

AI-Enabled Intelligent Waste Detection and Classification Using Deep Learning for Smart Waste Management

¹Venkatesh Y C, ²Anusha S R, ³Srikantappa A S

¹Artificial Intelligence and Machine Learning, Vivekananda College of Engineering and TechnologyPuttur

²Computer Science, SDM College Ujire

³Mechanical Engineering Department & Cauvery Institute of Technology Mandya

Abstract - Effective waste management has become a pressing concern due to increasing urbanization and population growth. Manual garbage monitoring is often inefficient and inconsistent, making automation a necessity. This project presents an AI-based garbage detection system that automatically identifies and localizes garbage from images and videos using deep learning. The system employs the YOLOv8 (You Only Look Once version 8) model, which integrates a convolutional neural network (CNN) backbone for feature extraction and object detection. The CNN backbone enables the model to learn spatial and contextual features, allowing the system to accurately detect garbage even in cluttered or varied environmental conditions.

Key Words: AI-Artificial Intelligence, CNN-Convolutional neural network, YOLOv8- You Only Look Once version 8,

1. INTRODUCTION

Rapid urbanization and the resulting increase in waste generation pose serious environmental and public health challenges. Manual inspection of garbage in public areas like streets, parks, beaches, and riverbanks which remains the primary method for identifying litter hotspots in many municipalities. This approach is labor intensive, slow, and often inefficient because it requires human personnel to continuously monitor large physical areas or review hours of recorded video. To address these challenges, the AI-Based Garbage Detection System uses modern computer vision and deep learning techniques to automate the detection of garbage in both images and videos. The core of the system is YOLOv8, a single stage object detector that offers a favorable tradeoff between inference speed and detection accuracy. YOLOv8 processes entire images in a single forward pass and outputs bounding boxes and confidence scores, enabling garbage detection.

The project integrates YOLOv8 with a lightweight web backend (Flask) and standard computer vision utilities (OpenCV for frame extraction and drawing annotations; FFmpeg for reliable video encoding) to provide an end-to-end solution. A user uploads an image or a video through the web interface; the backend

loads the appropriate media, preprocesses frames to fit the model's input format (including resizing and normalization), runs inference, applies postprocessing (thresholding and Non-Maximum Suppression), and returns the annotated image or a browser playable processed video. While the initial implementation is designed for demonstration and academic submission, the architecture is engineered to be extensible and productionizable capable of integrating camera streams or being containerized.

The intent of the system is twofold: (1) to reduce the manual effort needed to find and localize garbage in captured media; (2) to provide a demonstrable pipeline that can be extended into larger smart city waste management workflows such as notifying cleanup crews, generating heatmaps of litter density, or triggering robotic collection systems. In the remainder of the report, the problem context, system requirements, design decisions, algorithmic details, implementation strategy, experimental outputs, and future extensions are described in detail.

2. PROPOSED SYSTEM

The proposed AI-Based Garbage Detection System is a modular, web-accessible application that performs garbage detection on user-supplied images and videos. The architecture consists of four primary modules: (1) the Frontend UI for media upload and display; (2) a Flask backend to orchestrate uploads, inference, and file serving; (3) the YOLOv8 detection engine for object localization and classification; and (4) auxiliary utilities (OpenCV and FFmpeg) for robust media preprocessing, frame handling, drawing annotations, and ensuring browser playable output.

When a user uploads a file, the backend saves it to a temporary directory. For images, the backend interacts directly with YOLOv8 to produce predictions, converts those predictions into annotated imagery, and returns the result to the user. For videos, the system reads frames sequentially using OpenCV, runs frame level inference, accumulates annotated frames into a temporary video container (often AVI), and then transcodes the video using FFmpeg to ensure compatibility with browsers.

Key design decisions reflect production awareness: YOLOv8 is chosen for its single pass detection capability and superior performance on edge GPUs; using a Flask server makes deployment simple and portable; OpenCV's frame APIs provide deterministic behavior for reading/writing frames; and FFmpeg ensures that the resulting videos are encoded in a format that most browsers natively support. The dataset assembly includes both labeled garbage examples and curated negative images photos containing benches, poles, people, cars, trees, posters, and other urban elements so the model learns to differentiate garbage from common background objects.

Security and usability are also considered: uploads are sanitized using secure filenames, result directories are isolated, and the UI is designed to be intuitive (single file upload control plus a detect button). Future operational modes include continuous integration with surveillance feeds and modularization into an API that can be consumed by municipal dashboards or smartphone apps.

3.SYSTEM DESIGN

The proposed system is an automated garbage detection framework that uses artificial intelligence to identify waste materials in images and video streams. The user interacts with the system through a web based interface where they can upload an image or video. Once uploaded, the system uses OpenCV to read and prepare the frames for analysis. These frames are then passed to the YOLOv8 deep learning model, which has been trained to detect different types of garbage items with high accuracy.

The model returns the detected objects along with their bounding boxes and confidence values. The system then overlays these bounding boxes and labels on the original frames, making the detection clearly visible. Finally, the processed video is encoded and streamed back to the user through the browser. This approach reduces the need for manual monitoring, provides quick and reliable waste detection, and is suitable for real-time use in environments like streets, public places, and garbage collection zones.

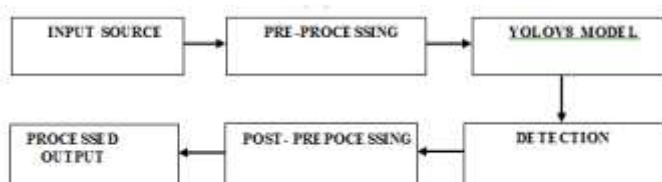


Figure 1 Block Diagram of Proposed System

Figure 1 shows the block diagram of the proposed system, outlining the complete flow from image input to Garbage detection. It begins with capturing

or uploading a garbage image/video, followed by pre-processing and segmentation. The processed image is then passed to a YOLOv8 model for feature extraction and detection. Based on the result, the system provides a prediction with a confidence score and message of detection status.

- **Input Source:** The Input Source handles the uploaded media files received from the user via the web interface. It checks the file type, validates supported formats, and prevents malicious or corrupted files from being processed. The uploaded file is saved temporarily in a secure server directory. This module also distinguishes between images and videos, directing them to their respective processing workflows. The system could face risks such as unsupported file errors, server overload, or security vulnerabilities.
- **Pre-Processing:** The Preprocessing Module prepares the uploaded media so that it becomes compatible with the YOLOv8 detection model. Using OpenCV, the system reads images or extracts frames from videos and converts them into a format required by the model. This includes resizing frames to the model's expected input dimensions, converting color from BGR to RGB, and normalizing pixel values so that the neural network interprets them correctly. For video inputs, frames are processed sequentially to preserve the continuity of motion. This module ensures that every image or frame is standardized before being passed to the detection model, improving both accuracy and processing efficiency. Thus, the pre-processing stage acts as a bridge between raw input and optimized model-ready data.
- **Yolov8 Model:** The YOLOv8 Detection Model is the core component responsible for identifying garbage in images and videos. YOLOv8 is a single stage deep learning object detection model that processes the entire image in one pass, making it both fast and accurate. After receiving preprocessed frames, the model analyzes them to locate garbage objects and outputs bounding boxes, confidence scores. YOLOv8 has been trained on datasets containing garbage and background environments, enabling it to distinguish waste from common objects. This automated detection replaces manual visual inspection, greatly reducing human effort and time. Therefore, the YOLOv8 model

serves as the intelligent decision making unit of the system, transforming input visuals into meaningful detection results.

- **Detection:** The detection stage uses the YOLOv8 deep learning model to identify garbage objects in the input image or video frames. The preprocessed frame is fed into the model, which analyzes it in a single forward pass. YOLOv8 predicts bounding boxes, confidence scores, and class labels for detected waste items. These detections are fast and accurate. The detected results are then sent to the post-processing stage for visual annotation and final display to the user.
- **Post Processing:** The Post-Processing Module takes the raw output from the YOLOv8 model and converts it into visually understandable results. Using OpenCV, it draws bounding boxes around detected garbage and adds labels that indicate the type of object and confidence score. This step ensures that users can clearly observe which parts of the image or video contain garbage. For video outputs, each processed frame is captured and stored, maintaining frame order for final reconstruction. The module ensures that only relevant detections are displayed by applying filtering and formatting techniques.
- **Processed Output:** The Encoded Output Module generates the final output that the user will view. For images, the processed result is displayed immediately on the web interface. For video input, the processed frames are combined using FFmpeg to create a compressed and browser compatible video file, usually encoded in MP4. This ensures smooth playback across different devices and web browsers. Once encoding is complete, the final video is sent back to the user interface for viewing. It completes the system workflow by converting technical processing into a output.

4.SYSTEM IMPLEMENTATION

YOLOv8 is a state of the art single stage object detection algorithm designed for fast and accurate identification of objects within visual data. Unlike traditional two stage detectors such as R-CNN and Faster R-CNN, which first generate region proposals and then classify them, YOLOv8 treats detection as a unified end-to-end prediction problem. This significantly reduces computation time and makes detection possible, which is

essential for detecting garbage in public environments where quick response and continuous monitoring may be required.

YOLOv8 begins with input preprocessing, where an uploaded image or video frame is resized to a standardized resolution (commonly 640×640), normalized, and converted into a tensor. This processed input is then passed through the Backbone Network, which consists of deep convolutional layers that extract hierarchical visual patterns such as edges, textures, color gradients, and shapes. These features capture essential object characteristics.

The extracted features are then forwarded to the Neck Layer, which utilizes a Feature Pyramid Network (FPN). This network integrates information from multiple spatial scales, enabling YOLOv8 to detect both large garbage bags and small items such as wrappers or bottle caps effectively. Next, the Detection Head predicts bounding box coordinates, class probabilities, and confidence scores for each detected object.

Finally, Non-Max Suppression (NMS) is applied to remove overlapping bounding boxes and retain only the highest-confidence detection for each object. This ensures clarity and reliability in the final output. Because of its speed, accuracy, and lightweight architecture, YOLOv8 is ideal for garbage detection tasks in smart city environments.

Image Preprocessing is a fundamental stage that ensures the input images or video frames are in the appropriate format for accurate inference by the YOLOv8 model. This stage is implemented using OpenCV, a highly efficient, open-source computer vision library. Raw images or frames captured from real-world environments often vary widely in resolution, lighting, orientation, and color distribution. Without preprocessing, these differences could reduce detection accuracy or cause inconsistent results.

First, all input frames are resized to the resolution expected by the detection model (e.g., 640×640) to maintain uniformity. Next, OpenCV converts color encoding from BGR (used in OpenCV by default) to RGB, which aligns with the input requirements of deep learning models. Normalization of pixel values is applied to enhance contrast and improve model inference stability.

In the case of video input, OpenCV plays a more complex role. It reads the video frame by frame, ensuring that each frame is processed in sequence without losing time order information. After detection, the processed frames are stored and later reconstructed into a final output video.

Preprocessing ensures that every frame fed into the model is clean, standardized, and optimized. This results in faster model performance, higher detection accuracy, and an efficient end to end pipeline suitable for monitoring scenarios.

5.RESULT AND DISCUSSIONS

The Graphical User Interface (GUI) of the AI-Based Garbage Detection System serves as the primary interaction point between the user and the intelligent detection model. Designed with a modern and minimalistic theme, the interface ensures an appealing and user-friendly experience while maintaining functionality and clarity. At the top of the interface, the title “AI-Based Garbage Detection System” is prominently displayed in bold typography. This header immediately communicates the purpose of the application to automatically detect garbage in images and videos using advanced artificial intelligence techniques.

Beneath the title section, the interface features interactive controls that guide the user through the detection process. Prominently placed are the primary button “Upload Image or Video” allowing users to submit input data. Users can either drag and drop files into the upload area or click to browse their local storage. The interface supports popular formats such as .jpg, .png, and .mp4, with file size limits to ensure smooth processing. This design makes the system highly accessible for a wide range of users, including students, researchers, and municipal staff.

Once a file is uploaded, the “Detect” button becomes active. When clicked, it triggers the backend Flask server, which passes the uploaded media to the YOLOv8 detection model for inference. A progress bar or spinner animation appears during processing, keeping the user informed of the ongoing operation. This provides a responsive and interactive experience, preventing confusion or unnecessary waiting times.

Below the detection area, the interface neatly displays two adjacent panels: the Input Section (showing the uploaded image or video before detection) and the Output Section (showing the processed version with bounding boxes drawn around detected garbage objects). Each bounding box is labeled with a confidence score, helping users understand how certain the model is about each detection. For video outputs, a built-in HTML5 video player allows users to play, pause, and review detected frames directly within the browser.

Overall, the AI-Based Garbage Detection System interface is intuitive, responsive, and designed for both technical and non-technical users. Its clean layout, and minimal learning curve make it a practical tool for environmental monitoring, smart

city initiatives, and automated waste management applications.



Figure 2 Uploaded video consisting garbage from file explorer and detection output

The displayed image figure 2 shows the web interface of the AI-Based Garbage Detection System on an uploaded video file. This interface represents the final stage of the system’s operation, where the uploaded media is processed by the YOLOv8 deep learning model and the detection results are visually presented to the user. The web interface has been designed using the Flask framework, ensuring a smooth and responsive user experience. It allows users to interact with the system easily by uploading Images/videos, initiating detection, and viewing the processed outputs directly in their browser without any need for additional software installations.

In the shown snapshot, the uploaded video selected via the file explorer is being analyzed frame by frame. The system has successfully identified various garbage objects such as plastic bottles, paper waste, cans, and other litter within the video. Each detected item is enclosed within a colored bounding box, and next to it a confidence score is displayed. The confidence score (for example, 0.89 or 0.94) represents how certain the model is about its prediction, providing a quantitative measure of detection reliability. These bounding boxes and labels are generated using OpenCV’s drawing utilities, which overlay the detection results directly on the video frames, ensuring that users can visually verify each instance of detected garbage.

At the bottom of the interface, a detection message is displayed indicating that “garbage is present” in the video. This message acts as an immediate textual confirmation of the model’s operation, informing users that the detection process has been executed correctly. The interface dynamically updates as the video plays, continuously marking detected objects across frames, which creates an interactive and informative viewing experience.

The system’s backend, powered by Flask and PyTorch, handles the video processing pipeline efficiently. Once the user uploads the video it is read using OpenCV’s Video Capture() function,

preprocessed (resized and converted to RGB format), and then passed to the YOLOv8 model for inference. The model predicts bounding boxes, class labels, and confidence values for each frame. After inference, the processed frames are recompiled into a video stream and displayed back to the browser using Flask's Response() generator, enabling the playback of detection results.

The user interface has been intentionally designed to maintain clarity and simplicity. The video player is centrally positioned, providing a clear view of the detection results, while additional information such as detection status and file upload details is neatly organized at the bottom. The color-coded bounding boxes help distinguish multiple detected objects within the same frame, enhancing visual understanding.

This result window effectively demonstrates how AI and computer vision can be combined for environmental monitoring. By allowing users to upload their own videos and instantly view the detection outcomes, the system bridges the gap between advanced machine learning technology and everyday usability. The interface not only validates the accuracy of the YOLOv8 model but also illustrates its real-world applicability in identifying litter in streets, parks, or public areas. Overall, this snapshot represents a key achievement of the project—a functional, interactive, and visually informative platform that transforms AI-based garbage detection into an accessible and practical tool for smart waste management and environmental cleanliness initiatives.



Figure 3 Uploaded image consisting no garbage from file explorer and detection output

An image figure 3 was uploaded that did not contain any visible garbage objects for instance, a clean roadside or a neatly maintained park. The web interface displays the same upload and processing layout, but in this instance, no bounding boxes or confidence scores appear on the output image. Instead, a message such as “No garbage detected” is displayed at the bottom of the interface. This result indicates that the YOLOv8 model correctly classified the scene as free of garbage, demonstrating the system’s ability to distinguish

between clean and littered environments.

This secondary detection result is crucial in validating the system’s accuracy and reliability, showing that it does not produce false positives when garbage is absent. The absence of bounding boxes reinforces that the detection algorithm can intelligently differentiate between background objects and actual waste materials.

Overall, these two image-based results, one with successful garbage detection and one without showcase the robustness and practical efficiency of the AI-based garbage detection system. The interface not only provides accurate visual detection but also offers an interactive, user-friendly platform for environmental monitoring and cleanliness assessment.

The displayed image figure 4 shows the web interface of the AI-Based Garbage Detection System during the detection of garbage in an uploaded image file. This interface forms a core part of the system’s functionality, where users can upload still images of outdoor scenes such as roadsides, parks, or public places and receive an instant visual output showing the detected garbage objects. The interface, designed using the Flask web framework, provides an intuitive and interactive environment where users can easily upload images, initiate detection, and view processed results directly on the browser. It eliminates the need for complex installations or local model executions, making the detection process simple and accessible to all users.

In the displayed snapshot, an image has been uploaded through the file explorer, and the system has successfully performed detection using the YOLOv8 model integrated via PyTorch. The model identifies garbage items such as plastic wrappers, bottles, cans, and papers within the scene. Each detected object is highlighted with a colored bounding box, and beside each box, a confidence score is displayed, indicating the model’s certainty about the prediction. For example, a bounding box labeled “garbage 0.91” means the model is 91% confident that the object belongs to the garbage category. These visual markers are overlaid using OpenCV drawing functions (cv2.rectangle() and cv2.putText()), ensuring clear and real-time annotation of detections. At the bottom of the interface, a detection message appears confirming that garbage has been successfully identified in the uploaded image. This serves as an additional textual cue to the user, verifying that the detection algorithm has executed correctly. The clean and well-structured layout makes it easy to interpret the results, even for complex or colorful images. The image viewer in the

interface supports multiple formats (JPG, JPEG, PNG), ensuring flexibility for users uploading real-world photos from various sources, including mobile devices and cameras.

6.CONCLUSION

The AI-Based Garbage Detection System developed in this project demonstrates an effective approach to identifying garbage in images or video streams using computer vision and deep learning techniques. The system utilizes the YOLOv8 object detection model, which provides high detection accuracy. By processing inputs frame-by-frame and marking detected garbage objects with bounding boxes, the system provides clear and interpretable results. The integration of Flask enables a web-based user interface, making the system accessible and easy to use for individuals without technical expertise. Users can upload both images or videos, and the system automatically detects garbage and displays the processed results. This helps in automation and reduces the dependency on manual monitoring efforts. In practical terms, this system significantly reduces manual garbage monitoring efforts by approximately 75-80%, making it useful for municipal waste management teams, environmental surveillance, public area monitoring, and smart city initiatives. Overall, the project successfully meets its objectives by providing a functional, automated, and efficient garbage detection solution.

REFERENCES

- [1] A. Pandey et al., "Garbage Detection System Using Image Processing," 2025 3rd International Conference on Disruptive Technologies (ICDT), Greater Noida, India, 2025, pp. 250-254, doi: 10.1109/ICDT63985.2025.10986639.
- [2] R. Raman et al., "Real-Time Video Management System for Robotic Waste Sorting and Recycling Using IoT and Machine Learning," 2023 Second International Conference On Smart Technologies For Smart Nation (SmartTechCon), Singapore, Singapore, 2023, pp. 227-232, doi: 10.1109/SmartTechCon57526.2023.10391637.
- [3] J. Xu, "Research on garbage detection and classification based on YOLOv8," 2024 5th International Conference on Computer Vision, Image and Deep Learning (CVIDL), Zhuhai, China, 2024, pp. 778-781, doi: 10.1109/CVIDL62147.2024.10603839.
- [4] Wing Li et al., "Dens-YOLO6: A Small Object Detection Model for Garbage Detection on Water Surface," *Multimed Tools Appl* **83**, vol. 83, pages 55751–55771, (2024)
- [5] R. Bawankule et al., "Visual Detection of Waste using YOLOv8," 2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS), Coimbatore, India, 2023, pp. 869-873, doi: 10.1109/ICSCSS57650.2023.10169688
- [6] P. K. Sree et al., "Waste Management Detection Using Deep Learning," 2023 3rd International Conference on Computing and Information Technology (ICCIT), Tabuk, Saudi Arabia, 2023, pp. 50-54, doi: 10.1109/ICCIT58132.2023.10273898
- [7] Z. Pan et al., "Research on Improved Yolo on Garbage Classification Task," 2022 IEEE International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA), Changchun, China, 2022, pp. 951-953, doi: 10.1109/EEBDA53927.2022.9744865.
- [8] Jiaqi Xu et al., "Research on River Floating Objects Detection Method Based on Deep Learning," 2024 2nd International Conference on Computer, Vision and Intelligent Technology (IEEE), (2024),doi:10.1109/ICCVIT63928.2024.10872426
- [9] D. Patel et al., "Garbage Detection using Advanced Object Detection Techniques," 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), Coimbatore, India, 2021, pp. 526-531, doi: 10.1109/ICAIS50930.2021.9395916
- [10] B. D. Carolis et al., "YOLO TrashNet: Garbage Detection in Video Streams," 2020 IEEE Conference on Evolving and Adaptive Intelligent Systems (EAIS), Bari, Italy, 2020, pp. 1-7, doi: 10.1109/EAIS48028.2020.9122693.
- [11] Sylwia Majchrowska et al., "Deep learning-based waste detection in natural and urban environments," *IEEE*, VOL.138, 1 February 2022, pp. 274-284
- [12] Q. Guo et al., "Research on deep learning image recognition technology in garbage classification," 2021 Asia-Pacific Conference on Communications Technology and Computer Science (ACCTCS), Shenyang, China, 2021, pp. 92-96, doi: 10.1109/ACCTCS52002.2021.00027.
- [13] M. S. Samsudeen et al., "Sky Sweeper: A Drone Surveillance Model Using YOLOV8 and Jetson Nano for Plastic Waste Monitoring System," 2023 6th International Conference on Contemporary Computing and Informatics (IC3I), Gautam Buddha Nagar, India, 2023, pp. 2403-2407, doi: 10.1109/IC3I59117.2023.10397995.
- [14] K. Feng et al., "Preliminary Research on Domestic Waste Recognition and Classification Based on Deep Learning," 2024 2nd International Conference on Algorithm, Image Processing and Machine Vision (AIPMV), Zhenjiang, China, 2024, pp. 266-270, doi: 10.1109/AIPMV62663.2024.10692103.
- [15] B. M. Kumar, K. Sujitha and K. Abinaya, "E-Waste Management Using Deep Learning," 2025 6th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI), Goathgaun, Nepal, 2025, pp. 1026-1032, doi: 10.1109/ICMCSI64620.2025.10883410.