EnviroScan: Community and NGO Waste Solution

Shamal Dhekale
Computer Department
Vivekanand Education Society's
Institute of Technology
Mumbai, India.
2021.shamal.dhekale@ves.ac.in

Bhagyashree Vaswani
Computer Department
Vivekanand Education Society's
Institute of Technology
Mumbai, India.
2021.bhagyashree.vaswani@ves.ac.in

Chandni Gangwani
Computer Department
Vivekanand Education Society's
Institute of Technology
Mumbai, India.
2021.chandni.gangwani@ves.ac.in

Rupali Soni
Computer Department
Vivekanand Education Society's
Institute of Technology
Mumbai, India
rupali.hande@ves.ac.in

***Abstract*—Waste, particularly plastic waste, poses a major challenge to environmental sustainability, overwhelming current management systems. The integration of advanced object detection technology offers a solution by accurately detecting and quantifying plastic waste from community-reported videos or photos of dumpsters. This data-driven approach enables NGOs to optimize resource allocation for waste collection and segregation, significantly improving efficiency and reducing operational bottlenecks. Additionally, the system promotes environmental engagement through features such as event announcements, educational content, and sustainability news, encouraging active community participation. Key findings highlight improvements in NGO operations, reduced waste collection times, and increased community involvement in sustainability efforts. The implications of this technology extend to scalable solutions for both urban and rural waste management, offering an innovative approach to addressing plastic waste challenges. By merging machine learning with community action, this solution bridges the gap between technology and environmental stewardship, empowering both NGOs and communities to work collaboratively towards effective waste management.**

Keywords: Waste Management, Plastic Waste, Object Detection, Community Engagement, Sustainability Practices

I. Introduction

According to [1], annual global waste generation currently stands at approximately 2.1 billion tonnes and is projected to increase significantly, reaching 3.4 billion tonnes by 2050. India, with 17.7% of the global population and just 2.4% of the world's surface area, contributes approximately 12% to this total. This staggering volume of waste is driven by rapid urbanization, population growth, economic development, and changing lifestyles. Unfortunately, more than 33% of all generated waste is not disposed of in an environmentally sound manner, leading to widespread illegal dumping, water and air pollution, and the degradation of land. These practices not only harm ecosystems but also impede the sustainable growth of urban areas and communities, creating significant environmental and health risks.

Developing countries are particularly vulnerable to the challenges of managing increasing levels of waste due to a lack of infrastructure, planning, and advanced technologies necessary for effective solid waste management (SWM). Despite efforts to develop smart cities and promote sustainable urban development, these regions continue to face significant challenges in handling waste properly, as highlighted in [2]. Inadequate waste management contributes to environmental hazards, illegal dumping, and health risks for local populations. However, as technology progresses, there is a growing recognition of the potential for machine learning (ML), deep learning (DL), and the Internet of Things (IoT) to revolutionize waste detection, classification, and recycling processes. These technologies offer opportunities to better forecast, collect, and process waste, reducing both health risks and environmental damage.

"EnviroScan: Community and NGO Waste Solution" aims to leverage these advancements, particularly YOLOv9 for object detection, to improve waste management at the community level. The project focuses on the role of NGOs and community engagement in addressing the growing issue of waste, particularly plastic waste, which is a major contributor to environmental degradation. By incorporating machine learning models for precise waste detection from videos or images submitted by the community, EnviroScan helps optimize waste collection, resource allocation, and segregation processes. This data-driven approach enables NGOs to reduce operational bottlenecks and streamline their waste management practices, resulting in more efficient and timely waste collection efforts.

The system also promotes environmental stewardship by engaging the community with features such as event announcements, educational content, and sustainability news. These elements encourage active participation from residents and foster a collective sense of responsibility toward environmental conservation. By integrating technology with community action, EnviroScan strengthens the connection between NGOs and the local population, enhancing collaboration for effective waste management. This approach not only benefits urban areas but is also scalable for use in rural settings, offering a comprehensive solution to waste management challenges across diverse geographic regions.

EnviroScan's integration of machine learning with community-driven efforts offers a sustainable and innovative solution to the plastic waste crisis. This collaboration between communities and NGOs not only improves waste management operations, with reduced collection times and increased community involvement, but also fosters a shared commitment to a cleaner, more sustainable future. The following sections will elaborate on the methodology, experimental results, and potential future directions, all of which hold great promise for transforming waste management through innovative, tech-driven solutions.

II. Related work

In recent years, object detection technologies have played a transformative role in waste management systems, enabling more effective classification and monitoring of waste. Patil et al. [3] developed an Android-based application utilizing a YOLO framework for identifying electronic waste, thus streamlining the process of recognizing discarded devices and their components for appropriate recycling. In another work, Karkar and Al-Maadeed [4] implemented a Faster R-CNN-based approach to differentiate between plastic waste and disposable diapers, attaining an accuracy of 91.2%.

A different study [5] employed a Single Shot Detector (SSD) alongside clustering methods to automatically distinguish between biodegradable and non-biodegradable waste, achieving a high mean average precision (mAP) of 0.965. Similarly, models based on EfficientNet have demonstrated high accuracy in waste classification tasks, with one study reporting 98% accuracy on the TrashNet dataset [6]. Another approach, YOLO-Green [7], introduced a lightweight real-time detection system capable of identifying seven waste categories and recorded a mAP of 78.04%.

Furthermore, drone-integrated waste detection systems [8] have shown promise in enabling real-time multi-object detection for segregation tasks, highlighting the synergy between image processing and deep learning in environmental monitoring. The need for scalable and cost-effective urban cleanliness monitoring has also driven innovation. Recent systems incorporate smart edge-based infrastructure on moving vehicles, equipped with cameras and edge processors, to capture and analyze street-level images in real-time. These systems leverage deep learning techniques to detect and categorize street litter, showcasing their value in smart city environments [19].

Advancements in dataset creation have also contributed significantly. The study in [9] addresses the limitation of single-object datasets by proposing Trash-Fusion, which combines multiple waste types in complex backgrounds using image fusion techniques. Alongside the Trash-Collect dataset, this enabled convolutional neural networks to be trained efficiently, achieving real-time performance of 60 FPS using YOLOv5. Another notable study [10] introduced a novel dataset composed of 3,192 urban waste images captured via Google Earth. A location-aware keypoint network was then applied, outperforming conventional detection systems with a recall of 71.8% and an average precision of 44.0%.

An extensive survey on deep learning applications in waste management was presented in [11], reviewing over twenty benchmarked datasets and numerous model architectures for waste detection and classification. For example, [12] proposed a real-time garbage detection method using small UAVs for nature reserves, reaching 91.34% accuracy with a YOLOv4 model. Deep learning models have also been explored for classifying e-waste via mobile applications, achieving classification accuracies ranging from 90% to 97% [13].

Moreover, a hybrid classification approach combining three pre-trained CNNs was proposed in [14], leading to classification accuracies of 96.5% and 94%, significantly improving traditional classification models. The relevance of automatic waste classification in sustainable recycling processes is further explored in [15], where the authors propose a "double fusion" strategy. This method combines early and late fusion techniques across deep learning models to improve classification performance, achieving a 3.58% boost over existing state-of-the-art systems.

In [16], deep convolutional neural networks (DCNNs) with four and five layers were utilized to classify different types of waste, achieving a maximum accuracy of 70% with the five-layer variant. Another notable contribution [17] introduced a smart waste sorting solution employing deep neural networks, which demonstrated outstanding performance on the TrashNet and VN-trash datasets with accuracies of 94% and 98%, respectively. A broad review of machine learning techniques in municipal solid waste management was presented in [18], analyzing more than 200 publications to pinpoint key challenges and opportunities for future research in sustainable waste solutions.

Lastly, research works such as [19] and [20] emphasize the impact of urban cleanliness and propose AI-driven methods for assessing street litter levels using mobile edge computing. These include systems that capture high-resolution images from moving vehicles and analyze them using deep learning models like Faster R-CNN, providing real-time feedback for optimal resource allocation in street cleaning operations.

III. Proposed Solution

EnviroScan is a mobile application developed to address the persistent issue of unmanaged waste by connecting the general public with NGOs involved in environmental cleanup. The system empowers users to report garbage-ridden or unclean areas by submitting geotagged photographs of those locations. This not only allows NGOs to identify problem spots and dispatch cleanup teams efficiently, but also plays a critical role in our research objective—crowdsourcing complex, real-world waste data to train and improve our object detection model.

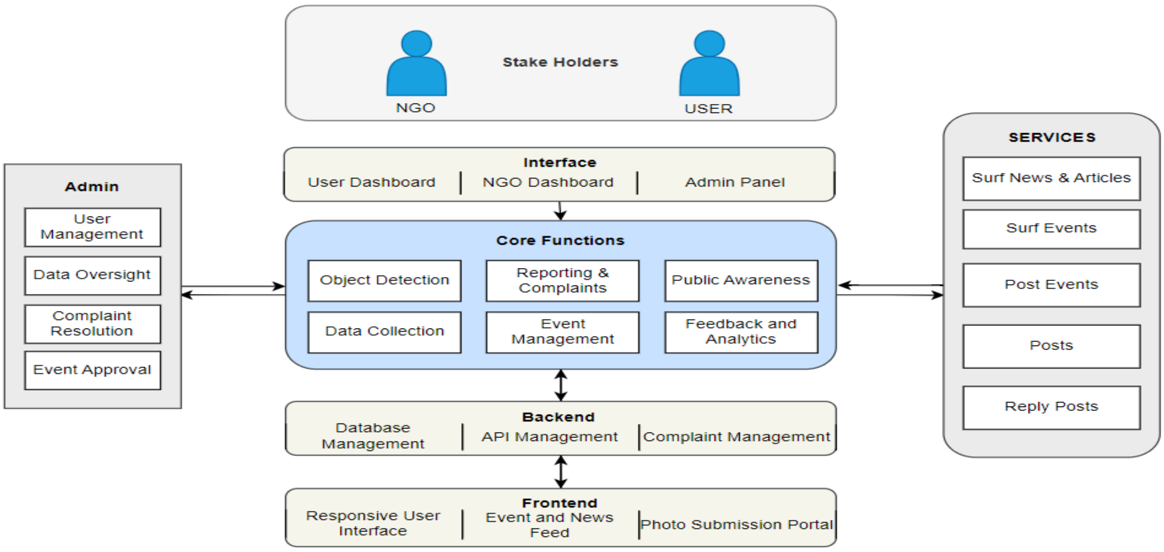

Fig 1. Overview of the System

An overview of the system architecture is illustrated in Fig. 1, which outlines the interaction flow between users and NGOs. Users submit geotagged images through the app interface, which are stored in a centralized database. NGOs access these reports via a dedicated dashboard to plan clean-up drives. Simultaneously, the backend infrastructure integrates this data stream with our waste classification pipeline for ongoing model enhancement.

To detect various waste materials within these submitted images, we have developed a custom object detection model based on the YOLOv12 architecture. According to recent research by Alif and Hussain [24], YOLOv12 introduces a series of significant advancements tailored for real-time object detection. These include a Residual Efficient Layer Aggregation Network (R-ELAN) that improves feature extraction and gradient flow, 7×7 separable convolutions that preserve spatial context with fewer parameters, and a FlashAttention-powered area-based attention mechanism that enhances focus on salient regions, particularly in cluttered scenes. These enhancements not only optimize the model for accurate detection of small, partially obscured, or overlapping objects, but also reduce inference latency, making it suitable for edge deployment and mobile processing environments.

To train the object detection model effectively, a diverse and representative dataset was used that composed of both controlled and real-world images:

A. Single Object Detection:

This consists of 1000 images taken from the TrashNet dataset, where each image contains a single type of waste object, making them ideal for clear and focused object identification.

B. Multiple Object Detection:

This set includes 65 images collected through real-time photography and internet sources like Google. These images depict real-world waste scenes with multiple and overlapping objects, offering more complex detection scenarios.

C. Combined Dataset:

A merged dataset containing a total of 1065 images, combining both single and multiple object images. Unlike standard bounding boxes, polygonal annotations were used for greater precision, especially for irregularly shaped waste items.

To optimize the training process and enhance the model's reliability, a series of preprocessing steps were applied to the image dataset. Each image was resized to a fixed dimension of 640×640 pixels to maintain consistency in input size across the model. To increase variability and prevent overfitting, data augmentation methods such as random rotation, changes in brightness, and horizontal flipping were implemented. Finally, the dataset was methodically divided into training, validation, and test sets to ensure thorough evaluation and balanced model performance.

The use of a combined dataset, particularly the inclusion of real-world images with overlapping waste objects, significantly improves the model's ability to recognize garbage in complex settings. Furthermore, the EnviroScan application acts as a continuous data collection platform. As users upload more geotagged waste images via complaints, the dataset can be regularly updated and expanded. This not only enhances the model's real-world performance but also builds a sustainable loop of improvement through real-time community input.

Through this system, we aim to combine advanced AI capabilities with grassroots participation, paving the way for smarter and more responsive waste management practices.

IV. Implementation and Results

Initially, a model was trained using single waste object images through Roboflow's default object detection framework (Roboflow 3.0 Fast). This approach achieved a notable mAP@50 of 93.4% and demonstrated strong performance on isolated waste items such as plastic, paper, cardboard, glass, metal, and trash. Roboflow 3.0 provides a user-friendly, end-to-end platform for dataset management, model training, and deployment, making it ideal for rapid prototyping and real-time applications. Its default model is Convolutional Neural Network-based (CNN) and built on pre-trained architectures like YOLO, optimized for quick inference and deployment through a no-code interface.

Similar to findings in research by Kohut et al. [21], the model performed well on clean, single-object scenarios but showed limitations when applied to complex real-world environments. In our case, its effectiveness significantly declined in dumping ground scenes where multiple waste objects often overlap, resulting in inaccurate detections and misclassifications. These observations emphasized the need for more advanced models capable of handling dense and cluttered waste scenarios.

To address these challenges, the training approach was adapted to include images featuring multiple, overlapping waste items. Complex scenes were annotated using polygonal masks to achieve greater labeling precision. After annotation, the images underwent preprocessing and various augmentation techniques. This approach provided the model with more representative, real-world scenarios, significantly boosting its capability to identify and categorize several objects in a single image.

Following this, the effectiveness of three state-of-the-art object detection models was systematically assessed:

A. RF-DETR (Region-Friendly Detection Transformer):

RF-DETR (Region-Friendly Detection Transformer) is an enhanced version of the original DETR (DEtection TRansformer) model introduced by Carion et al. [22], which pioneered the use of transformer architectures for object detection in a fully end-to-end manner—eliminating the need for anchor boxes and non-maximum suppression. RF-DETR extends this framework by introducing region-focused improvements aimed at addressing challenges in detecting small or tightly packed objects, areas where traditional DETR often underperformed due to its broad attention across spatial dimensions. Through refined attention mechanisms and upgraded positional encoding, RF-DETR delivers better accuracy in both object localization and classification, making it particularly well-suited for detecting waste in complex, cluttered scenes like landfills and dumping sites.

B. YOLO-NAS (You Only Look Once – Neural Architecture Search):

YOLO-NAS (You Only Look Once – Neural Architecture Search) is a convolutional neural network (CNN)-based object detection model developed by Deci AI. It is optimized through Neural Architecture Search (NAS), which automates the process of selecting the optimal architecture to balance accuracy, speed, and efficiency. Unlike traditional YOLO models that depend on manually designed architectures, YOLO-NAS generates lightweight and powerful models by automating this selection. The model is particularly effective for real-time applications and edge devices, providing quantization support and latency-aware optimization to perform well in low-resource environments. According to a study by Ali et al. [1], YOLO-NAS has proven its effectiveness in real-world applications, showcasing its adaptability and reliability in handling complex object detection tasks. In our specific case, YOLO-NAS demonstrated strong performance in cluttered waste environments, making it a suitable choice for practical deployments.

C. YOLOv12:

YOLOv12 is a recent and advanced object detection model in the YOLO series, based on Convolutional Neural Networks (CNN). It is designed to deliver high-speed, accurate performance in real-time applications. According to research by Alif et al. [24], YOLOv12 introduces several key innovations, including the R-ELAN backbone to enhance feature fusion, 7×7 depth wise separable convolutions that increase the receptive field while reducing computational costs, and a combination of Flash Attention and Area-Based Attention mechanisms to improve focus on small and overlapping objects. It also uses an anchor-free detection method, simplifying the model architecture and improving object localization. While the model's mAP@50 was slightly lower than YOLO-NAS in our tests, YOLOv12 excelled in detecting multiple waste objects and adapting to real-world waste segregation challenges. Its lightweight CNN-based structure makes it ideal for deployment on mobile and edge devices, providing an effective combination of speed, accuracy, and efficiency.

TABLE I. A COMPARATIVE ANALYSIS OF OBJECT DETECTION MODELS.

Feature	Roboflow 3.0 (Default)	YOLOv12	YOLO-NAS	RF-DETR
Target Scenario	Single waste images	Multiple waste images	Multiple waste images	Multiple waste images
Architecture	CNN-based	CNN-based	YOLO + NAS optimization	DETR Transformer + Region-aware
Accuracy (mAP@50)	93.5%	79.7%	81.8%	78.8%
Performance on Overlapping Objects	Low	Better	Good	Good

The comparison shown in Table I highlights that while Roboflow excels in single-object detection, YOLOv12 and YOLO-NAS are more suitable for real-world scenes with overlapping waste, with YOLOv12 offering a strong balance of speed, accuracy, and robustness.

Subsequently, the single and multiple waste image datasets were merged, forming a unified dataset comprising 1,065 images. Following data augmentation and pre-processing, this combined dataset was used to train the YOLOv12 model. The model demonstrated exceptional performance on both isolated and overlapping waste instances, achieving a mAP@50 of 97.5%, Precision of

97.4%, and Recall of 94.2%. Visual inspections of the predictions confirmed accurate localization and classification, even in cluttered scenes. The model also exhibited strong generalization across varied lighting conditions and backgrounds. Additionally, YOLOv12 maintained efficient inference speed, making it suitable for real-time deployment in practical waste management scenarios.

The performance graph below demonstrates that the model achieved stable accuracy within the first 100 epochs, with minimal variation observed thereafter.

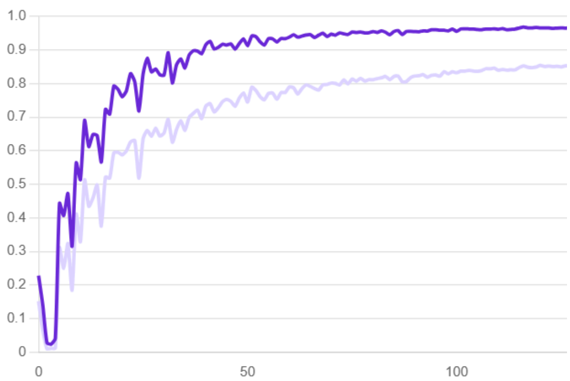

Fig 2. Number of Epochs vs Accuracy

Fig. 3 below illustrates YOLOv12's ability to accurately detect and classify multiple overlapping waste objects within a real-world test image.

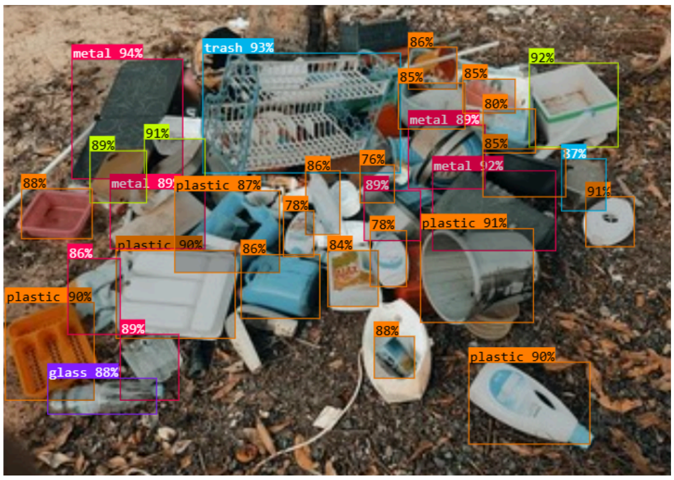

Fig 3. Model Prediction on Testing Image

The screenshots below showcase two key functionalities of the EnviroScan mobile application. The first interface allows users to lodge complaints by uploading images of unclean areas along with their geolocation, ensuring that the exact location of waste accumulation is accurately recorded. The second interface displays upcoming NGO-led environmental events and clean-up drives, allowing users to register as volunteers and stay informed.

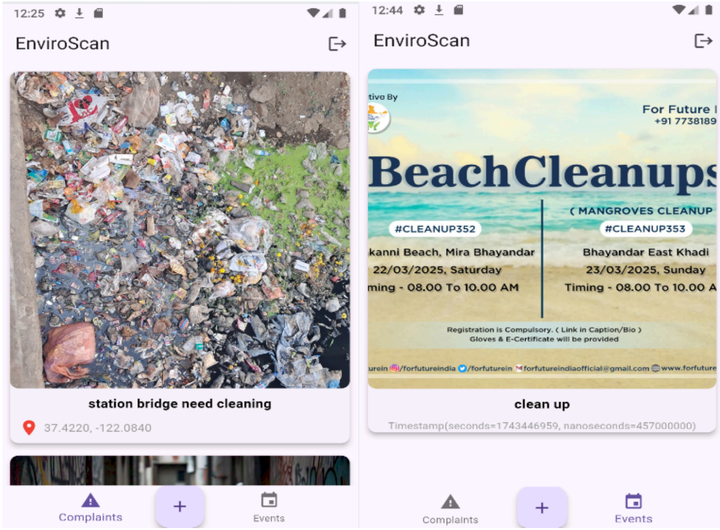

Fig 4. Application User Interface Snapshots

Through this application, we not only enable real-time waste reporting but also create an opportunity to continuously gather diverse, geotagged images of waste. When a user uploads a complaint image, it is automatically stored in the backend and would be added to our training dataset. This allows the number of images in our dataset to grow over time, helping the model adapt to newer waste patterns and real-world complexities. We aim to make this image collection and model improvement process dynamic and self-sustaining, ensuring that the system evolves continuously with real-world input.

V. Conclusion and Future Work

EnviroScan represents a significant advancement in bridging grassroots involvement with AI-driven waste detection. By training the YOLOv12 model on a balanced dataset of single and overlapping waste objects, the system achieved a high mAP@50 of 97.5%, showcasing its reliability in real-world conditions. The mobile app empowers users to report waste locations through geotagged images, while NGOs receive structured, actionable insights for planning clean-up operations. This combination of deep learning and civic engagement lays the groundwork for a scalable and responsive waste management framework.

Moving forward, a major focus will be on integrating the trained YOLOv12 model directly into the EnviroScan application for real-time, on-device waste detection and classification. This will allow users to receive instant feedback while submitting reports, enhancing interactivity and accuracy. Future updates also include continuous learning from new user-submitted images, severity-based prioritization of complaints and multilingual support to encourage broader community participation. These advancements aim to evolve EnviroScan into a smarter, more inclusive, and self-improving environmental reporting platform.

References

[1] United Nations Environment Programme. "Global Waste Management Outlook 2024." UNEP, 2024.

[2] G. Vinti and M. Vaccari, "Solid Waste Management in Rural Communities of Developing Countries: An Overview of Challenges

and Opportunities," Clean Technologies, vol. 4, no. 4, pp. 1138–1151, Nov. 2022.

[3] K. Alfatmi, F. S. Shinde, M. Shahade, S. S. Sharma, S. S. Aruja and T. Y. Chaudhari, "E-Safe: An E-waste Management and Awareness Application using YOLO Object Detection," 2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2023, pp. 1061-1066.

[4] J. Rashida, R. Hamzah, K. A. Fariza Abu Samah and S. Ibrahim, "Implementation of Faster Region-Based Convolutional Neural Network for Waste Type Classification," 2022 International Conference on Computer and Drone Applications (IConDA), Kuching, Malaysia, 2022, pp. 125-130.

[5] Karthikeyan, M., Subashini, T.S. & Jebakumar, R. SSD based waste separation in smart garbage using augmented clustering NMS. Autom Softw Eng 28, 17 (2021).

[6] A. Masand, S. Chauhan, M. Jangid, R. Kumar and S. Roy, "ScrapNet: An Efficient Approach to Trash Classification," in IEEE Access, vol. 9, pp. 130947-130958, 2021.

[7] W. Lin, "YOLO-Green: A Real-Time Classification and Object Detection Model Optimized for Waste Management," 2021 IEEE International Conference on Big Data (Big Data), Orlando, FL, USA, 2021, pp. 51-57.

[8] P. Saraswat and S. Lohia, "Smart Waste Management: Waste Segregation Using Machine Learning," on ResearchGate, 2023.

[9] Xiao, Y., Chen, B., Feng, C. et al. Recyclable solid waste detection based on image fusion and convolutional neural network. J Mater Cycles Waste Manag 26, 2043–2057 (2024).

[10] H. Li, C. Hu, X. Zhong, C. Zeng and H. Shen, "Solid Waste Detection in Cities Using Remote Sensing Imagery Based on a Location-Guided Key Point Network With Multiple Enhancements," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 16, pp. 191-201, 2023.

[11] H. Abdu and M. H. Mohd Noor, "A Survey on Waste Detection and Classification Using Deep Learning," in IEEE Access, vol. 10, pp. 128151-128165, 2022.

[12] Piotr Nowakowski, Teresa Pamuła, Application of deep learning object classifier to improve e-waste collection planning, Waste Management, Volume 109, 2020, Pages 1-9, ISSN 0956-053X.

[13] Chang, X. Chen, and J. Wu, "Energy-aware virtual machine placement for cloud data centers," Concurrency and Computation: Practice and Experience, vol. 29, no. 5, 2017.

[14] K. Ahmad, K. Khan and A. Al-Fuqaha, "Intelligent Fusion of Deep Features for Improved Waste Classification," in IEEE Access, vol. 8, pp. 96495-96504, 2020.

[15] Altikat, A., Gulbe, A. & Altikat, S. Intelligent solid waste classification using deep convolutional neural networks. Int. J. Environ. Sci. Technol. 19, 1285–1292 (2022).

[16] A. H. Vo, L. Hoang Son, M. T. Vo and T. Le, "A Novel Framework for Trash Classification Using Deep Transfer Learning," in IEEE Access, vol. 7, pp. 178631-178639, 2019.

[17] Xia W, Jiang Y, Chen X, Zhao R. Application of machine learning algorithms in municipal solid waste management: A mini review. Waste Management & Research. 2022;40(6):609-624.

[18] Ping, Ping & Xu, Guoyan & Kumala, Effendy & Gao, Jerry. (2020). Smart Street Litter Detection and Classification Based on Faster R-CNN and Edge Computing. International Journal of Software Engineering and Knowledge Engineering. 30. 537-553. 10.1142/S0218194020400045.

[19] P. Zhang, Q. Zhao, J. Gao, W. Li and J. Lu, "Urban Street Cleanliness Assessment Using Mobile Edge Computing and Deep Learning," in IEEE Access, vol. 7, pp. 63550-63563, 2019.

[20] G. Salvia, N. Zimmermann, C. Willan, J. Hale, H. Gitau, K. Muindi, E. Gichana, and M. Davies, "The wicked problem of waste management: An attention-based analysis of stakeholder behaviours," Journal of Cleaner Production, vol. 326, p. 129200, 2021.

[21] Kohut, N., Basystiuk, O., Shakhovska, N., & Melnykova, N. (2025). Detecting plant diseases using machine learning models. Sustainability, 17(1), 132.

[22] Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., Zagoruyko, S. (2020). End-to-End Object Detection with Transformers. In: Vedaldi, A., Bischof, H., Brox, T., Frahm, JM. (eds) Computer Vision – ECCV 2020. ECCV 2020. Lecture Notes in Computer Science(), vol 12346. Springer, Cham.

[23] Ali, A., Zaman, A., Abbas, M., & Shabbir, M. (2024). Firearm Detection Using YOLO-NAS LMS in Advanced Surveillance System. ResearchGate.

[24] M. Alif, M. N. Morshed, and M. A. Rahman, "YOLOv12: A Breakdown of the Key Architectural Features," arXiv preprint arXiv:2502.14740, 2024.