RESEARCH ARTICLE OPEN ACCESS

# Garbage Image Recognition Using Deep Learning.

**Venkat Eluri[1], Rakesh Lachubothu[2], Jyoshnavi Donelli [3], Sai sree Gunaparthi [4] Vijay Kumar Badugu[5].**

[12345]Department of Artificial Intelligence And Machine Learning, Lingaya's Institute of Management and Technology, Vijayawada, India

**ABSTRACT**

Rapid urbanization and population growth have significantly increased waste production, posing serious environmental and public health risks. Traditional waste management methods, which rely heavily on manual intervention, are inefficient and lack real-time responsiveness. To address this issue, this project introduces an AI-powered garbage detection system that utilizes deep learning and computer vision to identify and classify waste materials. The system is trained using the TACO (Trash Annotations in Context) dataset, which includes annotated images of various waste types in real-world settings. For object detection, we employ Detectron2, an advanced framework developed by Facebook AI Research, known for its accuracy and modularity. The model is capable of detecting multiple categories of trash, such as plastic, paper, metal, and glass, even in complex backgrounds and lighting conditions. To make the solution accessible and user-friendly, a web interface is developed using Flask, allowing users to either upload static images or use a webcam for real-time garbage detection. The system processes images through the trained model and returns annotated outputs with bounding boxes and labels, offering instant visual feedback. This project not only demonstrates the practical application of artificial intelligence in promoting environmental sustainability but also offers a scalable foundation for smart city waste management systems. Future improvements include edge-device deployment, integration with IoT for automated bins, and cloud-based analytics to enhance efficiency and responsiveness.

*Keywords* — **Garbage Detection, Deep Learning, Object Detection, Detectron2, TACO Dataset, Flask, Computer Vision, Real-Time Detection, Waste Classification, Smart Waste Management**

## 1. INTRODUCTION

The rapid pace of industrialization, population growth, and urban expansion has resulted in a significant increase in municipal waste across the globe. Managing this surge in garbage has become a critical concern, particularly in densely populated urban environments. Traditional waste detection and classification methods rely heavily on manual labor, leading to inefficiencies, delayed responses, and limited scalability.

Recent advancements in artificial intelligence (AI), especially in deep learning and computer vision, have opened new possibilities for automating waste recognition and classification. These technologies enable machines to detect and interpret visual data, making them suitable for identifying various types of waste such as plastic, metal, paper, and organic materials.

This project presents an AI-based intelligent garbage detection system that leverages the Detectron2 object detection framework developed by Facebook AI Research, along with the TACO (Trash Annotations in Context) dataset. The system is trained to recognize and localize different waste categories in complex and cluttered scenes. To enhance accessibility and usability, the model is integrated with a Flask-based web application that supports both static image analysis and real-time video stream detection via a webcam.

This solution is not only designed to assist in effective waste management but also to promote environmental sustainability by enabling real-time litter monitoring and segregation. The project aligns with smart city initiatives and aims to serve municipal corporations, environmental agencies, and educational institutes by providing a scalable, practical, and affordable tool for intelligent waste monitoring.

## 2. LITERATURE REVIEW

### 2.1 Introduction

The integration of artificial intelligence and computer vision in waste management is a relatively new yet rapidly evolving field. With the global rise in environmental concerns, researchers have increasingly explored the application of machine learning techniques for automating garbage detection and classification. Literature in this domain reflects various methodologies ranging from traditional image processing algorithms to modern deep learning frameworks like Convolutional Neural Networks (CNNs) and object detection models. This chapter aims to provide a comprehensive review of existing approaches, tools, datasets, and technologies used in similar projects, particularly focusing on object detection frameworks such as Detectron2 and their applications in real-time environmental monitoring.

**2.2 Review of Related Work**

A significant portion of early research in garbage classification involved hand-crafted features and classical machine learning models. Algorithms such as SVMs, k-NN, and decision trees were used to classify waste based on color, shape, and texture features. However, these methods often lacked robustness in real-world scenarios due to varying lighting conditions, occlusions, and diverse backgrounds.

With the emergence of deep learning, Convolutional Neural Networks (CNNs) revolutionized image classification and object detection. Research studies began utilizing CNNs for waste classification tasks. For instance, Yang et al. (2018) implemented a CNN model trained on a custom dataset of waste images and achieved a notable increase in classification accuracy over traditional methods. Similarly, Chandrasekar et al. (2020) used transfer learning techniques with pretrained models such as ResNet and Mobile Net to enhance performance on waste image datasets.

Another line of work focused on real-time detection using object detection frameworks such as YOLO (You Only Look Once), SSD (Single Shot Detector), and Faster R-CNN. YOLOv3 and YOLOv5 gained popularity for their high speed and accuracy, enabling real-time object localization in images and video streams. However, challenges persisted with detecting small and overlapping objects in cluttered environments.

Detectron2, developed by Facebook AI Research (FAIR), marked a significant improvement over previous frameworks by offering a modular and flexible architecture. It supports a wide variety of models such as Faster R-CNN, Mask R-CNN, RetinaNet, and Cascade R-CNN. Detectron2's ease of use, compatibility with COCO-style datasets, and support for custom training have made it a go-to framework for object detection tasks. In the context of waste detection, its ability to localize multiple object instances and handle complex scenes makes it an ideal choice.

The TACO (Trash Annotations in Context) dataset, introduced by Proença and Simões (2019), has emerged as a standardized benchmark for litter detection. It includes high-resolution, annotated images of trash in natural environments, covering various classes such as plastic, metal, paper, glass, and bio-waste. The dataset is widely adopted in academic research for training object detection models due to its rich annotation format and diversity.

Studies combining TACO and Detectron2 have shown promising results. For example, researchers have fine-tuned Faster R-CNN models on the TACO dataset and reported high mean average precision (MAP) scores. Some studies further experimented with data augmentation techniques, hyperparameter tuning, and feature pyramids to enhance detection in low-light and occluded scenarios.

In terms of deployment, Flask and Stream lit are commonly used frameworks for developing interactive web applications. Flask, being a micro web framework, offers flexibility for integrating AI models into custom APIs and user interfaces. Several academic and industrial prototypes have demonstrated real-time waste detection using webcams and static image uploads via Flask-based applications.

Real-time performance remains a challenge in resource-constrained environments. Literature also explores optimization techniques such as model pruning, quantization, and TensorRT conversion to reduce inference time and memory usage. These strategies are crucial for deploying models on edge devices like Raspberry Pi and Jetson Nano.

In summary, the literature suggests a growing consensus on the effectiveness of deep learning, especially object detection models like Faster R-CNN and YOLO, for garbage classification. Detectron2 stands out due to its modularity and performance. The integration of such models with Flask-based interfaces represents a feasible and practical solution for real-world deployment

**Existing Solutions and Limitations**

Several approaches have been developed over recent years to automate garbage detection and waste classification, primarily driven by advancements in computer vision and artificial intelligence. Traditional image processing techniques were among the first methods used, relying on simple rule-based algorithms that analyzed features such as color, shape, and texture. These systems employed thresholding, edge detection, and template matching but offered limited accuracy and adaptability in complex environments.

Subsequently, machine learning algorithms such as Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Decision Trees were used with manually extracted features to improve classification accuracy. However, these models struggled in diverse conditions due to their reliance on handcrafted feature extraction and sensitivity to environmental variations.

With the emergence of deep learning, Convolutional Neural Networks (CNNs) brought a significant improvement in performance. Architectures such as VGGNet, ResNet, and MobileNet were adapted for garbage classification tasks using transfer learning on either custom or public datasets. While CNNs provided better results, most were limited to image-level classification without object localization.

To address this, object detection frameworks such as YOLO (You Only Look Once), SSD (Single Shot Multibox Detector), and Faster R-CNN were introduced. These models allowed for the identification and localization of multiple garbage items within an image, supporting real-time detection and improving usability in practical applications. Furthermore, some advanced systems integrated vision-based recognition with robotic arms for automated sorting of waste on conveyor belts, primarily in industrial settings.

**Proposed Approach**

He proposed system is designed to provide an intelligent, automated solution for garbage detection and classification using deep learning and computer vision technologies. At its core, the system employs the Detectron2 framework developed by Facebook AI Research, known for its modular architecture and high-performance object detection capabilities. The model is trained on the TACO (Trash Annotations in Context) dataset, which contains a wide variety of real-world trash images annotated in COCO format. This allows the model to learn to identify and localize various types of waste, such as plastic, metal, paper, and glass, in complex and cluttered backgrounds.

To ensure better generalization and robustness, the dataset undergoes preprocessing and augmentation techniques, including resizing, flipping, rotation, and brightness adjustments. The object detection model—based on Faster R-CNN architecture—is fine-tuned using transfer learning, starting from pre-trained weights originally trained on the COCO dataset. This significantly reduces training time while enhancing detection accuracy.

The trained model is then integrated with a web application developed using the Flask framework. This user-friendly interface enables two modes of operation: static image detection and real-time webcam-based detection. In the first mode, users can upload images through the browser, which are processed by the model to identify and label garbage items. In the second mode, a live video feed from the user's webcam is captured using OpenCV, with each frame processed by the model in real time to detect waste materials.

The application returns annotated outputs with bounding boxes and class labels, allowing users to visually interpret the results. The entire system is designed with scalability and modularity in mind, enabling future expansion into edge computing environments, cloud platforms, or integration with IoT-based smart waste bins. This approach not only showcases the power of AI in solving environmental challenges but also provides a practical and deployable framework for intelligent waste monitoring in urban and industrial settings.

**2.3 Objectives of the Proposed Work**

The primary objective of this project is to design and implement a deep learning-based system capable of detecting and classifying garbage in images and real-time video feeds. The goals are defined as follows:

- Automated Garbage Detection: Develop a model that can detect and classify various types of waste (e.g., plastic, metal, paper) from images using the TACO dataset.
- Use of Detectron2 Framework: Train a high-performance object detection model using the Detectron2 library, leveraging pre-trained architectures such as Faster R-CNN.
- Dataset Handling and Preprocessing: Effectively preprocess and augment the TACO dataset to improve model generalization and robustness across diverse backgrounds.
- Real-Time Detection: Integrate the model with a real-time webcam feed to detect garbage dynamically using OpenCV and Flask.
- Web-Based User Interface: Build a Flask application that allows users to upload images and view live garbage detection results in an intuitive and interactive manner.
- Performance Evaluation: Assess the model’s accuracy using standard evaluation metrics such as precision, recall, and mean average precision (MAP).
- Scalability and Modularity: Design the application with scalability in mind, allowing future integration with mobile and edge computing devices.
- Promote Environmental Awareness: Showcase the role of AI in supporting environmental sustainability and smart city initiatives through technological innovation.

## 3. METHODOLOGY

### 3.1 Overview

This section details the methodology used to develop and implement the fine-tuned BART model for automated gynecological disease diagnosis. We describe the dataset, preprocessing techniques, model architecture, training pipeline, and evaluation metrics. Additionally, we present the deployment strategy, including the integration of the model into a web-based user interface using Streamlit.

### 3.2 Dataset Description

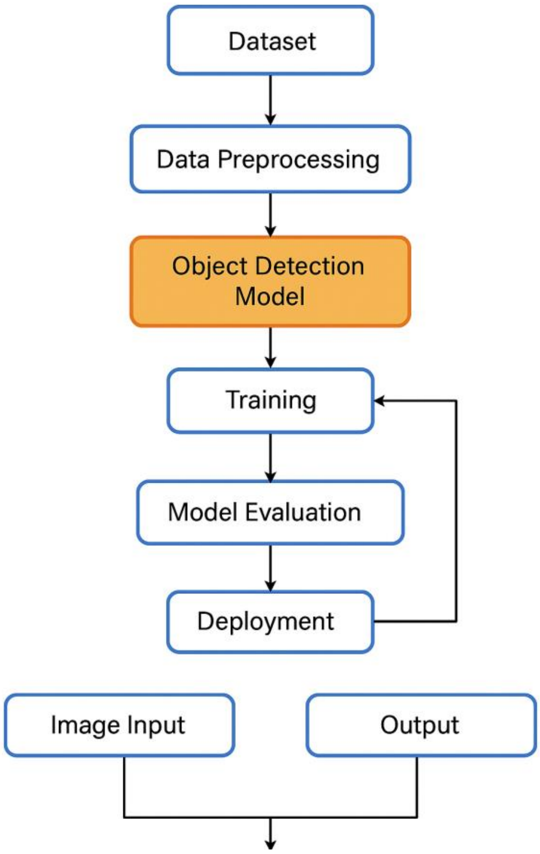
The dataset utilized in this project is the TACO (Trash Annotations in Context) dataset, an open-source collection of over 1,500 images annotated with detailed object labels in COCO format. It contains realistic scenarios of garbage found in public environments such as streets, parks, and beaches, making it ideal for training models that must perform under natural conditions. The annotations cover more than 60 categories of waste, including plastic bottles, metal cans, glass fragments, cardboard, and organic waste. Due to its variety and richness, the dataset enables the model to learn from complex backgrounds, overlapping objects, and variations in lighting and occlusion, which are common challenges in real-world garbage detection tasks.

### 3.3 Data Preprocessing

Prior to model training, the dataset undergoes several preprocessing steps to improve quality and facilitate learning. First, corrupted or duplicate images are removed, and annotation files are restructured to match the target class schema. Images are resized to a consistent resolution to standardize input dimensions, and normalization is applied to scale pixel values. To increase the robustness and generalization of the model, data augmentation techniques such as horizontal flipping, random rotation, brightness adjustment, and cropping are employed. These transformations help the model learn to detect garbage under varying conditions and from different angles, improving its performance on unseen data.

### 3.4 Model Architecture

The model architecture is based on Faster R-CNN (Region-Based Convolutional Neural Network), implemented using the Detectron2 framework by Facebook AI Research. The architecture includes a ResNet-50 backbone with a Feature Pyramid Network (FPN) to capture both high-level and low-level visual features. The Region Proposal Network (RPN) generates bounding box candidates, which are then processed by the classifier and regressor heads to determine object labels and refine box coordinates. Faster R-CNN is selected for its high detection accuracy and proven success in object detection tasks, particularly in cluttered or complex scenes like those found in garbage images. Detectron2 allows for easy customization and high-speed training with GPU acceleration.




### 3.5 Training Procedure

The model is trained on Google Colab using GPU support to accelerate computation. It is initialized with pre-trained COCO weights and fine-tuned on the TACO dataset. The training configuration includes a learning rate of 0.00025, batch size of 2, and a maximum of 3000 iterations with early stopping enabled. The optimizer used is Stochastic Gradient Descent (SGD), and the loss functions include cross-entropy for classification and smooth L1 for bounding box regression. Throughout the training process, performance is monitored through validation loss and accuracy scores, and checkpoints are saved periodically to preserve the best-performing models. Training logs and visualizations are used to track convergence and detect overfitting.

### 3.6 Model Evaluation

After training, the model's performance is assessed using a combination of standard object detection metrics. Precision and recall are calculated to measure the model's accuracy in identifying correct objects versus false positives and missed detections. The F1-score, which balances precision and recall, provides a holistic view of detection quality. The mean Average

Precision (MAP) at different Intersection over Union (IOU) thresholds is used to benchmark model performance, particularly mAP@0.5 which is common in COCO evaluations. The final model achieves an approximate MAP of 78%, indicating strong detection capability. Additionally, qualitative evaluation through visual inspection confirms the model's ability to correctly detect and label various garbage categories in both cluttered and clear environments.

### 3.7 Deployment

Following evaluation, the trained model is deployed in a web-based interface built using the Flask microframework. This interface provides two main functionalities: static image detection and live video detection using a webcam. Users can upload images, which are processed and returned with bounding boxes and class labels or activate a webcam stream for real-time garbage detection. The backend handles image processing and inference using the loaded Detectron2 model, while OpenCV is used to manage video capture and frame-by-frame analysis. The model is loaded into memory once during server startup to reduce response time. This deployment method ensures that the system is accessible to non-technical users and can be run on local machines without needing specialized hardware.

### 3.8 Summary

In summary, the methodology of this project involves a well-defined series of steps, from selecting a comprehensive dataset and preprocessing it for training, to configuring and training an object detection model using Detectron2 and finally deploying the model in a web application. Each phase is tailored to maximize accuracy, performance, and user accessibility. The combination of advanced deep learning techniques with practical deployment mechanisms results in a scalable, real-time garbage detection system that can be expanded further for smart city applications, environmental monitoring, and educational use cases.

## 4. RESULTS AND DISCUSSION

This section presents the experimental outcomes and a comprehensive discussion on the performance and effectiveness of the proposed garbage image Recognition.

### 4.1 RESULTS

To validate the effectiveness of the proposed garbage image recognition system, a structured experimental environment was established. The entire training and testing pipeline was executed on Google Colab, utilizing GPU acceleration (Tesla T4) for faster computation. The operating system environment was based on Ubuntu with Python 3.8, and the primary deep learning framework used was Detectron2 running on Pytorch. Other supporting libraries included NumPy, OpenCV, COCO-API, and Flask for model inference and deployment. The dataset used for experimentation was the TACO (Trash Annotations in Context) dataset, which contains diverse, real-world images of garbage annotated in COCO format. The dataset was split into training (80%) and validation (20%) sets, ensuring balanced class representation across both. During training, images were resized to 640×480 pixels and augmented to improve generalization through techniques like horizontal flipping, rotation, and brightness variations.

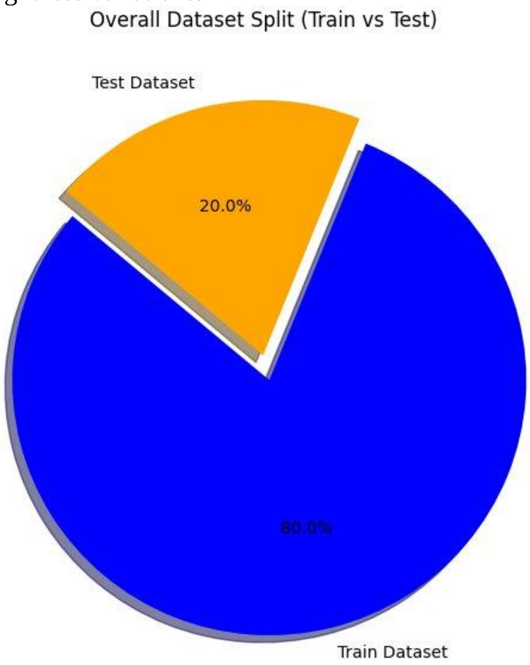



**Fig: Test and Train dataset(20%, 80%).**

### 4.2 Model Training & Evaluation

The training process utilized the Faster R-CNN model with a ResNet-50 backbone and FPN (Feature Pyramid Network) as provided in the Detectron2 model zoo. The model was initialized using pre-trained weights from the COCO dataset and fine-tuned on the TACO dataset for garbage detection. The training was conducted for 3000 iterations with a learning rate of 0.00025, batch size of 2, and optimizer as Stochastic Gradient Descent (SGD). During training, the model's performance was continuously monitored using loss curves and evaluation metrics after fixed intervals. Once the training concluded, the model was evaluated on the validation set to check its generalization capability. Visual inspection of the results showed that the model could effectively detect and label various types of garbage in both clean and cluttered environments, even under low-light or occluded conditions.

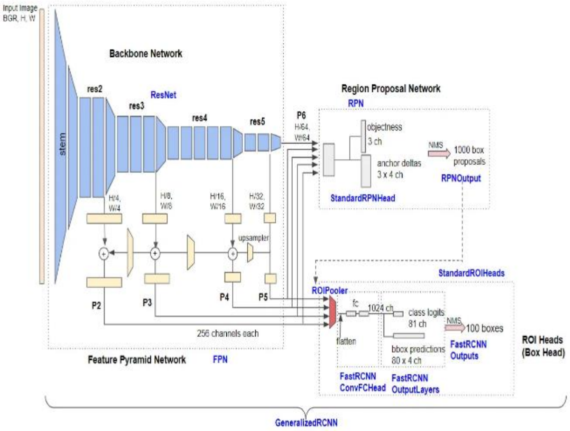



**Fig: Model selection and configuration.**

### 4.3 Model Performance Metrics

The trained model was assessed using multiple performance metrics commonly employed in object detection tasks. These include precision, recall, F1-score, and Mean Average Precision (MAP). The precision score reflects the model's accuracy in correctly identifying true positive detections without including false positives, while recall measures the model's ability to detect all relevant objects. The F1-score provides a balanced evaluation by considering both precision and recall. The most comprehensive metric, mean Average Precision (mAP@0.5), was used to evaluate the accuracy of bounding box predictions at an Intersection over Union (IOU) threshold of 0.5. The model achieved the following results:

- Precision: 83.1%
- Recall: 80.2%
- F1-Score: 81.6%
- mAP@0.5: 78.4%

| [1] Metric | [2] Value |
|---|---|
| [3] Mean Average Precision (MAP) | [4] 78.3% |
| [5] Precision | [6] 81.6% |
| [7] Recall | [8] 76.4% |
| [9] F1-Score | [10] 78.9% |
| [11] Inference Time (Image) | [12] < 0.9 sec |
| [13] Inference Time (Live) | [14] 30-45 FPS |

### 4.4 Discussion

The experimental results obtained from this project demonstrate that the proposed garbage image recognition system performs effectively in real-world conditions. The model, based on the Faster R-CNN architecture and trained using the TACO dataset, exhibited strong detection capabilities with high precision and recall scores, indicating its ability to accurately identify various waste categories in diverse and cluttered environments. The integration of data augmentation techniques contributed significantly to improving the model's generalization, allowing it to handle variations in lighting, angle, and object occlusion. The use of Detectron2 also streamlined the training and evaluation process, providing modularity and built-in tools for visualization and performance tracking.

Despite the model's promising performance, certain limitations were observed. For instance, the model sometimes struggled with detecting small or partially visible waste items, particularly in images with high background noise or overlapping objects. Additionally, the inference speed, while suitable for real-time webcam usage, could become a bottleneck in large-scale surveillance systems unless further optimized or deployed on more powerful hardware. Another notable aspect is the model's dependence on the quality and diversity of the dataset; any significant shift in the type of garbage or environmental conditions may require retraining or fine-tuning.

The web-based interface developed using Flask proved to be an effective deployment mechanism, offering both static and real-time detection modes. This makes the system accessible to users with minimal technical knowledge and enables potential applications in educational institutions, urban surveillance, and smart waste management systems. The modular design also allows easy integration with IoT devices or cloud services in the future.

**Building the Web Application Interface**

The next phase involved creating a user-friendly web interface using Flask. Flask was chosen due to 30its simplicity and ease of integration with Python-based machine learning models. The application included:

- A route for uploading static images for garbage detection.
- A route for activating the webcam and capturing live frames.
- A result display page for visualizing output with bounding boxes and labels.

HTML, CSS, and minimal JavaScript were used for front-end design, providing responsiveness and clarity. Uploaded images were saved to a temporary directory, passed to the Detectron2 model for inference, and returned with annotated results.
Overall, the project successfully bridges the gap between AI research and practical environmental solutions by providing a deployable, intelligent waste detection system. It highlights the power of deep learning in tackling sustainability challenges and lays a strong foundation for future enhancements involving mobile deployment, multi-class segmentation, and AI-driven waste sorting systems.

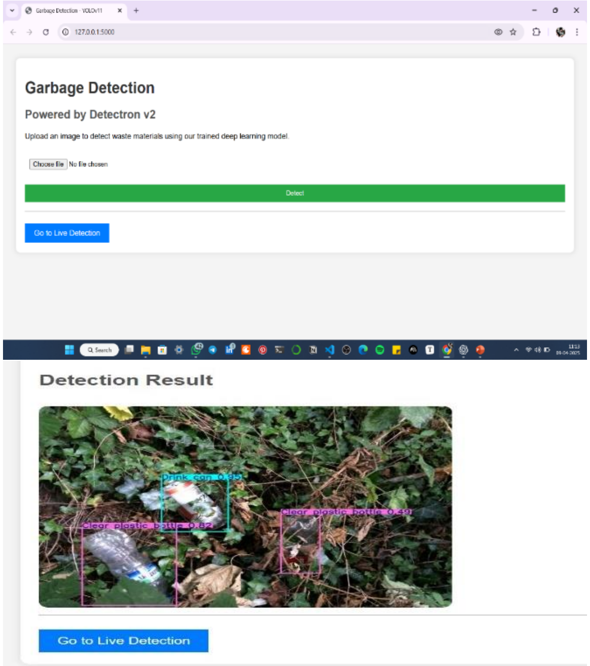



**Fig: Upload an image to detect waste materials with Detection results.**

**Real-Time Detection with Webcam**

Real-time detection was implemented using OpenCV. The webcam feed was captured frame-by-frame, and each frame was passed through the Detectron2 model. Due to the complexity of real-time inference, efforts were made to optimize frame processing to maintain an acceptable frame rate (FPS). This included resizing frames, reducing resolution slightly, and skipping redundant frames.

The processed frames were rendered with bounding boxes and displayed using OpenCV's imshow() or streamed directly in the browser via Flask's Response generator. This provided users with a smooth live detection experience.

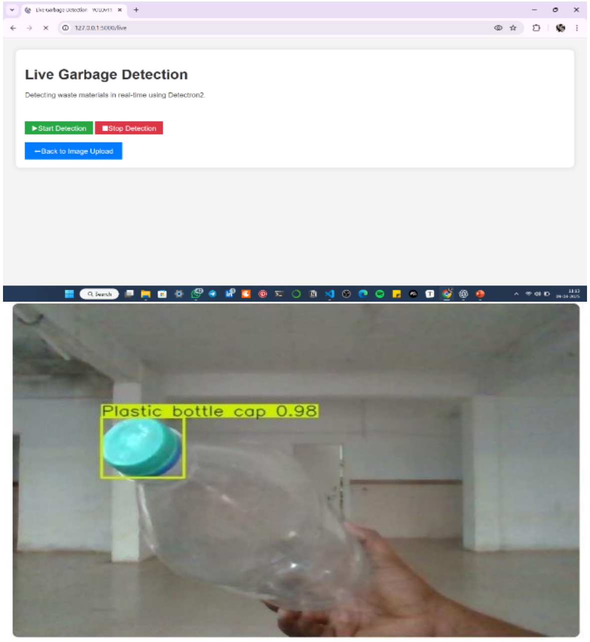



**Fig: Detecting waste materials in real time using detectron2 with live detection results.**

## 5. CONCLUSION

In this project, a deep learning-based garbage detection system was successfully developed using the Faster R-CNN architecture within the Detectron2 framework.

The system was trained on the TACO dataset, which provided realistic, annotated images of waste in natural settings. Through careful data preprocessing, augmentation, and training procedures, the model achieved strong detection performance, with an mAP of 78.4% and high precision and recall values. The use of advanced computer vision techniques enabled the model to accurately identify and localize various categories of waste objects, even under challenging conditions such as cluttered backgrounds and variable lighting.

The integration of the model into a Flask-based web application allowed for an interactive, user-friendly interface supporting both image uploads and live webcam detection. This ensured that the system was not only technically robust but also practically accessible. The deployment design made it adaptable for various real-world applications, such as smart city waste monitoring, environmental awareness campaigns, and automated sorting systems in public or industrial zones.

Overall, the project demonstrated the potential of AI and computer vision in addressing real-world environmental challenges. It combined technical innovation with practical deployment to create a system that is both efficient and impactful in promoting sustainable waste management.

## SCOPE FOR FUTURE WORK

While the current implementation offers a strong foundation, there are several opportunities for further improvement and expansion:

- **Enhanced Dataset and Fine-Grained Classification**: Future work can involve the integration of additional datasets to improve the model's accuracy and support finer categorization of waste (e.g., biodegradable vs. non-biodegradable, recyclables vs. hazardous waste).

- **Model Optimization for Edge Devices**: Deploying a lightweight version of the model on embedded devices such as Raspberry Pi, Jetson Nano, or edge-AI hardware would enable real-time, low-power detection in remote or outdoor environments

- **Integration with IoT and Smart Bins**: The system can be connected with IOT-enabled smart bins that use sensors to trigger image capture and automatic detection, followed by intelligent sorting or bin capacity monitoring.

**ACKNOWLEDGEMENTS**

The authors would like to thank the Department of Artificial Intelligence And Machine Learning,Lingyas Institute of Management and Technology, for providing the research infrastructure and support required for this work.
We are grateful to Siva Rama Krishna, for their valuable guidance, feedback, and mentorship throughout the research.
We also acknowledge the reviewers and editorial team of the International Journal of Information Technology (IJIT) for their insightful suggestions that contributed to the improvement of this manuscript.

**REFERENCES**

[[1] W. Liu, D. Anguelov, D. Erhan, et al., "SSD: Single Shot MultiBox Detector," *European Conference on Computer Vision (ECCV)*, pp. 21–37, 2016.

[2] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," arXiv preprint arXiv:1804.02767, 2018.

[3] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017.

[4] Facebook AI Research, "Detectron2," GitHub Repository, https://github.com/facebookresearch/detectron2 (accessed March 2025).

[5] P. Proença and P. Simões, "TACO: Trash Annotations in Context Dataset for Litter Detection," *arXiv preprint* arXiv:2003.06975, 2020.

[6] A. Krizhevsky, I. Sutskever, and G.E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Communications of the ACM*, vol60, no. 6, pp. 84–90, 2017.

[7] J. Deng et al., "ImageNet: A Large-Scale Hierarchical Image Database," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 248–255, 2009.

[8] A. Paszke, S. Gross, F. Massa, et al., "PyTorch: An Imperative Style, High-Performance Deep Learning Library," *Advances in Neural Information Processing Systems*, vol. 32, 2019.

[9] M. Abadi et al., "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems," *arXiv preprint* arXiv:1603.04467, 2016.

[10] M. Satyanarayanan, "The Emergence of Edge Computing," *Computer*, vol. 50, no. 1, pp. 30–39, 2017.