Classifying Waste: Multiclass Model for Litter Detection and Classification.

Uma Shankar N
Department of Mathematics
Indian Institute of Technology
Jodhpur, India
m22ai657@iitj.ac.in

Kathijaafrose.A
Department of Mathematics
Indian Institute of Technology
Jodhpur, India
m22ai564@iitj.ac.in

Vignesh Pitchaiah
Department of Mathematics
Indian Institute of Technology
Jodhpur, India
m22ai660@iitj.ac.in

***Abstract*—This research introduces an innovative system that utilizes advanced deep learning techniques, with a particular emphasis on YOLOv7, to achieve precise identification and classification of non-biodegradable waste in urban environments. By harnessing cutting-edge image recognition technologies, the system can detect and categorize non-biodegradable waste objects in images, marking a significant advancement in waste management practices. Through accurate quantification of volume and spatial coordinates of these waste items, the system has the potential to enhance cost-efficiency, optimize time, and improve resource allocation in waste management processes. The study employs an inventive approach that integrates multiple deep learning models, including YOLOv7, to elevate the accuracy and reliability of waste object detection. The system's capabilities span object detection, localization, and classification, enabling the automated identification of various non-biodegradable waste materials commonly found in urban areas. By capitalizing on YOLOv7's strengths, including real-time waste object detection and precise bounding box localization, the model is fine-tuned to excel in swiftly and effectively identifying non-biodegradable waste items while maintaining a balance between speed and accuracy. Moreover, the research utilizes the TACO dataset, curated for waste object detection and segmentation, providing diverse real-world urban waste scenarios for training and evaluation.This system shows great promise in revolutionizing waste management practices by reducing the need for human intervention and enhancing the efficiency of urban waste handling.**

***Index Terms*—Deep Learning, YOLOv7, Waste Management, Object Detection, Non-Biodegradable Waste, Image Recognition,Model Architecture ,Optimization Technique , Model Training, Waste Segregation**

I. Introduction

Solid waste management stands as a formidable and pressing challenge in today's world. The unrelenting surge in waste production driven by human activities poses grave threats to our environment, public health, and overall societal well-being. According to estimates from the World Bank, global annual waste generation is projected to soar from 2.01 billion tonnes in 2016 to a staggering 3.40 billion tonnes by the year 2050 [2]. This ominous trajectory underscores the urgency of addressing the multifaceted complexities of waste management.

Compounding this challenge is the ever-increasing diversity and intricacy in the composition of waste materials. This diversity adds layers of complexity, making the task of sorting and disposing of waste a daunting endeavor. Failing to manage waste properly carries dire consequences, including but not limited to pollution, the release of greenhouse gases, depletion of precious resources, land degradation, and heightened risks of disease transmission. In response to these pressing concerns, our study ventures into the critical domain of waste management, employing innovative approaches and advanced technologies. At the core of our solution lies a powerful deep learning model, which forms the linchpin of our waste management strategy.

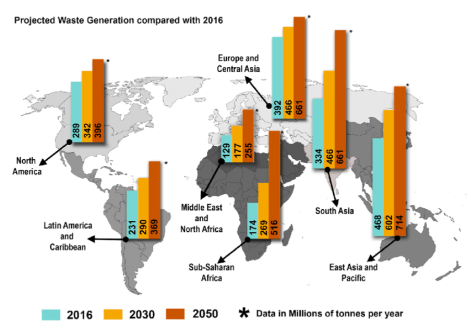

Fig. 1. Global Waste Generation Survey: [12]

Our research primarily focuses on the comprehensive performance analysis of the YOLOv7 deep learning model in the specific domain of waste object detection and segmentation. YOLOv7, an advanced deep learning model, leverages cutting-edge techniques in computer vision and image recognition to excel in swiftly and accurately detecting, categorizing, and sorting waste materials. By analyzing waste images, YOLOv7 can rapidly and precisely identify non-biodegradable materials, facilitating their effective segregation from biodegradable ones. This breakthrough not only streamlines waste sorting processes but also contributes to minimizing the environmental impact by ensuring proper disposal and recycling.

II. Literature Survey

In the realm of computer vision for non-biodegradable waste detection and classification, we have identified notable works conducted within the past five years that have contributed significantly to this field:

Early Work by Ashim Dey et al. (2021) [6]: Ashim and his team presented a computer vision-based system explicitly designed for non-biodegradable waste detection and classification, with a focus on garbage management and recycling. They employed a convolutional neural network (CNN) to classify waste objects into six distinct non-biodegradable categories: organic, plastic, glass, paper, metal, and cardboard. Their dataset comprised 25,215 images of non-biodegradable waste objects obtained from the Garbage Classification dataset. Remarkably, their system achieved an impressive accuracy rate of 93.6 Percentage using a sigmoid classifier.

Limitations: The paper's method may not work well for some waste objects that have more than one type of material, like a bottle with a metal cap as they uses a sigmoid classifier as the final layer of the CNN. It also does not show how it compares to other methods that can do the same task.

Work by Nadish Ramsurrun et al. (2022) [7]: They proposed the deep learning approach using computer vision to automatically identify the type of waste and classify it into five main categories: plastic, metal, paper, cardboard, glass1 and described a conceptual system of an automated recycling bin that opens the lid corresponding to the type of waste identified1.They used a pre-existing images to train 12 variants of the convolutional neural network (CNN) algorithm over three classifiers: support vector machine (SVM), sigmoid and softmax1. Using pre built VGG19 with softmax classifier they achived an accuracy of 88.3 Percentage

Limitations: The paper does not compare its results with any existing methods or baselines1 and failed to evaluate the performance of the system in real-world scenarios or test its robustness to noise, occlusion, or varying lighting conditions1.

Work by Islam and Alam (2022) [8]: Islam and Alam proposed a pioneering deep learning-based model tailored for non-biodegradable waste management, particularly in smart cities and IoT environments. Their approach featured object detection using state-of-the-art models, EfficientDet and YOLOv5, combined with a functional link neural network (FLNN) for object classification. They curated an extensive dataset comprising 2,200 images of waste dumps, encompassing 135,541 annotated objects from five non-biodegradable categories: plastic, metal, paper, cardboard, and glass. Their results showcased a mean average precision (mAP) of 66.08 Percentage at an intersection over union (IoU) threshold of 0.35, achieving inference speeds of 55-58 milliseconds on a single GPU.

Limitations : The system employs two distinct object detection models, EfficientDet and YOLOv5, which may compromise the computational efficiency and resource utilization of the system. It also utilizes a functional link neural network (FLNN) for object classification, which may struggle to cope with variations in the input images, such as lighting, occlusion, or noise. A more sophisticated system may adopt other deep learning techniques or methods to enhance the accuracy and reliability of the waste classification.

Fig. 2. deep learning-based model tailored for non-biodegradable waste management. Image source: [6]

Work by Miguel Valente (2019) [9]: They aim to identify different kinds of waste containers using computer vision. It tries two methods: one based on feature detection and description, and another based on deep learning. The first method uses VLAD, a technique that summarizes local features of an image. The second method uses YOLO, a deep neural network that can detect objects in real time. This method achieves a high accuracy of around 90 Percentage.

Limitations : The first method does not work well for this task so they went with the second one but they did not provide sufficient details about the implementation and evaluation of the methods.

These seminal works collectively advance the domain of non-biodegradable waste detection and classification, demonstrating the potential for computer vision and deep learning to revolutionize waste management practices. Their insights, methodologies, and results contribute significantly to the development of intelligent and sustainable waste management solutions for non-biodegradable materials.

III. Dataset

The dataset is called "TACO" (Trash Annotations in Context) [16], and it is a dataset specifically curated for object detection and segmentation tasks related to waste objects. TACO is designed to support computer vision research, particularly in the context of waste management and environmental applications.

OverView of Dataset: The dataset contains a wide variety of images captured in real-world urban environments. These images showcase different types of waste objects commonly found on streets, sidewalks, and public spaces of total 6000 Images. The waste objects include items like plastic bottles, cardboard boxes, cans, and various other non-biodegradable materials.

IV. Work Implementation

Data Preparation :

- **Data Collection:** We started with an existing dataset focused on tacos, and to enhance it, we incorporated images specifically showcasing bottle lids. These new images were meticulously labeled and annotated by hand, ensuring that each bottle lid was accurately identified within the dataset.Next we Standardize image size, format, and quality across the dataset to ensure consistency for efficient annotation.suitable annotation tool compatible with taco object annotation is selected. Tools like LabelImg bounding box annotation is used for annotations.

Annotation Process for Customized Data:

- **Bounding Box Annotation:** Imported images into the LabelImg annotation tool.Manually draw bounding boxes around individual tacos in each image, ensuring the boxes tightly enclose the taco without including unrelated background elements.
- **Class Labeling:** Assigned a specific class label 'Aluminium foil', 'Bottle cap', 'Bottle', 'Broken glass', 'Can', 'Carton', 'Cigarette', 'Cup', 'Lid', 'Other litter', 'Other plastic', 'Paper', 'Plastic bag - wrapper', 'Plastic container', 'Pop tab', 'Straw', 'Styrofoam piece', 'Unlabeled litter' to each annotated bounding box to indicate the object category.

Data Preprocessing :

- **Converting Dataset into YOLO Format:** Current Dataset Format is in COCO-like format, where images and annotations are stored in separate files. The goal is to create a new dataset with two folders, "images" and "labels." Image files are stored in the "images" folder, and corresponding text files with the same filenames are placed in the "labels" folder. The text files follow YOLO format. YOLO format represents each image with a corresponding text file. The text file contains rows with class IDs, bounding box coordinates (x, y, width, height), and other details.
- **Implementation:** The process involves importing required packages, reading JSON annotation files, processing images, and creating helper functions to extract image and annotation details. It normalizes bounding box properties and saves them in YOLO format text files.The outcome is a successfully converted dataset in YOLO format, ready for object detection model training.all images were resized to a fixed dimension of 416x416, employing a stretch method to maintain consistency in size and aspect ratio, facilitating efficient processing for subsequent stages.
- **Dataset Splitting:** Divide the annotated dataset into training, validation, and test sets. Allocate images carefully to each set to ensure a balanced representation of tacos across all subsets.

Fig. 3. TACO Data Performance Result:

V. Method Used

YOLO:

YOLO is a deep learning-based object detection algorithm known for its speed and accuracy. It's designed to detect and locate objects within images or video frames efficiently. YOLO's key innovation is its ability to make real-time object detection possible by dividing the image into a grid and predicting bounding boxes and class labels for objects within each grid cell in a single forward pass of the neural network. YOLO has seen several iterations. It has been widely adopted in various applications, including waste object detection, where it plays a crucial role in automating waste sorting and management processes.

Object Detection: Object detection can be classified into two main types:

- **Single-Stage Detectors:** These models directly predict bounding boxes and class labels in a single step. YOLO (You Only Look Once) is a popular example of a single-stage detector.

- **Two-Stage Detectors:** These models first propose regions of interest and then classify and refine the bounding boxes within those regions. Faster R-CNN and R-CNN are examples of two-stage detectors.

Motivation to Use YOLOv7: The decision to opt for YOLOv7 was primarily driven by its demonstrated superiority over YOLOv3 and YOLOv5 in the realm of waste detection. Its architectural improvements, notably the integration of VOVNET and ELAN alongside other techniques, play a pivotal role. In previous Versions of YOLO small Objects like Bottle Cap won't get detect,But in our project main aim is to make detect the small objects like bottle cap.

YOLOv7 for Waste Object Detection:

YOLO V7 offers an advantage by efficiently identifying waste items using fewer parameters. It utilizes nine anchor boxes to detect various waste shapes, effectively reducing false detections. YOLO V7 introduces a focal loss function for improved detection of smaller waste objects. With a higher image resolution (608x608 pixels), it excels in pinpointing smaller waste items swiftly, processing images at 155 frames per second. This balance between speed and accuracy makes it ideal for real-time waste management applications, despite being slightly less accurate than other detectors that require more computational resources.

Architechure:

YOLO V7 comprises four distinct layers: the Input Layer, Backbone Layer, Neck Layer, and Head Layer.

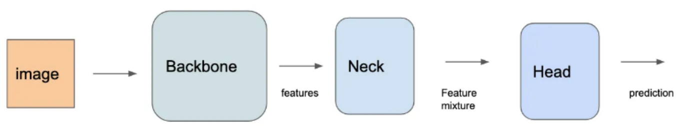

Fig. 4. Architechure: [17]

- **Input Layer:**
  The Input Layer is where the network receives the initial data, in this case, the image for waste object detection. It resizes and preprocesses the input image to a format suitable for deep learning. The image is often divided into small, fixed-size regions called grid cells, which serve as the basis for object detection.

- **Backbone Layer:** The Backbone Layer plays a fundamental role in the YOLO V7 architecture, responsible for feature extraction from the Input Image. It employs advanced deep neural network architectures, including VOVNET and ELAN. These architectures are designed to analyze the input image at multiple scales, allowing the model to capture features and details of different sizes.For example, VOVNET is known for its ability to process images efficiently across various scales, which is vital for detecting waste objects of different sizes in urban environments. The Backbone Layer focuses on identifying intricate patterns, shapes, and structures within the image, which are essential for the precise and accurate detection of non-biodegradable waste materials.

- **Neck Layer:** The Neck Layer connects the Backbone Layer to the Head Layer and plays a role in feature fusion and enhancement. It often combines features from different scales to ensure that objects of various sizes are detected effectively. Techniques like skip connections, feature pyramids, and feature concatenation are commonly used in this layer to improve the network's ability to detect waste objects.

- **Head Layer:** The Head Layer is the ultimate decision-making component within YOLO V7's architecture for waste object detection. This layer takes advantage of the information extracted by the Backbone and Neck Layers to make predictions regarding the presence, precise location, and category of non-biodegradable waste objects. To achieve this, YOLO V7 leverages the concept of anchor boxes. Anchor boxes are predefined bounding boxes with specific aspect ratios used in the detection process. They allow the model to make highly accurate identifications and classifications of waste materials in various shapes and sizes. The Head Layer manages a crucial post-processing step called non-maximum suppression. This step is responsible for eliminating redundant detections by considering the most confident predictions and discarding others. Non-maximum suppression ensures that the final output of the model includes only the most accurate and relevant information regarding the waste objects identified in the input image.

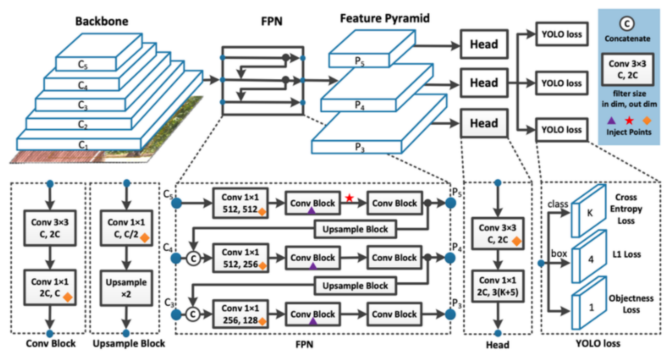

Fig. 5. Architechure With Layers: [18]

Model Training: In YOLOv7 model training, several key innovations were introduced to enhance object detection accuracy and efficiency:

- **Extended Efficient Layer Aggregation:** To improve efficiency, YOLOv7 uses layer aggregation techniques, specifically the E-ELAN computational block. This approach optimizes memory usage and gradient propagation through layers, resulting in more powerful learning.

- **Model Scaling Techniques:** YOLOv7 scales network depth and width together while concatenating layers. This allows the model to adapt to different accuracy and inference speed requirements, keeping the architecture optimal for various sizes.

- **Re-parameterization Planning:** Gradient flow propagation paths are used to determine which modules in the network should employ re-parameterization strategies. This technique enhances the network's ability to capture general patterns.
  Gradient flow propagation can be understood as the way gradients (derivatives) are back-propagated through a

neural network during training. It's a fundamental concept in gradient-based optimization algorithms like stochastic gradient descent. The mathematical formula for gradient flow propagation typically involves the chain rule of calculus, which calculates how gradients change as they move backward through the layers of a neural network.

$$\Delta w_{i,j} = -\eta \frac{\partial E}{\partial w_{i,j}}$$

- $\Delta w_{i,j}$ represents the change in weight for the connection between neuron i and neuron j.
- η is the learning rate, which controls the step size for weight updates.
- $\frac{\partial E}{\partial w_{i,j}}$ is the partial derivative of the error (loss) function E with respect to the weight $w_{i,j}$.

The process of backpropagation involves calculating $\frac{\partial E}{\partial w_{i,j}}$ for each weight in the network, starting from the output layer and moving backward through the hidden layers. This is done iteratively for each training example to adjust the weights and minimize the error during the training process.

The specific mathematical expression for $\frac{\partial E}{\partial w_{i,j}}$ depends on the choice of loss function and the architecture of the neural network. It is calculated by applying the chain rule and may involve the derivative of the activation functions and the weights of the subsequent layers. The goal is to find the optimal weights that minimize the error function E, which represents the discrepancy between the network's predictions and the actual target values.

- **Auxiliary Head Coarse-to-Fine:** YOLOv7 introduces an auxiliary head in the middle of the network. During training, this auxiliary head is supervised along with the final head. The authors experiment with different levels of supervision, settling on a coarse-to-fine approach, where supervision is passed back from the lead head at various granularities.

Model Loss Functions: In YOLOv7, the loss function is a critical element for training the model to enhance its object detection capabilities. This loss function consists of three key components:

- **Classification Loss:** The classification loss, often referred to as Categorical Cross-Entropy, quantifies the disparity between the predicted class probabilities and the actual class, which can be either the object class or the background. It plays a pivotal role in guiding the model to assign high probabilities to the correct object class. The mathematical formulation for the classification loss is as follows:

$$L_{\text{cls}} = -\sum_{i=1}^{N}\sum_{j=1}^{C} y_{ij}\log(\hat{y}_{ij})$$

- N represents the total number of grid cells.
- C signifies the number of distinct classes.
- y_{ij} assumes a value of 1 if the object belongs to class j in grid cell i, and 0 otherwise.
- $\hat{y}_{ij}$ represents the model's predicted probability for class j in grid cell i.

This loss function's purpose is to minimize the difference between predicted class probabilities and the true object class to enhance the model's classification accuracy.

- **Localization Loss:** Measures the error in predicting the boundary box coordinates (x, y, width, height) for the object.Encourages the model to predict accurate bounding box locations. The mathematical formula for the localization loss (smooth L1 loss) is:

$$L_{\text{loc}} = \sum_{i=1}^{N}\sum_{j=1}^{B} \left[\text{SmoothL1}(\hat{t}_{ij} - t_{ij})\right]$$

- N is the number of grid cells.
- B is the number of bounding boxes per grid cell.
- $\hat{t}_{ij}$ is the predicted bounding box coordinates for box j in grid cell i.
- t_{ij} is the ground truth bounding box coordinates.

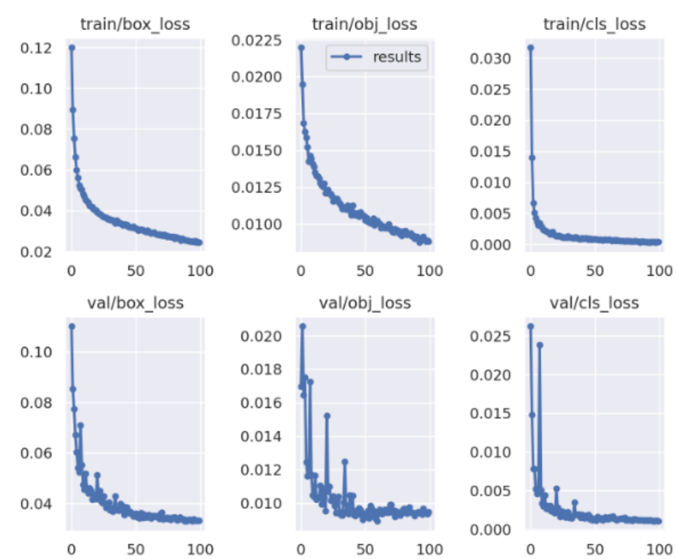

Fig. 6. Output Graph of Training and Validation Loss vs Epochs(Loss Decreases in increasing the Epochs of Training): [18]

- **Confidence Loss :**
The confidence loss, often referred to as the objectness loss, measures the network's capability to discern if an

object is present within a particular bounding box. It uses a mathematical formula to compute this loss:

$$L_{\text{conf}} = -\sum_{i=1}^{N}\sum_{j=1}^{B}\sum_{k=1}^{S} \mathbb{1}_{ij}^{k}\left[\log(\hat{p}_{ij}^{k}) + (\alpha)\log(1-\hat{p}_{ij}^{k})\right]$$

- N is the total number of grid cells.
- B represents the number of anchor boxes.
- S is the grid size (usually $S = 13$ in YOLO).
- $\mathbb{1}_{ij}^{k}$ is an indicator function that equals 1 when the k-th anchor box in grid cell i is responsible for predicting the object's presence, and 0 otherwise.
- $\hat{p}_{ij}^{k}$ denotes the predicted confidence score of the k-th anchor box in grid cell i.
- α is a weighting factor, typically used to balance the loss components.

The confidence loss serves to encourage the model to predict high confidence scores when the anchor box is responsible for object prediction and low scores when it is not.

Performance Evaluation:

- **Intersection over Union (IoU):**The performance of an object detection model is often measured using several metrics, and one of the key metrics is the Intersection over Union (IoU). IoU computes the intersection over the union of the predicted bounding box and the ground truth bounding box. It is defined mathematically as:

$$IoU = \frac{\text{Area of Intersection}}{\text{Area of Union}}$$

- The "Area of Intersection" is the region where the predicted bounding box and the ground truth bounding box overlap.
- The "Area of Union" is the total area encompassed by both the predicted bounding box and the ground truth bounding box.
An IoU of 1 indicates that the predicted and ground-truth bounding boxes perfectly overlap, while an IoU of 0 means there is no overlap.To determine if an object detection is valid or not based on IoU, a threshold value, often denoted as $IoU_{\text{threshold}}$, is set. Commonly, a threshold value of 0.5 is used. The classification is as follows:
- If $IoU \geq IoU_{\text{threshold}}$, the object detection is classified as a True Positive (TP), indicating that the predicted bounding box is a valid detection that aligns well with the ground truth.
- If $IoU < IoU_{\text{threshold}}$, the object detection is considered a False Positive (FP), meaning that the predicted bounding box does not align well with the ground truth. This threshold allows us to control the trade-off between precision and recall, making it possible to tailor the object detection model's performance to specific needs.

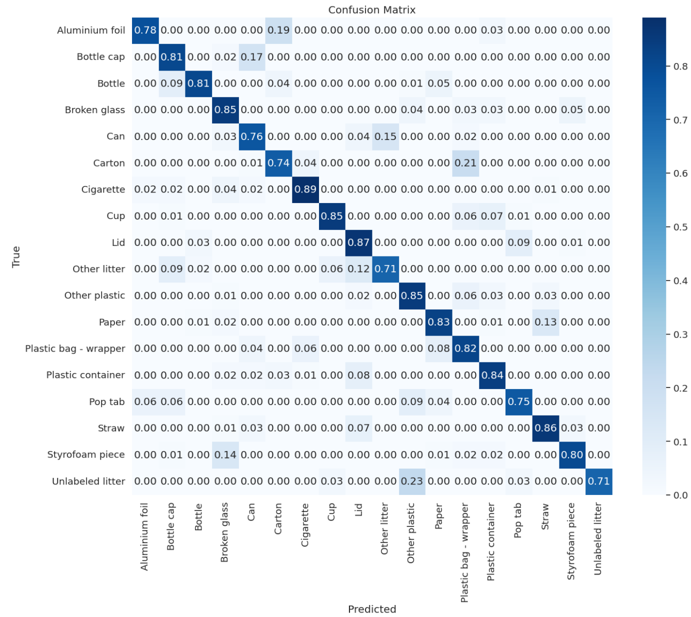

Fig. 7. Confusion Matrix: [18]

The confusion matrix provided indicates the model's classification performance across various waste item classes. It illustrates the proportions of correct and incorrect predictions for each class, detailing how the model assigned instances to different classes. For instance, "Broken glass" had 85% correct classifications, primarily predicted as itself, but with some confusion with "Can" and "Styrofoam piece." Similarly, "Can" had 76% accurate predictions but exhibited misclassifications with "Other litter" and "Styrofoam piece." From this matrix, precision and recall scores for each class can be computed, reflecting the model's ability to accurately identify specific waste items against misclassifications and overall performance across all classes.

- **Mean Average Precision (mAP):** The Mean Average Precision (mAP) for object detection is calculated by first computing the Average Precision (AP) for each individual class and then taking the mean of these AP values across all classes. The formula for mAP is as follows:

$$mAP = \frac{1}{n}\sum_{i=1}^{n} AP_i$$

- mAP is the Mean Average Precision.
- n is the number of classes.
- AP_i is the Average Precision for class i.
Each AP_i is computed based on precision-recall values at different Intersection over Union (IoU) thresholds, and the average of these values across all classes provides a comprehensive assessment of the model's overall performance in object detection.
Achieving precision and recall scores of 0.8 for both metrics on the training data showcases the model's ability to make accurate detections while minimizing

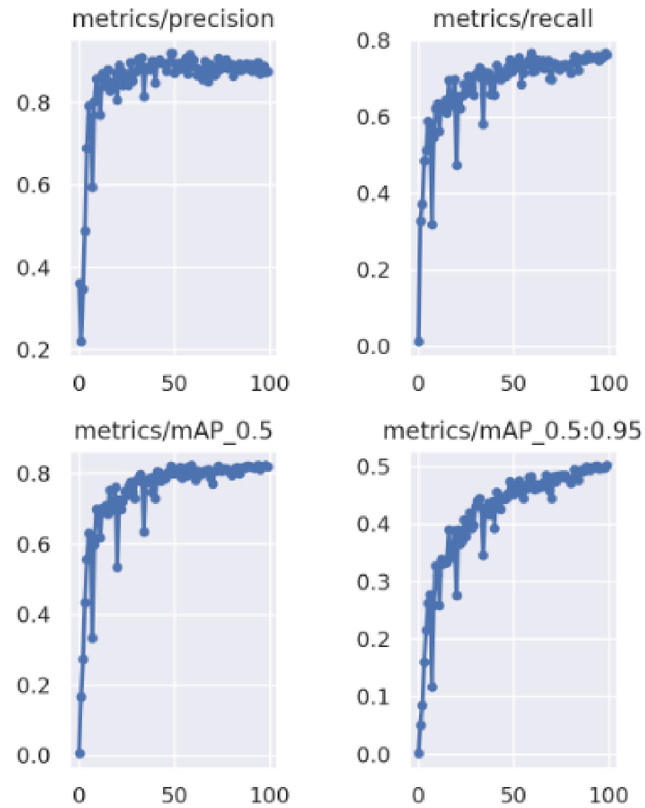

Fig. 8. Precision Recall Curve in Each Epochs: [18]

false positives and false negatives. This signifies a high level of confidence in the model's predictions within the training set. Moreover, maintaining a precision of 0.8 on the test data indicates the model's generalizability and robustness in identifying non-biodegradable waste items. Although the test recall at 0.7 is slightly lower, the overall performance remains strong, suggesting the model's capability to effectively recognize relevant waste objects while maintaining a good level of precision in real-world scenarios.

Model Training :

We initiated the training process using a partial dataset, and we are actively analyzing the model's performance. Our rationale for employing a partial dataset is to develop a finely tuned model that operates efficiently with limited computational resources, minimizing errors. Our ongoing work aims to yield optimal results by implementing the techniques outlined in Section 3 with complete Dataset.

Conclusion: In this study, we've undertaken a systematic approach to tackle non-biodegradable waste management through computer vision and deep learning techniques. Leveraging the TACO dataset and augmenting it with meticulously annotated bottle lid images, we've prepared a tailored dataset crucial for training our model. The adoption of YOLOv7, specifically chosen for its advancements in detecting smaller waste items, reflects our dedication to precision. Through meticulous data preparation, conversion to YOLO format, and careful dataset splitting, we've structured a robust workflow. Our model's architecture, incorporating VOVNET, ELAN, and innovative techniques like layer aggregation and re-parameterization planning, aims to optimize efficiency and accuracy. Understanding loss functions and key performance metrics like IoU and mAP has guided our assessment of model performance. While our initial training used a partial dataset to fine-tune the model, ongoing work involves comprehensive training with the complete dataset. This methodological approach lays a strong foundation for an intelligent waste detection system poised to revolutionize waste management practices.

Acknowledgment

We extend our heartfelt thanks to the diligent reviewers who dedicated their time to review our paper and provide insightful feedback. Your valuable comments have been instrumental in refining our work. We've carefully addressed your suggestions and resubmitted the paper for further review. Additionally, we would like to highlight that, as part of our commitment to innovation, we have conducted a thorough validation to ensure that the idea we are proposing is novel and not currently in existence. Thanks for your valuable input!

References

[1] Code: https://github.com/N-UmaShankar/Project

[2] S. Sudha, M. Vidhyalakshmi, K. Pavithra, K. Sangeetha and V. Swaathi, "An automatic classification method for environment: Friendly waste segregation using deep learning," 2016 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR), Chennai, India, 2016, pp. 65-70, doi: 10.1109/TIAR.2016.7801215.

[3] Gupta, T., Joshi, R., Mukhopadhyay, D. et al. A deep learning approach based hardware solution to categorise garbage in environment. Complex Intell. Syst. 8, 1129–1152 (2022). https://doi.org/10.1007/s40747-021-00529-0.

[4] Brintha, V.P., Rekha, R., Nandhini, J., Sreekaarthick, N., Ishwaryaa, B., Rahul, R. (2020). Automatic Classification of Solid Waste Using Deep Learning. In: Kumar, L., Jayashree, L., Manimegalai, R. (eds) Proceedings of International Conference on Artificial Intelligence, Smart Grid and Smart City Applications. AISGSC 2019 2019. Springer, Cham. https://doi.org/10.1007/978-3-030-24051-6_83

[5] Jerie, Steven. (2016). Occupational Risks Associated with Solid Waste Management in the Informal Sector of Gweru, Zimbabwe. Journal of Environmental and Public Health. 2016. 1-14. 10.1155/2016/9024160.

[6] R. Faria, F. Ahmed, A. Das and A. Dey, "Classification of Organic and Solid Waste Using Deep Convolutional Neural Networks," 2021 IEEE 9th Region 10 Humanitarian Technology Conference (R10-HTC), Bangalore, India, 2021, pp. 01-06, doi: 10.1109/R10-HTC53172.2021.9641560.

[7] N. Ramsurrun, G. Suddul, S. Armoogum and R. Foogooa, "Recyclable Waste Classification Using Computer Vision And Deep Learning," 2021 Zooming Innovation in Consumer Technologies Conference (ZINC), Novi Sad, Serbia, 2021, pp. 11-15, doi: 10.1109/ZINC52049.2021.9499291.

[8] Islam, S.M.Y., Alam, M.G.R. (2022). Computer Vision-based Waste Detection and Classification for Garbage Management and Recycling. In: Hossain, S., Hossain, M.S., Kaiser, M.S., Majumder, S.P., Ray, K. (eds) Proceedings of International Conference on Fourth Industrial Revolution and Beyond 2021 . Lecture Notes in Networks and Systems, vol 437. Springer, Singapore. https://doi.org/10.1007/978-981-19-2445-3_10

[9] M. Valente, H. Silva, J. M. L. P. Caldeira, V. N. G. J. Soares and P. D. Gaspar, "Computer Vision Approaches to Waste Containers Detection," 2019 14th Iberian Conference on Information Systems and Technologies (CISTI), Coimbra, Portugal, 2019, pp. 1-4, doi: 10.23919/CISTI.2019.8760862.

[10] Wang, Chien-Yao & Bochkovskiy, Alexey & Liao, Hong-yuan. (2022). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. 10.48550/arXiv.2207.02696.

[11] https://www.rts.com/blog/the-benefits-of-waste-management-for-business/

[12] Kaza, S., Yao, L.C., Bhada-Tata, P. and Van Woerden, F. (2018) What a Waste 2.0: A Global Snapshot of Solid Waste Management to 2050. Urban Development. World Bank, Washington DC. https://doi.org/10.1596/978-1-4648-1329-0 https://openknowledge.worldbank.org/handle/10986/30317

[13] https://www.newindianexpress.com/states/kerala/2023/apr/14/kerala-removes-over-8600-tonnes-of-non-biodegradable-waste-in-two-months-2565787.html

[14] https://www.ucl.ac.uk/global/case-studies/2021/sep/training-robot-sort-recycling-using-deep-learning

[15] https://www.thebetterindia.com/180488/delhi-plastic-waste-shirt-bags-management-india/

[16] https://github.com/pedropro/TACO/tree/master/data/

[17] https://medium.com/@nahidalam/understanding-yolov7-neural-network-343889e32e4e/

[18] https://blog.roboflow.com/yolov7-breakdown/