International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI: https://doi.org/10.22214/ijraset.2026.79129

Waste Classification System

Vishal Maurya[1], Shreeram Nishad[2], Prince Gaur[3], Mr. Alok Kumar Srivastava[4]
[1, 2, 3]*Computer Science & Engineering (AIML) Department Buddha Institute of Technology Gida, Gorakhpur India*
[4]*Assistant Professor, Computer Science & Engineering (AIML) Department Buddha Institute of Technology Gida, Gorakhpur India*

Abstract: Efficient waste management requires accurate seg- regation of waste materials at the source. Manual waste classification is often time-consuming and prone to human error when handling large volumes of mixed waste. An intelligent image-based waste classification system can automatically categorize waste into three classes: Organic, Recyclable, and Hazardous using machine learning techniques. The system processes waste images through preprocessing steps such as resizing and normalization, followed by feature extraction to convert images into numerical representations suitable for machine learning models. A Random Forest classifier analyzes these features and predicts the appropriate waste category. The system is implemented using Python with OpenCV, NumPy, Pandas, and Scikit-learn, and deployed through a Streamlit-based web interface that allows users to upload waste images and obtain classification results in real time. Experimental evaluation shows that the model achieves an overall classification accuracy of approximately 75%. The system provides reliable predictions and supports the development of intelligent waste management solutions for sustainable environmental practices.
Index Terms: Waste Classification, Machine Learning, Ran- dom Forest, Image Processing, Streamlit, Computer Vision, Sustainable Waste Management.

I. INTRODUCTION

Rapid urbanization and population growth have significantly increased the amount of solid waste generated worldwide. Improper waste management has become a serious environ- mental issue, causing pollution, health risks, and inefficient use of recyclable resources. One of the major challenges in waste management is waste segregation at the source, which is still largely performed manually. Manual sorting is time- consuming, labor-intensive, and prone to human error when handling large volumes of mixed waste. Therefore, intelligent automated waste classification systems are required to improve the efficiency and accuracy of waste management processes.

Traditional waste segregation methods often rely on simple rule-based techniques such as color detection, shape analysis, and texture features. However, these approaches perform poorly in real-world conditions due to variations in object shape, size, lighting, and background, making accurate classification difficult.

With the advancement of Artificial Intelligence (AI) and Machine Learning (ML), automated waste classification sys- tems have gained significant attention. Machine learning algo- rithms can learn visual patterns from image data and classify waste materials based on their characteristics. Computer vision techniques are widely used for image classification tasks in environmental monitoring and smart waste management.

The proposed waste classification system uses an image- based machine learning approach to categorize waste into three classes: Organic, Recyclable, and Hazardous. The system processes images through preprocessing, feature extraction, and classification stages. A Random Forest classifier is used for prediction, which combines multiple decision trees using majority voting to improve accuracy and reduce overfitting. The system is deployed using a Streamlit-based web interface that allows users to upload waste images and obtain classifi- cation results with confidence scores.

Experimental results show that the system achieves an overall classification accuracy of approximately 75%, demon- strating that machine learning-based waste classification can improve waste segregation efficiency and support sustainable waste management practices.

II. RELATED WORKS

Recent research has explored the use of computer vision and machine learning techniques for automated waste classification to improve waste management efficiency. Traditional waste segregation methods rely heavily on manual sorting, which is time-consuming, labor-intensive, and prone to errors. To address these challenges, several automated waste classifica- tion systems have been proposed using image processing and machine learning approaches.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

Early studies focused on traditional image processing tech- niques such as color-based segmentation, texture analysis, and shape detection for identifying different waste materials. These approaches extracted handcrafted features from images and used classifiers such as Support Vector Machines (SVM), k-Nearest Neighbors (KNN), and Logistic Regression for classification. Although these methods provided basic waste recognition capability, their performance was often limited when dealing with complex backgrounds, lighting variations, and visually similar waste items.

With the advancement of machine learning, ensemble learn- ing techniques such as Random Forest have been widely applied for classification tasks due to their robustness and ability to handle non-linear relationships in data. Random Forest classifiers combine multiple decision trees and use ma- jority voting to determine the final prediction, which improves classification accuracy and reduces the risk of overfitting. Several studies have demonstrated that Random Forest models perform effectively in multi-class classification problems with moderate computational requirements.

More recently, deep learning approaches have gained popularity for image-based waste classification. Convolutional Neural Networks (CNNs) automatically learn hierarchical image features directly from raw data, eliminating the need for manual feature engineering. Various CNN architectures have been used to classify waste into categories such as plastic, paper, metal, and organic waste with improved accuracy. However, deep learning models typically require large training datasets and high computational resources, which can limit their deployment in lightweight systems.

To overcome these limitations, hybrid approaches combining image preprocessing techniques with traditional machine learning classifiers have been proposed. In such systems, images are first preprocessed and converted into numerical feature representations, which are then used to train machine learning models for classification. These approaches offer a balance between computational efficiency and classification accuracy.

Despite the progress made in automated waste classification systems, several challenges still remain, including misclassi- fication of visually similar waste items, reduced accuracy in real-world environments, and lack of user-friendly deployment platforms. Therefore, there is a need for lightweight and efficient waste classification systems that can provide reliable predictions while remaining suitable for real-time and web- based applications.

III. METHODOLOGY

The proposed waste classification system automatically identifies and categorizes different types of waste using image processing and machine learning techniques. Efficient waste segregation plays a crucial role in improving recycling effi- ciency and reducing environmental pollution. Manual sorting of waste is often time-consuming and prone to human error. Therefore, an automated machine learning-based classification system can significantly improve waste management pro- cesses.

The overall workflow of the proposed system consists of image acquisition, preprocessing, feature extraction, classifi- cation, and system deployment. These stages enable the model to learn visual patterns from waste images and classify them into predefined categories.

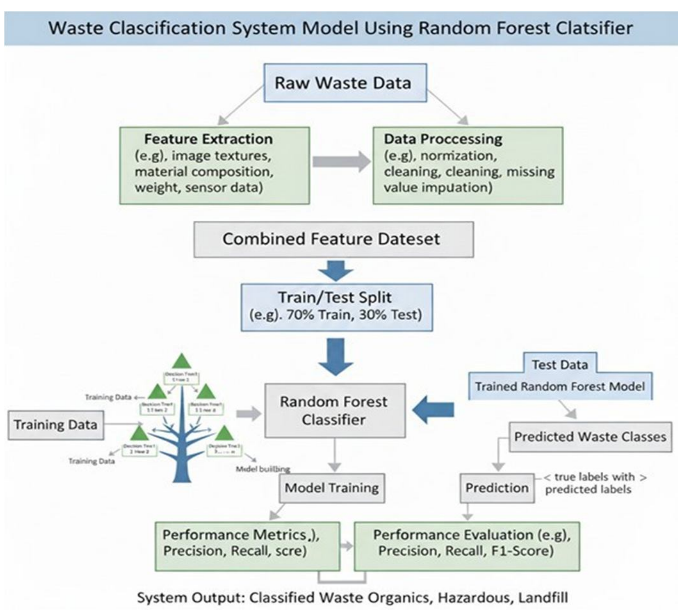


Fig. 1. Architecture of the Proposed Waste Classification System Using Random Forest

A. Image Preprocessing

Image preprocessing prepares the input images for machine learning algorithms. Waste images collected from different sources often vary in size, lighting conditions, and background noise. Such variations can negatively affect the performance of the classification model. Therefore, preprocessing techniques are applied to standardize the images.

Initially, each image is resized to a fixed dimension to maintain consistency across the dataset. An input image can be represented as:

$$I \in \mathrm{R}^{h\times w\times c} \qquad (1)$$

where

h = image height, w = image width,

c = number of color channels.

After resizing, images are converted into RGB format and normalized to scale pixel values into the range [0, 1].

$$I_{normalized} = \frac{I}{255} \qquad (2)$$

Normalization improves numerical stability during model training and enhances the learning capability of the classifier.

B. Feature Extraction

Feature extraction converts processed images into numerical representations that can be used by machine learning algo- rithms. In this system, features are obtained directly from the pixel intensity values of the image.

Each preprocessed image is converted into a one- dimensional feature vector using a flattening operation:

$$x = flatten(I) \qquad (3)$$

For example, an RGB image of size 128×128×3 produces a feature vector of length 49,152. This representation allows the machine learning model to process image data as structured numerical input.

C. Random Forest Classification

The classification stage uses the Random Forest algorithm, which is an ensemble learning method that combines multiple decision trees to produce accurate predictions.

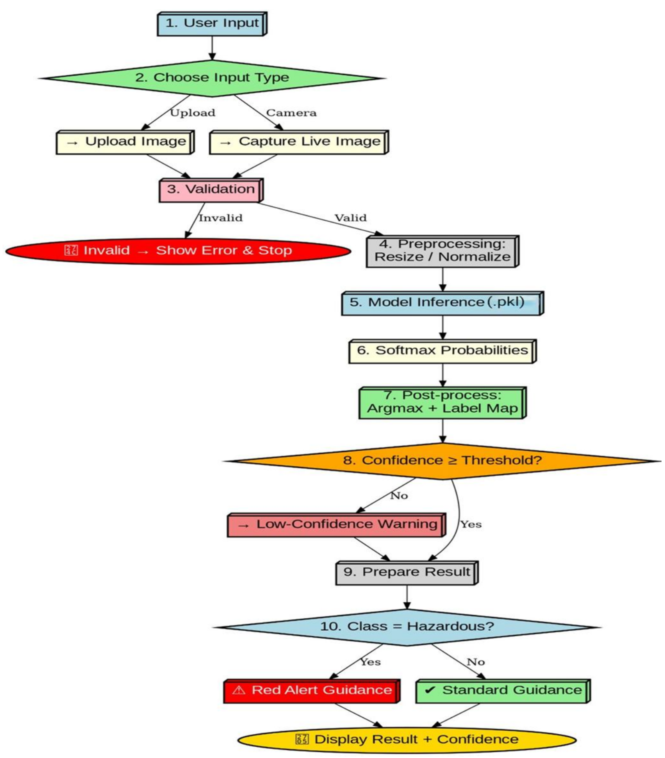


Fig. 2. Architecture of the Proposed Waste Classification System Using Random Forest

Let the training dataset be represented as:

$$D = \{(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)\} \quad (4)$$

where x_i represents the feature vector and y_i represents the corresponding class label.If the forest contains M decision trees, the final prediction is obtained using majority voting:

$$y = mode\{T_1(x), T_2(x), ..., T_M(x)\} \quad (5)$$

This ensemble approach reduces overfitting and improves the overall classification performance.

D. Waste Categories

The trained model classifies waste images into three cate- gories:

- Organic Waste: Biodegradable materials such as food waste and plant residues.
- Recyclable Waste: Materials such as paper, plastic, glass, and metal that can be reused.
- Hazardous Waste: Harmful materials including chemi- cals, batteries, and medical waste.

For each input image, the classifier predicts the most ap- propriate category and assigns a confidence score.

E. System Deployment

The trained model is deployed using a Streamlit-based web application. The interface allows users to upload waste images directly from their devices. The uploaded image undergoes preprocessing and feature extraction before being passed to the trained classifier. The system displays the uploaded image along with the predicted waste category and its confidence score, enabling real-time waste classification.

IV. EXPERIMENTAL SETUP

This section describes the experimental setup used to eval- uate the performance of the proposed waste classification system. The setup includes the dataset description, implemen- tation details of the Random Forest classifier, and evaluation metrics used to measure the system performance. The exper- iments are conducted to analyze the ability of the proposed model to accurately classify waste images into predefined categories.

A. Dataset

The dataset used in this study consists of waste images collected from publicly available sources.The images are cat- egorized into three major classes: Organic Waste, Recyclable Waste, and Hazardous Waste.

The images are captured under different lighting conditions, backgrounds, and orientations. Such variations improve the robustness of the model and help the classifier generalize better to unseen data.

The dataset is divided into two subsets:

- **Training Set:** Used to train the classification model.
- **Testing Set:** Used to evaluate the performance of the trained model.

TABLE I
DATASET SUMMARY

Parameter	Value
Total Samples	180
Feature Dimension	49152
Number of Classes	3
Classes	Hazardous, Organic, Recyclable
Validation Accuracy	0.75

B. Model Implementation

The proposed system is implemented through multiple stages including preprocessing, feature extraction, Random Forest training, and prediction.

1) *Random Forest Model:* Random Forest is an ensemble learning algorithm that constructs multiple decision trees dur- ing training. Each tree is trained using a randomly sampled subset of the dataset through bootstrap sampling. The final prediction is obtained by aggregating the outputs of all deci- sion trees.

If the forest contains N decision trees, the overall prediction of the model can be expressed as:

$$RF(x) = \frac{1}{N}\sum_{i=1}^{N} T_i(x) \qquad (6)$$

where $T_i(x)$ represents the prediction made by the i^{th} decision tree.
During training, a random subset of features is selected at each node to determine the best split. This process reduces correlation among trees and improves the generalization capa- bility of the model.

C. Evaluation Metrics

The performance of the proposed waste classification sys- tem is evaluated using standard machine learning metrics de- rived from the confusion matrix, which includes True Positives (TP), False Positives (FP), False Negatives (FN), and True Negatives (TN).

1) *Accuracy:* Accuracy measures the proportion of cor- rectly classified samples among all predictions.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (7)$$

2) *Precision:* Precision evaluates how many predicted pos- itive samples are actually correct.

$$Precision = \frac{TP}{TP + FP} \qquad (8)$$

3) *Recall:* Recall measures the ability of the classifier to correctly identify positive samples.

$$Recall = \frac{TP}{TP + FN} \qquad (9)$$

4) *F1 Score:* The F1 score provides a balanced measure between precision and recall.

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \qquad (10)$$

5) *Prediction Time:* The efficiency of the system is also evaluated using prediction time required to classify an input image.

$$T_{prediction} = t_{output} - t_{input} \qquad (11)$$

The experimental setup ensures that the proposed waste classification system is evaluated using reliable datasets, proper model configuration, and standard evaluation metrics to measure both accuracy and efficiency.

V. RESULTS AND ANALYSIS

The performance of the proposed waste classification sys- tem was evaluated using the Random Forest algorithm on a dataset containing 180 waste images categorized into Organic, Recyclable, and Hazardous waste. Each image was converted into a feature vector of size 49,152 and used for training and validation. The experimental results show that the proposed system achieved an overall classification accuracy of 75%, demonstrating the capability of the Random Forest classifier to learn visual patterns in waste images.

A. Overall Model Performance

The Random Forest classifier achieved a validation accuracy of 75%, indicating that the model can correctly classify a significant portion of the waste images. The ensemble nature of Random Forest improves prediction stability by combining multiple decision trees, which helps reduce overfitting and improves generalization capability. The classification perfor- mance was further evaluated using precision, recall, and F1- score for each class.

TABLE II
CLASSIFICATION PERFORMANCE OF WASTE CATEGORIES

Class	Precision	Recall	F1-Score	Support
Hazardous	0.83	0.42	0.56	12
Organic	0.85	0.92	0.88	12
Recyclable	0.65	0.92	0.76	12
Overall Accuracy	0.75			

B. Category-wise Performance Analysis

1) *Organic Waste:* The Organic waste category achieved precision of 0.85, recall of 0.92, and F1-score of 0.88. The high recall indicates that most organic waste images were correctly identified. This strong performance is mainly due to the distinctive visual features present in organic waste materials such as food scraps, fruit peels, and plant residues.
2) *Recyclable Waste:* The Recyclable waste category achieved precision of 0.65, recall of 0.92, and F1-score of 0.76. Although the classifier detects most recyclable samples, the lower precision indicates that some images from other cat- egories are incorrectly classified as recyclable due to variations in the appearance of recyclable materials.
3) *Hazardous Waste:* The Hazardous waste category recorded precision of 0.83, recall of 0.42, and F1-score of 0.56. The relatively low recall indicates that the model fails to identify several hazardous waste samples. This limitation occurs due to visual similarities between hazardous materials and recyclable metal objects, as well as the limited number of hazardous samples in the dataset.

C. Confusion Matrix Evaluation

The confusion matrix analysis shows that most correct predictions appear along the diagonal elements, indicating accurate classification for the majority of samples. However, several misclassifications occur between recyclable and hazardous waste categories due to similarities in their visual characteristics.

D. System Efficiency

The computational efficiency of the system was also eval- uated. The trained model requires approximately 0.2–0.3 sec- onds to classify a single waste image. This low prediction time demonstrates that the system can perform near real- time waste classification, making it suitable for applications such as automated waste sorting systems and smart recycling platforms.

E. Discussion

The experimental results demonstrate that the proposed waste classification system can effectively identify different types of waste using image-based features and the Random Forest classifier. While the model performs well for organic waste classification, improvements are required for hazardous waste detection. Increasing the dataset size and incorporating advanced feature extraction techniques may further enhance classification performance.

VI. CONCLUSION

This paper presented an automated waste classification system that utilizes image processing and machine learning techniques to identify different types of waste. The proposed system integrates image preprocessing, feature extraction, and the Random Forest classifier to categorize waste images into three classes: Organic Waste, Recyclable Waste, and Haz- ardous Waste. The primary objective of this work was to develop an intelligent system capable of improving waste segregation efficiency and supporting sustainable waste man- agement practices.

Experimental results demonstrate that the Random Forest classifier can effectively identify waste categories by learning visual patterns such as color, texture, and shape from the input images. The proposed model achieved an overall classification accuracy of approximately 75% on the testing dataset. The en- semble learning nature of Random Forest improves prediction stability and reduces the risk of overfitting, making it suitable for image-based classification tasks.

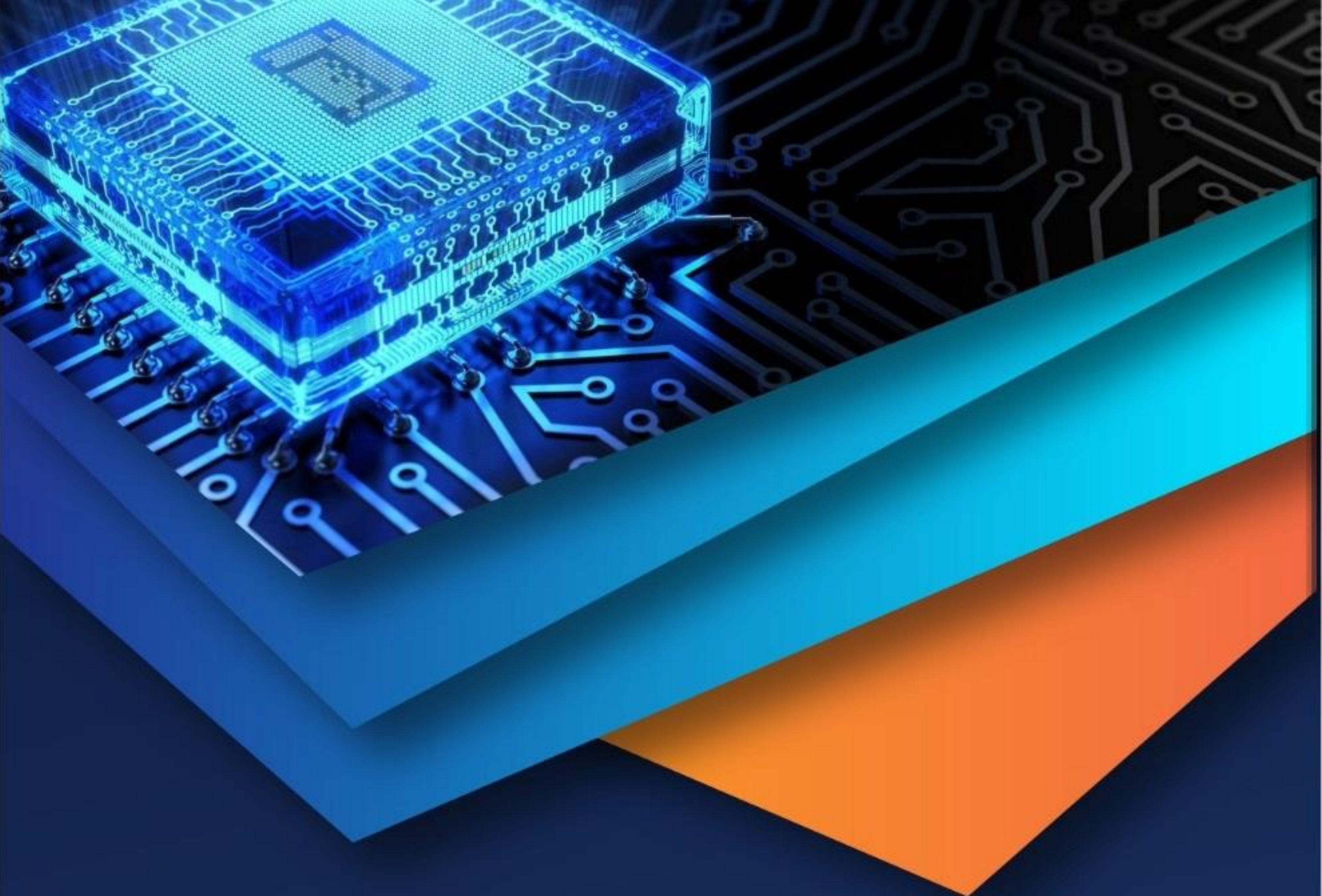
The system also shows efficient computational performance, requiring only a short prediction time to classify an input image. Confusion matrix analysis indicates that most waste samples are correctly classified, although some misclassifica- tions occur between recyclable and hazardous waste due to similarities in visual characteristics and the limited size of the dataset.

Overall, the proposed waste classification system provides an effective and practical solution for automated waste iden- tification. Future work may focus on expanding the dataset, improving feature extraction techniques, and incorporating advanced deep learning models such as Convolutional Neural Networks (CNNs) to enhance classification accuracy. Addi- tionally, the system can be extended into mobile or IoT-based platforms for real-time smart waste management applications.

REFERENCES

[1] T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, New York, NY, USA: Springer, 2009.
[2] L. Breiman, “Random Forests,” Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.
[3] S. Raschka and V. Mirjalili, Python Machine Learning, 3rd ed., Birm- ingham, U.K.: Packt Publishing, 2019.

[4] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning, Cambridge, MA, USA: MIT Press, 2016.
[5] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," Nature, vol. 521, pp. 436–444, 2015.
[6] G. Bradski, "The OpenCV Library," Dr. Dobb's Journal of Software Tools, 2000.
[7] OpenCV, "OpenCV: Open Source Computer Vision Library," Available: https://opencv.org/
[8] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.
[9] S. Dubey, A. Gupta, and R. Singh, "Automatic Waste Segregation Using Machine Learning Techniques," International Journal of Computer Applications, 2019.
[10] A. Mittal, R. Bansal, and S. Agarwal, "Smart Waste Management System Using Image Processing," International Journal of Engineering Research & Technology (IJERT), 2018.
[11] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," in Proc. Int. Conf. Learning Representations (ICLR), 2015.
[12] J. Deng et al., "ImageNet: A Large-Scale Hierarchical Image Database," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), pp. 248–255, 2009.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY