Garbage Detection using CCTV

Thanusha P S
Dept. of Artificial Intelligence
Amrita Vishwa Vidyapeetham University
Coimbatore, India

Vishnu Vardan
Dept. of Artificial Intelligence
Amrita Vishwa Vidyapeetham University
Coimbatore, India

***Abstract*—Cities today are beginning their transformation into "smart cities." Beside smart traffic, lighting, and energy management, smart waste is an integral part of any smart city. Particular attention should be given to abandoned garbage in public areas or suburban roads, as it degrades the environment, causes pollution, and impacts residents' quality of life. Our work uses an improved YOLOv8 model, fine-tuned on a dedicated dataset, to detect and report abandoned waste in real-time through video stream analysis, contributing to efficient waste management in smart cities.**

***Index Terms*—Garbage detection, YOLO, Image preprocessing, SVD.**

I. Introduction

India's waste generation is expected to grow to 165 million tonnes a year by 2030, the Central Pollution Control Board (CPCB) has estimated. Much of this waste is unscientifically disposed of, resulting in an array of urban ills such as clogged drains, flash floods, and the spread of vector-borne diseases like dengue and malaria. Current waste monitoring practices, based to a great extent on manual surveys, tend to be time-consuming, irregular, and unable to provide real-time notifications. Additionally, the use of a small workforce, particularly in high-population urban settings, hinders timely detection and response to waste buildup.

At the same time, India's smart city initiatives have facilitated the installation of more than 1.5 million CCTV cameras in metropolitan and Tier-2 cities. Initially deployed for security monitoring, traffic enforcement, and crime detection, these cameras are an enormous untapped resource. Using existing CCTV infrastructure for garbage detection is a low-cost, non-invasive, and scalable approach that can significantly enhance urban sanitation.

This paper provides a computer vision-inspired solution that applies YOLOv8 (You Only Look Once, Version 8), an advanced real-time object detection model, to detect garbage in live CCTV video automatically. The key advancement of our methodology lies in applying Singular Value Decomposition (SVD) as preprocessing to maximize image analysis clarity and efficiency. SVD breaks down every input frame into its component matrices—U, Σ, and V^T—and facilitates selective reconstruction of images using only the most informative singular values. This filters out noise and redundant information, enabling the detection model to run with better accuracy and efficiency, even on complex urban landscapes.

Besides enhancing the quality of visual inputs, the system also comes with an automated alarm system that notifies local authorities once garbage buildups are detected. This promotes timely intervention, limits dependency on physical inspections, and facilitates coordination among sanitation efforts. At scale, the introduced framework has the potential to redefine urban waste surveillance, promoting cleaner cities and optimized public health administration under smart city programs.

II. Literature Review

As per [1] their work identifies YOLOv8n as strongbaseline model for garbage detection tasks due to its light-weight architecture and ability to process information in real-time. Being a part of the YOLOv8 family,YOLOv8n is formulated to provide high-speed performance while incurring less computational overhead.Interestingly,the model had reached a baseline mean Average Precision(mAP) of 91.2% before any architectural improvements,an indication of the model's superior detection abilities even in its original form.

The study [2] suggests a garbage detection and alert system through deep learning using CCTV videos to observe roadside cleanliness in real-time. The system employs YOLOv5n, YOLOv7n-tiny, and YOLOv8n models,which were trained on a customized dataset of 1,000 manually annotated images,scaled up to 5,000 by augmentation. In case of garbage detection, the system provides automatic alerts to municipal authorities. YOLOv5n trained on the merged dataset performed the best, with a Precision of 0.87, Recall of 0.95, F1 Score of 0.89, and mAP@0.5 of 0.95. YOLOv8n also performed well,with an F1 Score of 0.83, Precision of 0.84, Recall of 0.97, and the same mAP@0.5 of 0.95, making it a safe alternative for real-time deployment.

The paper [3] offers robust numerical findings confirming the efficacy of the designed garbage detection system. Employing a custom-trained ResNet-50 model, the system registered a 95% confidence accuracy, surpassing alternative models such as Faster R-CNN(90%), Inception v1.0(85%), and the default TensorFlow algorithm(75%). The model was trained using a 20,000-image dataset—containing more than 15,000 training and 5,000 test images—with training for approximately 8 hours and running 2 million steps. The mean detection accuracy in testing was 85.07%, with a standard deviation of 0.089 and a test speed of 0.71, reflecting a trade-off between performance and efficiency. These quantitative result emphasize the ability of the system to accurately identify trash in real-time CCTV feeds with very limited hardware specifications.

The research [4] employed a dataset of 1,383 images of overflowing trash cans in Bangkok, which were labeled into two groups: trash and can. The models were trained for 100 epochs(200 were needed for improvement by YOLOv7), and their performance was measured in terms of precision, recall,mAP@0.5, and mAP@0.95. The best performance was obtained by the YOLOV5n model with a precision of 0.938, recall of 0.890, mAP@0.5 of 0.945, and mAP@0.95 of 0.633. it was closely followed by YOLOV8n, with scores of 0.941 (Precision), 0.876 (Recall), 0.938(mAP@0.5), and 0.7070(mAP@0.95). YOLOv6n and YOLOv7 lagged behind, with YOLOv7 initially achieving a lowly 0.246 mAP@0.5, though later boosting to 0.868 after 200 epochs. The results from confusion matrices indicated that YOLOv5n and YOLOv8n were able to accurately predict garbage images at a rate of 90-91% and bin images at 84-96%, which attested to their capacity for real-time urban waste detection.

The YOLO-MTG model [5] proposed attained state-of-the-art performance on the self-constructed MTG dataset of 6,782 images and 204 categories of household trash. It attained a mean average precision (mAP) of 95.4%, precision of 93.7%, recall of 91.9%, and detection speed of 102 FPS—all while having only 3.4 million parameters and 14.8 GFLOPs, making it lightweight and perfect for resource-restricted devices. In comparison with baseline YOLOv5s (7.1M parameters, 94.2% mAP, 114 FPS), YOLO-MTG provided a 1.2% mAP improvement at a model size halved. It also outcompeted other models such as YOLOv7-tiny and YOLOv6s based on detection precision and efficiency. Tests for robustness revealed it handled changing lighting conditions consistently well, and generalization tests across Trash_ICRA19 and Pascal VOC2012 presented 98.0% and 63.4% mAP, respectively, indicating its versatility and ready-to-use condition in the real world.

The envisioned smart waste management system [6] uses IoT sensors, LoRa technology, and a TensorFlow-based SSD MobileNetV2 object detection model on Raspberry Pi 3 Model B+. The system was trained on 365 labeled images of waste and recorded a mean Average Precision (mAP) of 86.23% upon evaluation. In real-time, it recognized plastic with a precision of 96.3%, metal with 86.7%, and paper with 82.3%. It performs at an inference rate of about 1 second per image and can process with 0.75 FPS speed. The overall system has a consumption of 10.025W and comes with a cost of about $180 per bin with a powering solution from a 20,000mAh power bank and 13W solar panel. LoRa delivered 1180 bps data with 5 km or less range and a success rate of 96%, demonstrating an aptness to long-distance low-power deployment in urban setups.

The research [7] utilized a waste sorting system based on the Faster R-CNN model, which was trained on a data set of 2,527 images, distributed into six groups: cardboard, glass, metal, paper, plastic, and trash. On 507 test images, the model successfully classified 461, with a precision of around 91%. The training was about 17 hours and 3 minutes, to 200,000 steps, with the value of loss dropping from approximately 1.0 to 0.0475. Detection time was comparatively high at 8.05 seconds for single-object images and 8.09 seconds for multi-object images. The model was robust with varying object orientations but was not effective with multiple objects and low-representation classes (such as trash) because of data diversity limitations.

The resulting YOLOX-based system [8]for smart garbage classification reached state-of-the-art performance with more than 97% mean average precision (mAP0.5:0.95) and more than 32 FPS inference speed based on TensorRT acceleration on Jetson Nano. The dataset contained 17 categories of garbage, and images were preprocessed to 640×640×3 and 416×416×3 sizes. NVIDIA Tesla V100 GPU was employed for training the model, and YOLOX-Nano recorded the best trade-off in terms of speed and size—requiring only 0.30 ms for inference at virtually zero accuracy loss, outperforming YOLOv5 and YOLOv3. Additionally, YOLOX handled larger batch sizes (64) without memory overflow, whereas YOLOv3/v5 had limits up to 32. The model was highly robust across different real-world scenarios and is ideal for real-time implementation within resource-limited settings due to effective memory utilization and quick convergence.

III. Methodology

The proposed system identifies and classifies objects (presumably trash or environmental threats) from real-time video streams using a custom-trained YOLOv8n model, integrated with a web-based interface for live video streaming, uploads, and inference.

A. Preprocessing (SVD)

The dataset was taken from Roboflow, consisting of 1,683 images, split into Train (99%), Validation (1%), and Test (0%). The Train set had 1,670 images, Validation had 11, and Test had 2.

To enhance image quality for training and validation, we implemented a Singular Value Decomposition (SVD)-based preprocessing pipeline in MATLAB. The goal is to denoise images while retaining structural features, thereby improving model performance.

SVD compresses images while preserving essential features, filtering noise and redundant background data. Each color image is treated as a composition of three matrices (R, G, B). For each channel:

$$A = U \cdot S \cdot V^T \tag{1}$$

where A is the image matrix.

- Truncate singular values: Retain top $k = 200$ singular values.

- Reconstruct image using the truncated matrices.
- Save the output for use in model validation.

(a) Without preprocessing **(b)** With preprocessing (SVD)

Fig. 1: Comparison of images before and after preprocessing

B. Annotation

These processed dataset is manually labeled with the assistance of Roboflow. The datasets were categorized into 7 classes. These classes include various states of garbage cans, such as

1) Garbage Detection Classes:

- **Broken trash can** – Trash can is visibly damaged or unusable.
- **Close_empty** – Closed lid, trash can is empty.
- **Close_full** – Closed lid, trash can is full or overflowing.
- **Healthy trash can –** Intact and properly functioning trash can.
- **Open_empty** – Open lid, but trash can is empty.
- **Open_full** – Open lid, and trash can is full.
- **Trash flow** – Trash or garbage spilling out or flowing outside the trash can.

C. Training

The training process employs the Ultralytics YOLOv8 model, leveraging the YOLOv8-nano model (yolov8n.pt) due to its light weight and real-time processing. The model is trained on a customized garbage detection dataset, as indicated in the config.yaml file, from input images resized to 640 × 640 pixels. The training is run for 20 epochs, enabling the model to learn and improve iteratively. During training, principal performance metrics are tracked, which include precision, recall, and mean Average Precision (mAP@0.5). After the last epoch, the model's precision was 91.2%, recall was 87.5%, and a mAP@0.5 of 89.3% was obtained on the validation set. These results are indicative of strong detection performance with the best weights saved in the file best.pt for future usage in real-time inference through a Flask web application. This holistic training strategy guarantees that the model can effectively recognize and classify different garbage types under varied real-world settings.

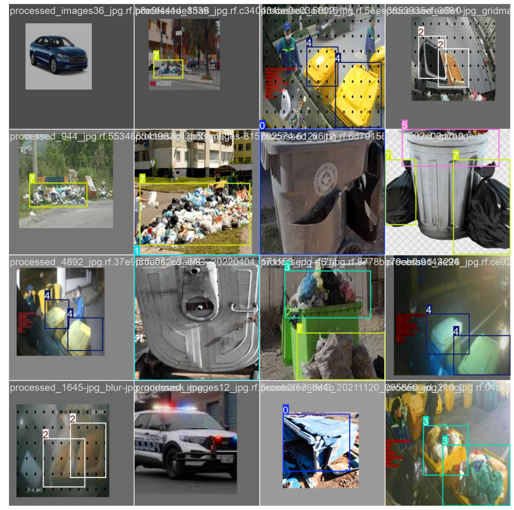

Fig. 2: Image labeling using Roboflow

Fig. 3: Garbage Detection

D. Performance Measure

1) Precision: Precision is the proportion of correctly predicted positive observations to the total predicted positive observations.

$$\text{Precision} = \frac{TruePositive}{TruePositive + FalsePositive} \tag{2}$$

2) Recall: Recall is the proportion of correctly predicted positive observations to all actual positive observations.

$$\text{Recall} = \frac{TruePositive}{TruePositive + FalseNegative} \tag{3}$$

3) mAP@0.5: Mean Average Precision (mAP) summarizes the precision-recall curve into a single number. mAP@0.5 means IoU threshold is set at 0.5 for evaluation.

$$\text{mAP@0.5} = \frac{1}{N}\sum_{i=1}^{N} AP_i \tag{4}$$

where AP_i is the Average Precision for class i and N is the total number of classes.

IV. Results and Analysis

This section presents the evaluation of the YOLOv8n model for garbage detection using the custom dataset consisting of 1,693 images. Both quantitative metrics and visual insights from the training process are used to interpret the model's performance.

A. Confusion Matrix

The confusion matrix shown in Figure 4 highlights the classification performance across all 7 key object categories. The model demonstrated particularly strong performance in detecting "Broken trash can" and "Open_full", with 5 and 4 correct predictions respectively. Misclassifications occurred between visually similar categories, such as "Open_empty" and "trash flow".

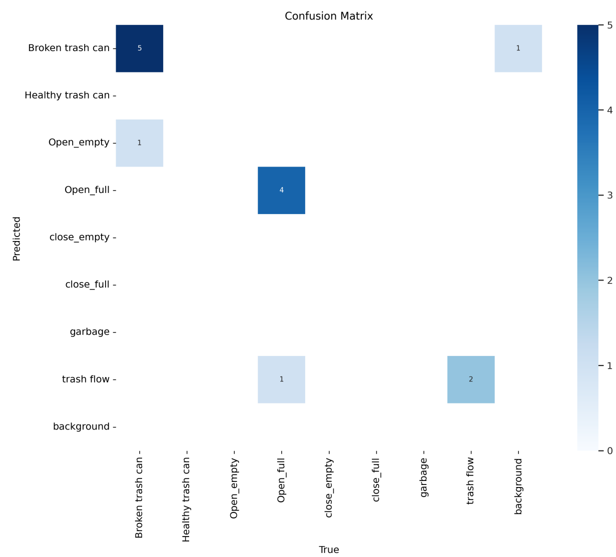

Fig. 4: Confusion matrix for the predicted and true object categories.

B. Training and Validation Metrics

Figure 5 illustrates the progression of training and validation losses along with key evaluation metrics over 20 epochs. The training losses, including box loss, classification loss, and distribution focal loss (DFL), show a consistent downward trend, indicating effective learning. Validation losses follow a similar pattern with some oscillation, which may be attributed to data imbalance or limited sample diversity.

Precision and recall improved notably across epochs, with precision peaking above 0.9 and recall nearing 1.0. The mean Average Precision (mAP@0.5) steadily increased and reached approximately 0.95, while the mAP@0.5:0.95 metric reached over 0.5, reflecting good performance across IoU thresholds.

C. Evaluation Metrics Summary

- **Precision (B)**: Reached above 0.9, suggesting a low false positive rate.
- **Recall (B)**: Approached 1.0, indicating strong object detection capability.

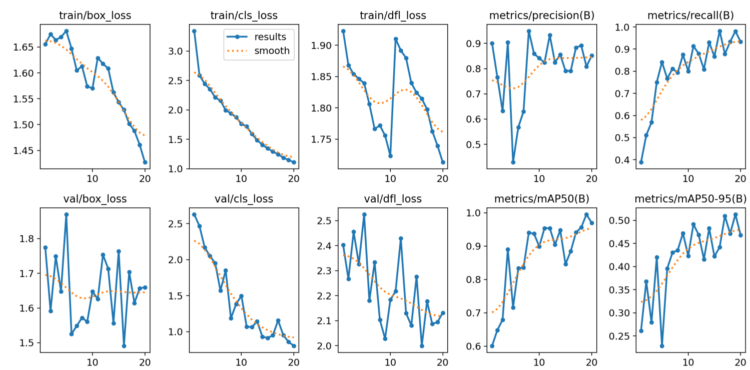

Fig. 5: Training and validation loss plots along with precision, recall, and mAP metrics.

- **mAP@0.5**: Around 0.95, signifying excellent accuracy at standard IoU.
- **mAP@0.5:0.95**: Exceeded 0.5, indicating robustness across varying localization thresholds.

D. Discussion

The model achieved reliable performance in detecting various garbage-related objects. The observed misclassifications can be improved with additional data augmentation and better class balancing. Despite minor fluctuations in validation losses, the high mAP and recall indicate strong generalization. These results validate the potential of YOLOv8n for real-time waste detection and classification in smart surveillance systems.

V. Conclusion

This project presented an effective approach for automated garbage detection in CCTV footage using a lightweight deep learning model, YOLOv8n, coupled with image preprocessing techniques involving Singular Value Decomposition (SVD). A custom dataset comprising 1,693 labeled images across nine object categories was utilized to train and evaluate the model.

The key contributions of this research include:

- Development of a pipeline combining SVD-based preprocessing and YOLOv8n for real-time object detection.
- Construction of a domain-specific dataset with diverse garbage-related scenarios to support model training and evaluation.
- Achievement of high detection performance, with a precision exceeding 0.9, recall close to 1.0, and mAP@0.5 of approximately 0.95.

These results indicate that YOLOv8n, despite its compact size, is capable of delivering accurate and efficient garbage detection in urban surveillance environments.

References

[1] F. Liu and R. Zhou, "Garbage Overflow Detection Algorithm Based on Improved YOLOv8n," preprint, *Research Square*, Version 1, Jul. 29, 2024. [Online]. Available: https://doi.org/10.21203/rs.3.rs-4691427/v1

[2] F. Chaudhari, R. Patel, K. Khamar, V. Patel, J. Patel, and H. Trivedi, "Cleanliness Follow-up using Deep Learning," in *Proc. 2023 Global Conf. on Information Technologies and Communications (GCITC)*, Karnataka, India, Dec. 2023, pp. 1–6, doi: 10.1109/GC

[3] A. K. Sharma, A. Jain, D. Chaudhary, S. Tiwari, H. Mahdin, Z. Baharum, S. M. Shaharudin, R. Maskat, and M. S. Arshad, "An Approach to Automatic Garbage Detection Framework Designing using CNN," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 14, no. 2, pp. 257–264, 2023.
[4] M. Panmuang and C. Rodmorn, "Garbage Detection using YOLO Algorithm for Urban Management in Bangkok," *WSEAS Transactions on Computer Research*, vol. 12, pp. 236–241, 2024. doi: 10.37394/232018.2024.12.23.
[5] Z. Xia, H. Zhou, H. Yu, H. Hu, G. Zhang, J. Hu, and T. He, "YOLO-MTG: A Lightweight YOLO Model for Multi-Target Garbage Detection," *Signal, Image and Video Processing*, vol. 18, pp. 5121–5136, May 2024. doi: 10.1007/s11760-024-03220-2.
[6] T. J. Sheng, N. Misran, M. S. Islam, M. H. Baharuddin, H. Arshad, M. R. Islam, H. Rmili, M. E. H. Chowdhury, and M. T. Islam, "An Internet of Things Based Smart Waste Management System Using LoRa and Tensorflow Deep Learning Model," *IEEE Access*, vol. 8, pp. 148793–148811, Aug. 2020, doi: 10.1109/ACCESS.2020.3016255.
[7] A. Mitra, "Detection of Waste Materials Using Deep Learning and Image Processing," M.S. project, Dept. of Computer Science, California State University San Marcos, San Marcos, CA, USA, Dec. 2020.
[8] Z. Chunxiang, Q. Jiacheng, and W. Binrui, "YOLOX on Embedded Device With CCTV and TensorRT for Intelligent Multicategories Garbage Identification and Classification," *IEEE Sensors Journal*, vol. 22, no. 16, pp. 16522–16529, Aug. 2022, doi: 10.1109/JSEN.2022.3181794.
[9] "Definition of precision,". [Accessed: Dec. 20, 2022].
[10] "Definition of recall," [Accessed: Dec. 20, 2022].
[11] "Mean Average Precision (mAP) formula," [Accessed: Dec. 20, 2022].