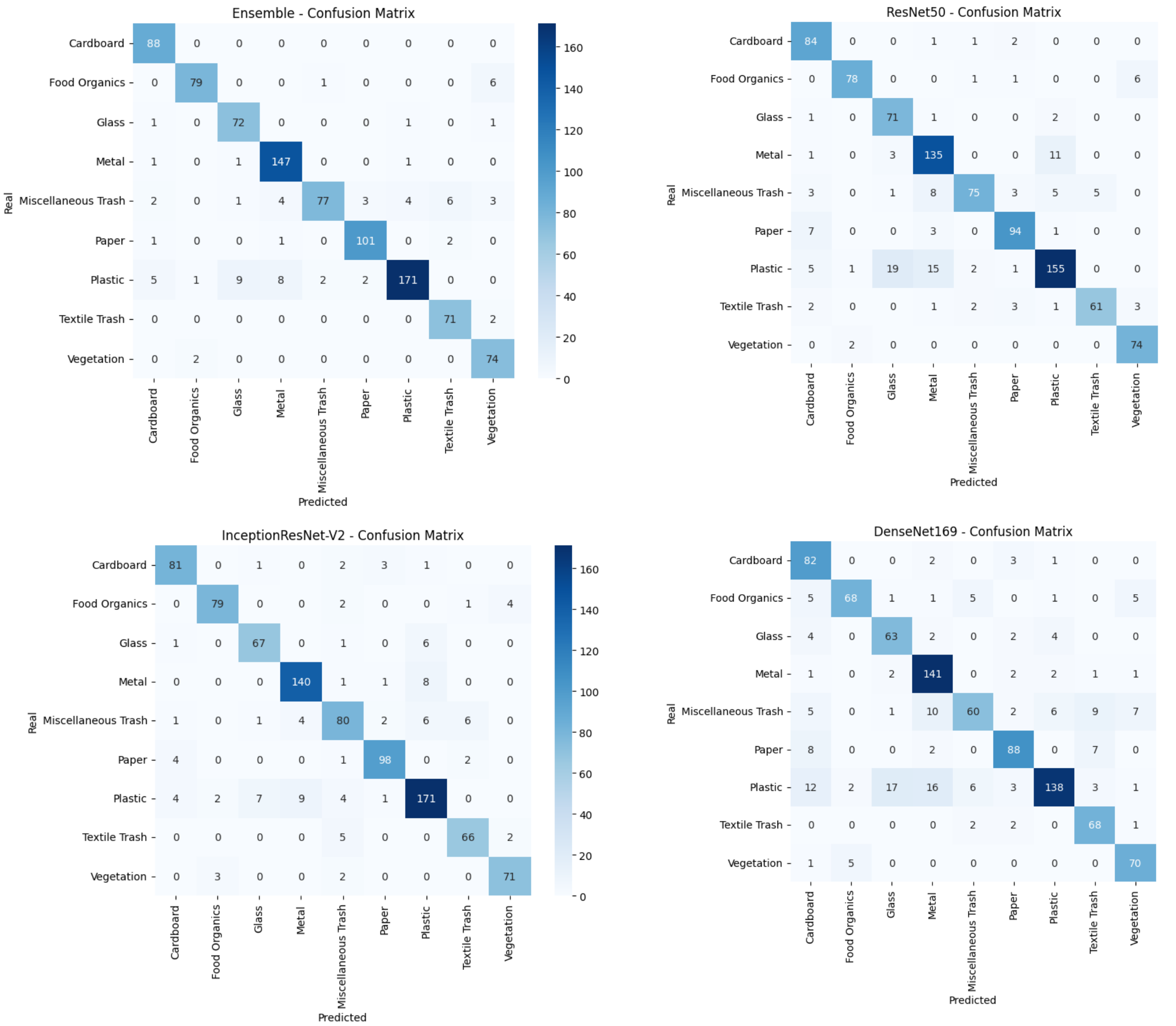

Figure 13. Confusion Matrices of ResNet50, InceptionResNet-V2, DenseNet169, and Ensemble DL Models in Dataset-2

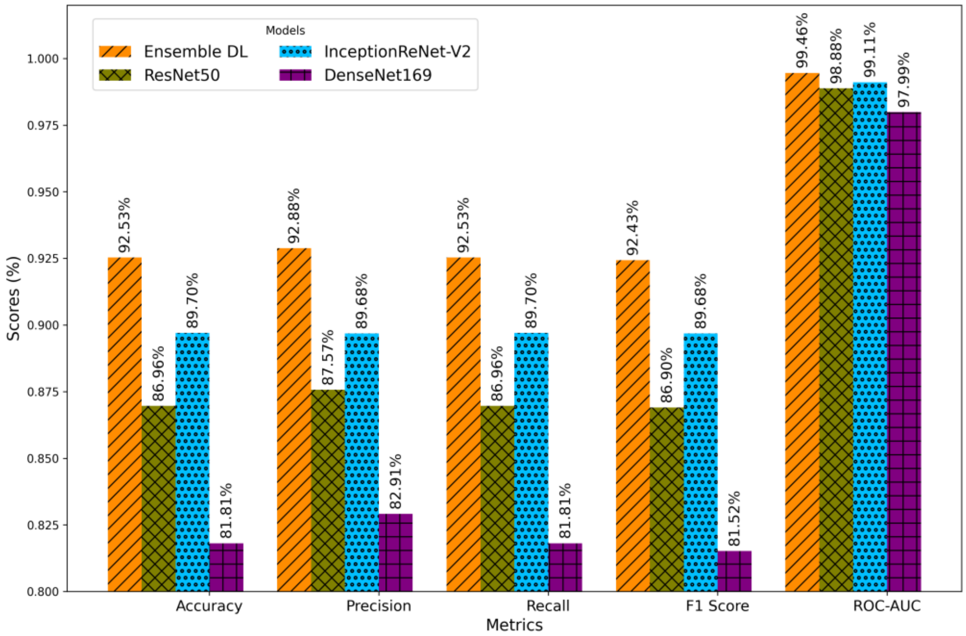

Figure 14. Performance Metrics of ResNet50, InceptionResNet-V2, DenseNet169, and the Ensemble DL Models in Dataset-2

Figure 15 presents the accuracy and loss curves of the models during the training and validation. The train accuracy graph shows that the accuracy rates of all models increased over time, reaching high values after 50 epochs. In the validation accuracy curve, the validation performance of the models follows a trajectory parallel to their training performance. Examining the loss curves, both train loss and validation loss decrease as the number of epochs increases, indicating that the optimization process is progressing effectively. Notably, the ensemble DL model outperforms the other models by achieving lower validation loss and higher accuracy, demonstrating the best performance. These results confirm that the models were successfully trained and did not encounter issues such as overfitting during the validation. Furthermore, the superior performance of the ensemble DL model compared to other individual models indicated that it generalizes better on Dataset-2.

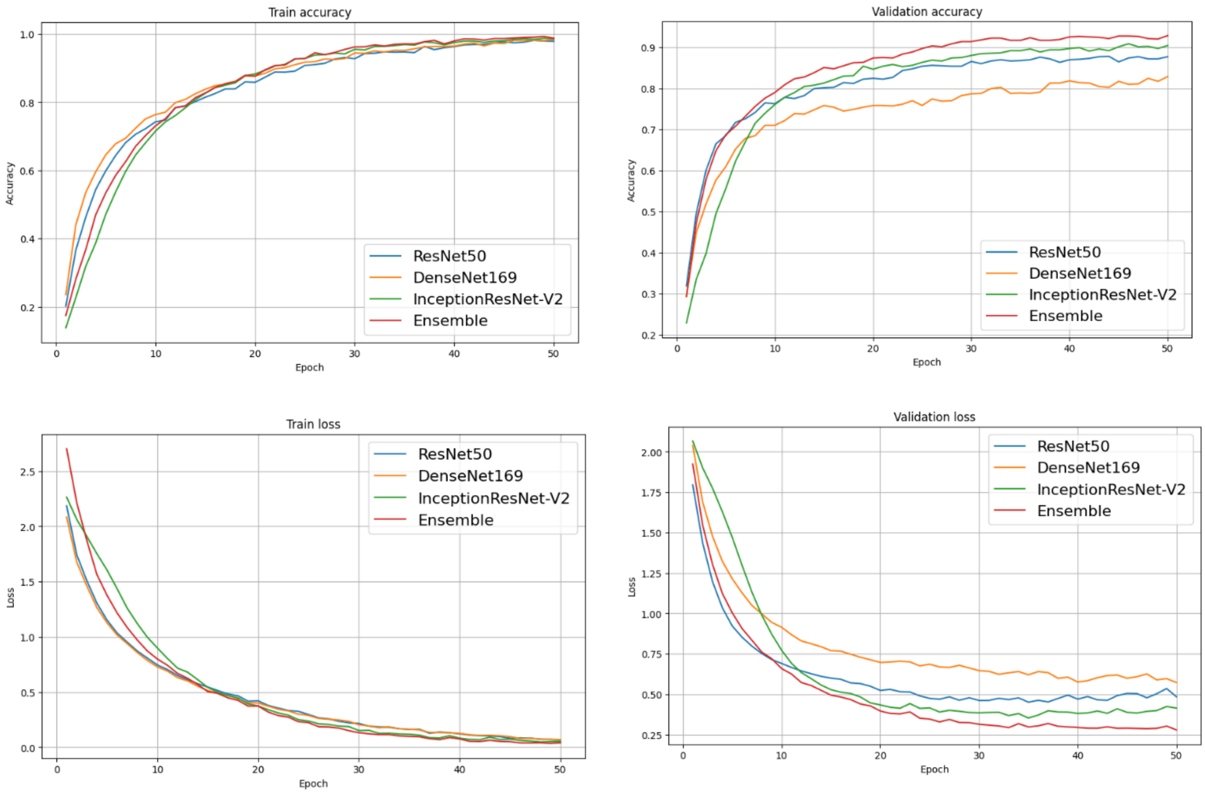

Figure15. Accuracy and Loss of DL and Ensemble Models During the Train and Test in Dataset-2.

In Dataset-2, the ML-based approaches were applied, using ANOVA, VT, MI, RF, Lasso, RFE, PCA, and RR methods to create different feature subsets by increasing the number of selected features in increments of five between 5 and 200. SVM, MLP, and RF classification algorithms are optimized and utilized on selected features. For ensemble ML models, SVM, MLP, and RF were used as base classifiers. The models were comprehensively evaluated using various metrics, including accuracy, precision, recall, F-score, ROC-AUC, confusion matrices, and classification reports. The obtained results were analyzed to compare the performance of the feature selection methods and classification algorithms, as well as to determine the most suitable model-feature count combinations. When all classification results were examined, as shown in Figure 16, the ANOVA method, with 170 selected features and the soft voting model, achieved the highest performance with an accuracy of 93.17%. Dataset-2 is imbalanced in terms of class distribution. Moreover, the number of samples in the plastic and metal classes is significantly different from the other classes. Therefore, the effectiveness of ANOVA's F-test based on inter-class variance in capturing the distinctive features of Dataset-2 has led to superior performance compared to other methods.

Figure 17 presents the comparison of accuracy, precision, recall, F1-score, and ROC-AUC metrics among the models. The proposed soft voting model demonstrated the best performance across all metrics, achieving 93.17% accuracy, 93.28% precision, 93.17% recall, 93.08% F1-score, and 99.32% ROC-AUC. The ensemble DL model, however, achieved a higher ROC-AUC value of 99.46%, outperforming the ensemble ML soft voting model in this specific metric. Meanwhile, the hard voting model also exhibited a competitive performance, with 92.74% accuracy, 92.74% recall, and 92.66% F1-score. Among the various models, the RF attained a notable ROC-AUC score of 99.37%, although it underperformed in other metrics compared to soft voting. SVM, MLP, and XGBoost models showed balanced performance, though XGBoost obtained relatively lower values compared to the other models. In Dataset-2, it is seen that the best classification performance after soft voting and hard voting is in the SVM algorithm.

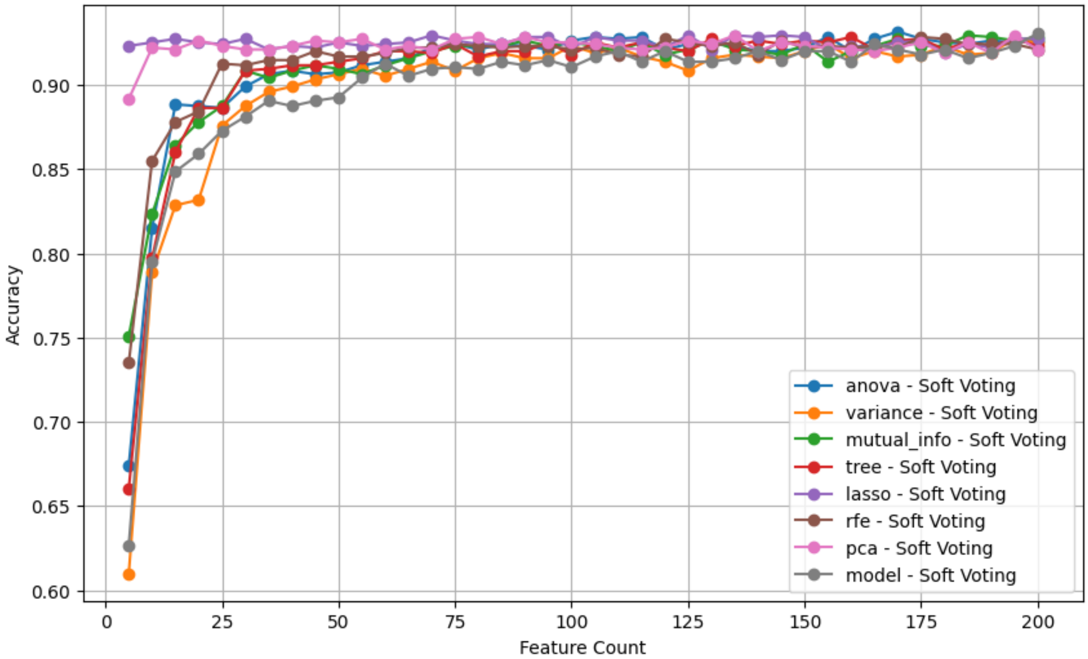

Figure16. Accuracy Value of the Soft Voting Classifier in Dataset-2 Based on the Feature Selection Method and the Number of Selected Features

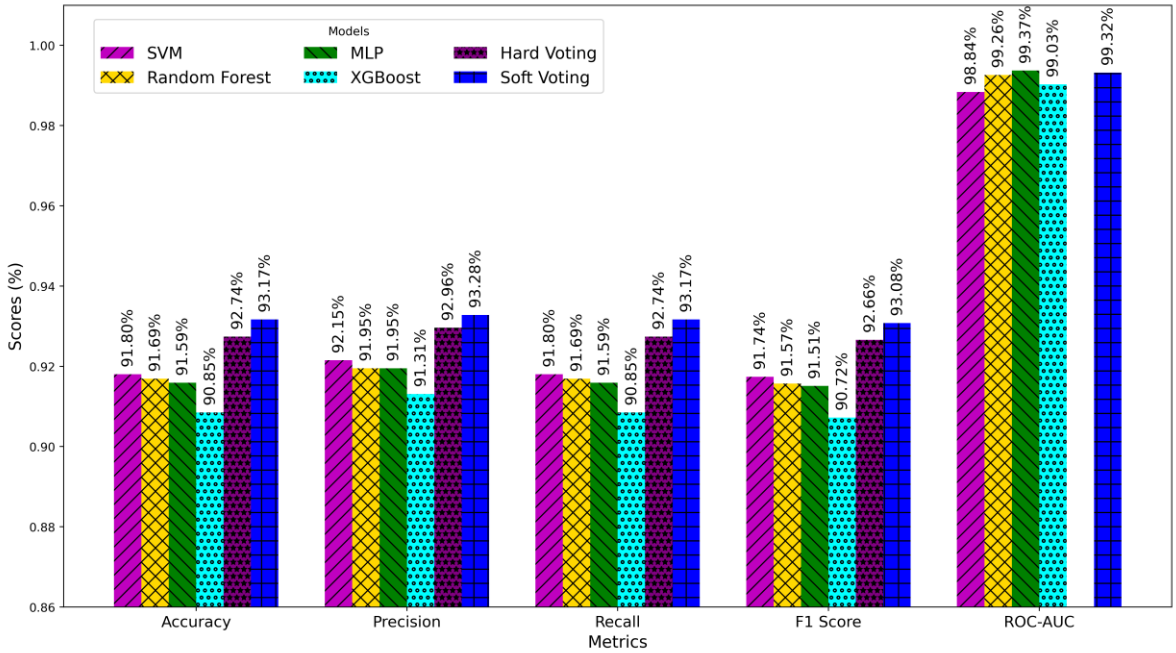

Figure17. Performance Metrics of ML Models and Ensemble ML Models in Dataset-2

The confusion matrix for the ANOVA feature selection method, with 170 selected features and the soft voting ensemble ML method, is presented in Figure 18. The proposed soft voting ensemble method demonstrates high classification accuracy. Among the total of 951 images analyzed, 65 were identified as misclassified, resulting in 886 images being accurately classified. The soft voting ensemble model demonstrated superior classification performance by correctly classifying 6 additional images when compared to the DL ensemble method.

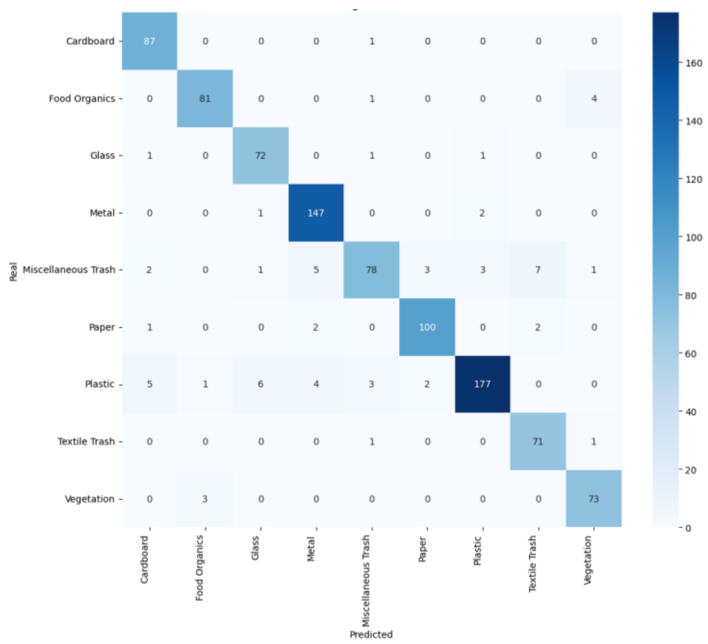

Figure18. Confusion Matrix of the Soft Voting Model with ANOVA -170 Features in Dataset-2

A heatmap summarizing the precision, recall, and F1-score performance metrics for different classes in the ML ensemble method is presented in Figure 19. The cardboard, food organics, glass, metal, paper, plastic, textile trash, and vegetation classes all achieved metric values above 89%, indicating that the model effectively differentiates these classes. This suggests that the model has successfully learned the characteristics of these waste types and can accurately classify them. However, the miscellaneous trash class exhibited the lowest performance, with a recall value of 78%, highlighting a relative difficulty in accurately identifying this class. According to the classification metrics values in Figure 19, it is possible to state that the proposed model fails the most in the miscellaneous trash class in the dataset.

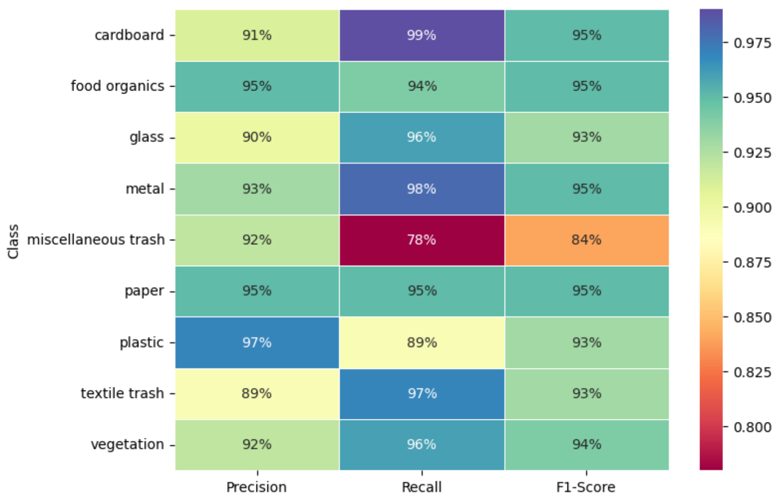

Figure 19. Heatmap of the Soft Voting Model with Lasso-145 Features in Dataset-2

Figure 20 presents a comparison of the performance metrics of the ensemble DL and soft voting ML models. The results indicate that the soft voting model outperforms the ensemble DL model in terms of all metrics. The soft voting model achieved 93.17% accuracy, 93.28% precision, 93.17% recall, 93.08% F1-score, and 99.32% ROC-AUC, demonstrating its superior classification performance. In comparison, the ensemble DL model attained 92.53% accuracy, 92.88% precision, 92.53% recall, 92.43 F1-score, and 99.46 ROC-AUC. The fact that soft voting achieved better results in accuracy, precision, recall, and F1-score highlights its potential to accurately classify positive classes and enhance overall classification performance.

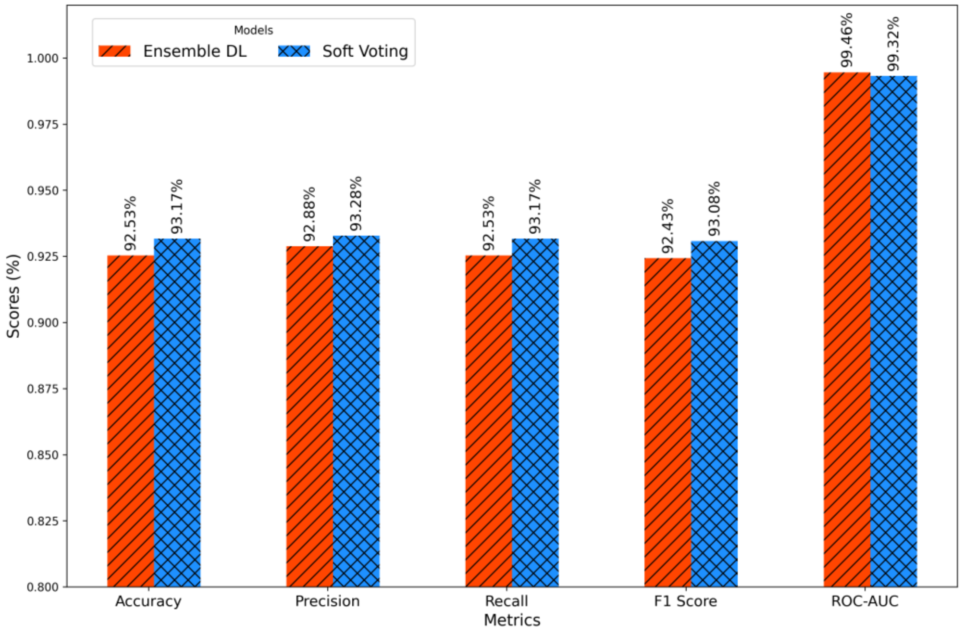

Figure 20. Comparison of the Performance Metrics of the Ensemble DL Model and the Soft Voting Model in Dataset-2

The ensemble ML model has performed better than the ensemble DL model in both Dataset-1 (Figure 12) and Dataset-2 (Figure 20). It is possible to state that both FS and model parameter optimization affect the performance of the ensemble ML model. Moreover, the results of this proposed end-to-end model show that ensemble models improve the classification performance both in the feature extraction stage and in the classification stage, and that the imbalance problems of the dataset are least affected. However, the proposed method has some disadvantages. Firstly, ensemble methods require more computation than individual methods. The fact that the proposed method adopts an ensemble approach in both the feature extraction and classification can be considered a deficiency in terms of computational requirements. Secondly, since GridSearchCV, which is used in ML parameter optimization, uses a brute-force strategy, it can be considered a disadvantage in terms of both time and computation in large search spaces.

Comparison of the proposed model results with the previous studies

Figure 21 shows that the ensemble learning-based method created for sorting environmental trash did as well as or better than other studies that have been done. The proposed method showed 95.26% accuracy on Dataset-1 with six classes. Salur et al., using the same dataset with four classes, achieved 95.22% accuracy (Salur et al., 2024). Furthermore, Sürücü et al. obtained an accuracy of 94.69% using an InceptionV3-based model, which is inferior to the performance of the proposed method (Sürücü & Ecemiş, 2022). On the other hand, the accuracy of 95.93% obtained by Tatke et al. is a remarkable achievement. However, the fact that this result was obtained using a single model suggests that its effectiveness on imbalanced datasets will be limited compared to ensemble models. (Tatke et al., 2021). The model proposed by Alsubaei et al. achieved an accuracy of 98.61%, outperforming the proposed method in terms of classification performance (Alsubaei et al., 2022). However, when compared to the 89.75% accuracy achieved by Yıldız et al., the proposed method demonstrates superior performance. Additionally, the 98% accuracy reported by Al Duhayyim et al. can be attributed to the fact that their dataset was classified into only two categories: "hazardous" and "non-hazardous", which simplifies the classification task (Al Duhayyim et al., 2023). In light of all these comparisons, the proposed method stands out as a strong and effective alternative in the field of environmental waste classification, offering high accuracy rates and superior generalization capability. Moreover, these compared studies do not have a plug-and-play structure as in our proposed model. One of the advantages of the model proposed in this study is that it allows the use of different models in the feature extraction, feature selection, and classification stages.

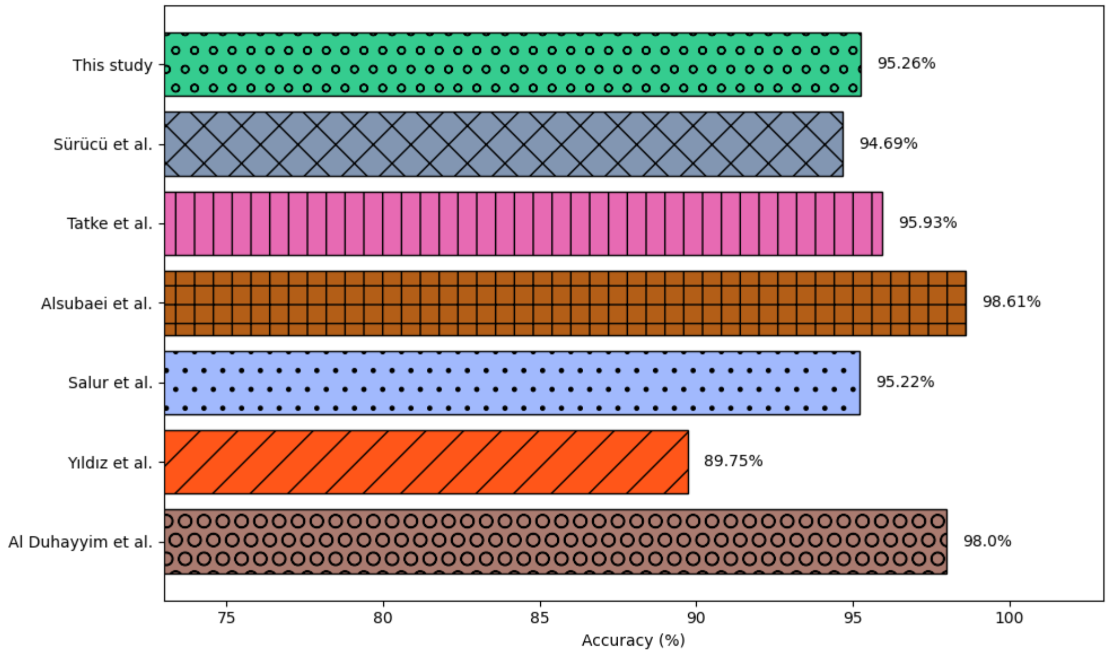

Figure 21. Comparison of the Performance of the Proposed Model on Dataset-1 with Previous Studies

Previous studies in the field of waste classification have been thoroughly analyzed, and the accuracy rates of existing methods have been compared. While some approaches in the literature have achieved high accuracy through techniques such as transfer learning and hyperparameter optimization, they have often been tested on a smaller number of classes or narrow-scope datasets. As shown in Figure 22, the proposed method achieved 93.17% accuracy on the Dataset-2, which is significantly higher than the 89.19% accuracy reported by Single et al. This discrepancy is mainly because Single et al. relied solely on individual models for waste classification, limiting their performance. On the other hand, the 96.31% accuracy achieved by Younis and Obaid using the EfficientNet model slightly surpasses the classification performance of our proposed method.

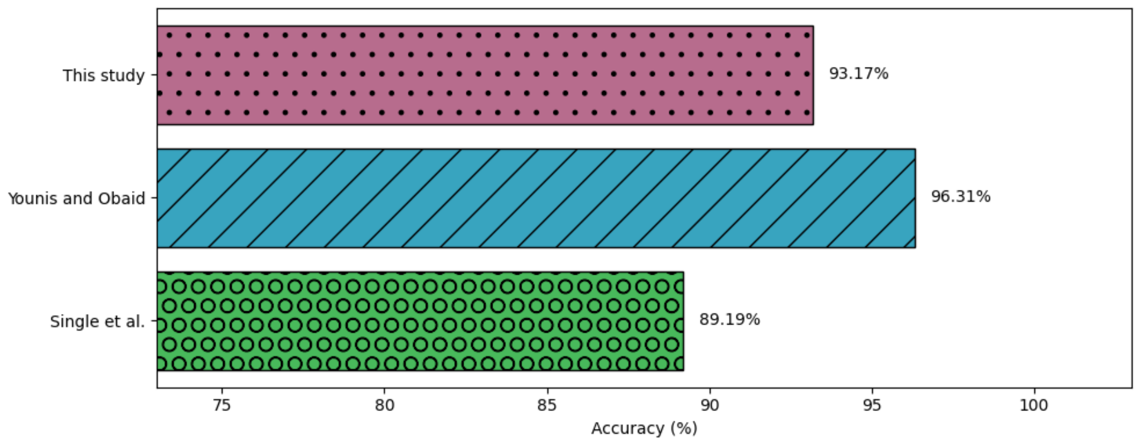

Figure 22. Comparison of the Performance of the Proposed Model on Dataset-2 with Previous Studies

CONCLUSION

The classification of environmental waste is crucial for enhancing the efficiency of sustainable waste management and recycling processes. Proper classification of waste facilitates the segregation of recyclable materials, contributing to natural resource conservation and energy savings. Moreover, it plays a vital role in ensuring the proper disposal of hazardous waste, thereby reducing environmental pollution and potential health risks. This study presents an end-to-end approach that integrates DL and ML techniques for environmental waste classification. By combining DL and ML with ensemble learning methods, the proposed approach enables accurate and efficient classification of environmental waste. To determine the effectiveness of the proposed model, tests were performed on two datasets (Dataset-1 and Dataset-2) for environmental waste classification, including several classification tasks.

In the pipeline of the proposed model, fine-tuning was applied to ResNet50, InceptionResNet-V2, and DenseNet169 models. These individual models were then combined to form an ensemble DL model. The extracted features from DL methods were subjected to various feature selection techniques, and the selected features were subsequently fed into ML algorithms as input, ensuring a comprehensive classification process. In the feature selection phase, methods such as ANOVA, Lasso, PCA, and MI were employed to eliminate irrelevant or low-importance features from the datasets. This process not only reduced model complexity but also enhanced classification performance. For ML classification, algorithms such as SVM, MLP, and RF were utilized, along with ensemble methods like XGBoost, hard voting, and soft voting, achieving high accuracy rates. The experiments conducted on Dataset-1 and Dataset-2 demonstrated the generalization capability of the proposed method. The ensemble ML model achieved 95.26% accuracy on Dataset-1, outperforming individual models. Similarly, an accuracy of 93.17% was obtained on Dataset-2. The superior performance of ensemble models over individual models clearly highlights the advantages of combining different model architectures. Additionally, the contributions of feature selection and optimization processes in ML classification were supported by thorough analysis. The proposed method represents a significant step toward the automation of waste management and recycling processes. This approach accelerates waste classification, reduces reliance on manual labor, and contributes to the development of a more efficient waste management system. The applicability of the proposed model is particularly noteworthy in increasing recycling rates and reducing waste management costs. The innovative methodologies introduced in this study serve as a foundation for future research aimed at enhancing environmental sustainability. Future studies may focus on integrating the proposed model into real-time waste management systems, testing it on diverse datasets with a wider variety of waste classes, and exploring its integration with sensor-based or robotic applications for large-scale solutions.

ACKNOWLEDGEMENTS

This study was supported by Gaziantep Islam Science and Technology University Scientific Research Projects Coordination Unit under Project No: 2024-GAP-Mühe-0008.

REFERENCES

Abdu, H., & Noor, M. H. M. (2022). A survey on waste detection and classification using deep learning. IEEE Access, 10, 128151–128165. https://doi.org/10.1109/ACCESS.2022.3226682

Al Duhayyim, M., Alotaibi, S. S., Al-Otaibi, S., Al-Wesabi, F. N., Othman, M., Yaseen, I., Rizwanullah, M., & Motwakel, A. (2023). An Intelligent Hazardous Waste Detection and Classification Model Using Ensemble Learning Techniques. Computers, Materials & Continua, 74(2). https://doi.org/10.32604/cmc.2023.033250

Alsubaei, F. S., Al-Wesabi, F. N., & Hilal, A. M. (2022). Deep Learning-Based Small Object Detection and Classification Model for Garbage Waste Management in Smart Cities and IoT Environment. Applied Sciences (Switzerland), 12(5). https://doi.org/10.3390/app12052281

Chatterjee, S., Hazra, D., & Byun, Y. C. (2022). IncepX-Ensemble: Performance Enhancement Based on Data Augmentation and Hybrid Learning for Recycling Transparent PET Bottles. IEEE Access, 10, 52280–52293. https://doi.org/10.1109/ACCESS.2022.3174076

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining, 785–794. https://doi.org/10.1145/2939672.2939785

Chen, W., Yang, K., Yu, Z., Shi, Y., & Chen, C. L. (2024). A survey on imbalanced learning: latest research, applications and future directions. Artificial Intelligence Review, 57(6), 1–51. https://doi.org/10.1007/s10462-024-10759-6

Dietterich, T. G. (2000). Ensemble methods in machine learning. International Workshop on Multiple Classifier Systems, 1–15. https://doi.org/10.1007/3-540-45014-9_1

Filtreleme yöntemleri· Miuul. (2021). https://miuul.com/blog/filtreleme-yontemleri-ile-makine-ogrenmesinde-degisken-secimi Accessed 11.01.2025.

Garbage Classification. (2021). https://www.kaggle.com/datasets/asdasdasasdas/garbage-classification Accessed 10.01.2025. https://doi.org/10.34740/kaggle/ds/81794

Genuer, R., Poggi, J.-M., Genuer, R., & Poggi, J.-M. (2020). Random forests. Springer. https://doi.org/10.1007/978-3-030-56485-8_3

Greenacre, M., Groenen, P. J. F., Hastie, T., d'Enza, A. I., Markos, A., & Tuzhilina, E. (2022). Principal component analysis. Nature Reviews Methods Primers, 2(1), 100. https://doi.org/10.1038/s43586-022-00184-w

Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. Journal of Machine Learning Research, 3(Mar), 1157–1182.

Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. Machine Learning, 46, 389–422. https://doi.org/10.1023/A:1012487302797

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 770–778.

Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 4700–4708.

Hutchinson, M. L., Antono, E., Gibbons, B. M., Paradiso, S., Ling, J., & Meredig, B. (2017). Overcoming data scarcity with transfer learning. ArXiv Preprint ArXiv:1711.05099. https://doi.org/10.48550/arXiv.1711.05099

Kanevski, M., Pozdnukhov, A., Canu, S., & Maignan, M. (2002). Advanced spatial data analysis and modeling with Support Vector Machines.

Kaza, S., Yao, L., Bhada-Tata, P., & Van Woerden, F. (2018). What a waste 2.0: a global snapshot of solid waste management to 2050. World Bank Publications.

Ma, X., Chen, L., Deng, Z., Xu, P., Yan, Q., Choi, K.-S., & Wang, S. (2023). Deep image feature learning with fuzzy rules. IEEE Transactions on Emerging Topics in Computational Intelligence, 8(1), 724–737. https://doi.org/10.1109/TETCI.2023.3259447

McDonald, G. C. (2009). Ridge regression. Wiley Interdisciplinary Reviews: Computational Statistics, 1(1), 93–100. https://doi.org/10.1002/wics.14

Meyen, S., Göppert, F., Alber, H., von Luxburg, U., & Franz, V. H. (2021). Specialists Outperform Generalists in Ensemble Classification. ArXiv Preprint ArXiv:2107.04381. https://doi.org/10.48550/arXiv.2107.04381

Ouedraogo, A. S., Kumar, A., & Wang, N. (2023). Landfill Waste Segregation Using Transfer and Ensemble Machine Learning: A Convolutional Neural Network Approach. Energies, 16(16), 5980. https://doi.org/10.3390/en16165980

Pinkus, A. (1999). Approximation theory of the MLP model in neural networks. Acta Numerica, 8, 143–195. https://doi.org/10.1017/S0962492900002919

Poudel, S., & Poudyal, P. (2022). Classification of waste materials using CNN based on transfer learning. Proceedings of the 14th Annual Meeting of the Forum for Information Retrieval Evaluation, 29–33. https://doi.org/10.1145/3574318.3574345

Santoso, H., Hanif, I., Magdalena, H., & Afiyati, A. (2024). A Hybrid Model for Dry Waste Classification using Transfer Learning and Dimensionality Reduction. JOIV: International Journal on Informatics Visualization, 8(2), 623-634. https://dx.doi.org/10.62527/joiv.8.2.1943

Salur, M. U., Elmas, N., Koçak, A. N., & Kaymaz, M. (2024). Derin Öğrenme ile Çevresel Atıkların Sınıflandırılmasına Dayalı Akıllı Çöp Konteyneri Tasarımı ve Prototipinin Geliştirilmesi. Adıyaman Üniversitesi Mühendislik Bilimleri Dergisi, 11(24), 547–563. https://doi.org/10.54365/adyumbd.1557588

Shahab, S., Anjum, M., & Umar, M. S. (2022). Deep learning applications in solid waste management: A deep literature review. International Journal of Advanced Computer Science and Applications, 13(3). https://doi.org/10.14569/ijacsa.2022.0130347

Single, S., Iranmanesh, S., & Raad, R. (2023). RealWaste: A Novel Real-Life Data Set for Landfill Waste Classification Using Deep Learning. Information (Switzerland), 14(12). https://doi.org/10.3390/info14120633

Sürücü, S., & Ecemiş, İ. N. (2022). Garbage classification using pre-trained models. Avrupa Bilim ve Teknoloji Dergisi, 36, 73–77. https://doi.org/10.31590/ejosat.1103628

Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. (2017). Inception-v4, inception-resnet and the impact of residual connections on learning. Proceedings of the AAAI Conference on Artificial Intelligence, 31(1). https://doi.org/10.1609/aaai.v31i1.11231

Şirin, E. (2017). Ensemble Yöntemler: Basit Teorik Anlatım ve Python Uygulama - Veri Bilimi Okulu - Veri Bilimi Okulu. https://www.veribilimiokulu.com/ensemble-yontemler-topluluk-ogrenmesi-basit-teorik-anlatim-ve-python-uygulama/ Accessed 10.01.2025

Tatke, A., Patil, M., Khot, A., & Karad's, P. J. V. (2021). Hybrid approach of garbage classification using computer vision and deep learning. International Journal of Engineering Applied Sciences and Technology, 5(10). https://doi.org/10.33564/IJEAST.2021.v05i10.032

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society Series B: Statistical Methodology, 58(1), 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Toğaçar, M., Ergen, B., & Cömert, Z. (2020). Waste classification using AutoEncoder network with integrated feature selection method in convolutional neural network models. Measurement: Journal of the International Measurement Confederation, 153. https://doi.org/10.1016/j.measurement.2019.107459

Vergara, J. R., & Estévez, P. A. (2014). A review of feature selection methods based on mutual information. Neural computing and applications, 24, 175-186. https://doi.org/10.1007/s00521-013-1368-0

Wu, N., Wang, G., & Jia, D. (2024). A Hybrid Model for Household Waste Sorting (HWS) Based on an Ensemble of Convolutional Neural Networks. Sustainability, 16(15), 6500. https://doi.org/10.3390/su16156500

Yıldız, E. N., Bingöl, H., & Yıldırım, M. (2023). Önerilen Derin Öğrenme ve Makine Öğrenmesi Tabanlı Hibrit Model ile Çevresel Atıkların Sınıflandırılması. Fırat Üniversitesi Mühendislik Bilimleri Dergisi, 35(1), 353–361. https://doi.org/10.35234/fumbd.1230982

Younis, H., & Obaid, M. (2024). Performance Comparison of Pretrained Deep Learning Models for Landfill Waste Classification. International Journal of Advanced Computer Science & Applications, 15(11). https://dx.doi.org/10.14569/IJACSA.2024.0151166

Yulita, I. N., Ardiansyah, F., Sholahuddin, A., Rosadi, R., Trisanto, A., & Ramdhani, M. R. (2024, January). Garbage Classification Using Inception V3 as Image Embedding and Extreme Gradient Boosting. International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSIS) (pp. 1394-1398). IEEE. https://doi.org/10.1109/ICETSIS61505.2024.10459560

Zheng, H., & Gu, Y. (2021). Encnn-upmws: Waste classification by a CNN ensemble using the UPM weighting strategy. Electronics (Switzerland), 10(4), 1–21. https://doi.org/10.3390/electronics10040427