Smart City Surveillance: Automated Street Waste Detection Using Live Camera

MSc Research Project
Master of Science in AI

Rehman Sarwar
Student ID: x23245875

School of Computing
National College of Ireland

Supervisor: Syed Abidi

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Rehman Sarwar
Student ID:	X23245875
Programme:	Master in Artificial Intelligence **Year:** 2024-2025
Module:	Research Master in AI
Supervisor:	SM Raza Abidi
Submission Due Date:	12 Dec 2024
Project Title:	Smart City Surveillance: Automated Street Waste Detection Using Live Camera
Word Count:	6054 **Page Count**:19

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Rehman Sarwar

Date: 12 Dec 2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	☑
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☑
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☑

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Smart City Surveillance: Automated Street Waste Detection Using Live Camera

Rehman Sarwar
X23245875

Abstract

The study proposes an Automated Waste Source Separation system that utilizes the YOLOv8 models in detecting and classifying waste on the streets of modern cities. The system has two versions, YOLOv8s and YOLOv8l, both of which are built off more than six thousand images obtained from a diverse collection of cities. The images include organic waste, metal waste, glass waste, paper waste, and plastic waste. Such specific application aims to simplify low-complexity detection tasks while achieving reasonable performance with the YOLOv8s model with encoder of 0.9058 precision, 0.94279 recall and 0.96042 mean Average Precision score. On the other hand, the YOLOv8l model is reported to be less accurate in some scenes but provides better results in the task of detecting and classifying objects in cluttered scenes. Two models YOLOv8l and YOLOv8s are tested concerning precision, recall, F1 scores and mAP, whereby the concept of transfer learning was applied to make use of the pre-trained weights for minimization of training cycles. The system provides the highest performance results and hence is able to provide a great solution for a fast and effective waste segregation within the contemporary cities. The inclusion of YOLOv8 models in this system allows efficient waste detection which will ensure that the urban landscape is clean and free from dangerous diseases.

Keywords: Real-time waste segregation, YOLOv8small, YOLOv8large, waste detections

1 Introduction

Inadequate infrastructure in large and overpopulated cities makes climate and waste challenges to be one of the major issues. However, as urban community extends this also means that a great deal of Waste is produced, leading to major ecological and even health challenges. An uncontrolled waste can cause land pollution affecting pests and diseases spreads thereby endangering the people in the neighbourhood at health risks (Cook et al., 2024).

There is also the demographic perspective to consider, since cleaning and maintenance tends to be very costly and continues to become a allow from keeping them up to the waste that is ever increasing (Kumar et al., 2017). Due to lack of remote management of waste facilities collection this can lead to a mass accumulation of waste and therefore lead to a more complex issue (Wang et al., 2024). This is however changing in a bigger number of cities adopting new methods of dealing with waste, that touches on the sustainability of the environment. It overrides the methods used nowadays due to its promising future with automated street waste detection that uses cameras and video trained models (Erin et al., 2022).

This technology makes it possible to monitor the state of garbage containers without any difficulties, which improves the efficiency of their collection. The use of such implementations would allow the cities to free the negative impact that the waste has on the environment and even the health making the city a better and cleaner place for the people

The efficient growth of new urban areas increases the need to come up with the novel approaches of waste management, especially in the highly populated (Mishra et al., 2019). As consumer population and purchases as well go up, waste management has become one of the global challenges. Inadequate waste disposal affects people's health and pollution control, hence the need for cleaner measures Kumar et al., 2017).

To deal with this concern, innovative solutions are being developed for easy identification and management of waste. For example, vision systems based on computer images can easily recognize wastes without human involvement, providing effective and timely waste retrieval response (Mitra and Li, 2020). Operational cameras are also great as they allow for effective and quick enabling and spotting of garbage around the area. Incorporation of smart systems into smart city programs is better as it helps provide cities with avenues for development (Tamakloe and Rosca, 2020). Therefore, by doing away with the search for the location of wastage, you will see that cities will get the best capacity hence leading to cleaner and healthier cities in the long run.

1.1 Research Objectives

- Develop and deploy an automated waste detection system utilizing YOLOv8 to effectively recognize and categorize different types of waste in real-time.
- To get high accuracy, develop and optimize the YOLOv8 model to ensure high accuracy in waste classification
- To evaluate the performance and precision of purposed model in identifying and sorting various types of waste

1.2 Research Questions

- In what ways can YOLOv8 be optimally specified and deployed to build a smart waste identification framework that can distinguish several types of waste after they have been captured in real time
- How can the techniques record high accuracy on the YOLOv8 algorithm to improve the waste detection across the different types of waste and across the environments?
- What is the effectiveness of the YOLOv8 algorithm in detection and classification of different types of wastes for smart cities?

1.3 Contribution

This research introduces an automated street waste detection system accompanied by an optimized Computer Vision model referred to as YOLOv8. It is true this system can monitor by various live camera feeds, and therefore enhances the speed and accuracy of observing wastes in the urban areas. They enable municipalities to identify littering areas well before these areas accumulate large quantities of litter; hence, appreciable efficiency and timely responses are made possible in waste management. Furthermore, the lessons learned from this study will be of benefits when developing improved cheap waste collection.

1.4 Structure of the Paper

The study is structured as follows: In Section 2, the related works concerning the waste detection techniques and efforts are discussed. Section 3 of paper is devoted to the methodologies, the design and the construction of the proposed automated waste detection system. The result of the study is presented in section 4 and the effectiveness of the applied method and modelling techniques used is also evaluated here. Finally, there is a discussion in section 5 that accompanies evaluation of the model and section 6 presents the conclusions and limitations of the study

2 Related Works

Despite its global significance the proper management of solid waste remains a challenge due to increased generation and insufficient disposal facilities. Various scholars and organizations have done their studies via new strategies that use other technologies such as YOLO deep learning algorithm for waste recognition and sorting (Erin et al., 2022), and other extended the scopes of stereo camera system with add loan4 to improve the sorting of waste while (Zailan et al., 2022) a modified YOLO algorithm for floating debris detection in rivers. Finally, concerning problems of this kind solving such problems, many works including those described above of a descriptive nature, use of robotic or automated systems in waste management, as well as challenges related to recyclables and pollution.

In the study of (Erin et al., 2022) explores the issue of waste management since the level of waste generation is rising while the landfill space is diminishing. The findings of this study provide an effective approach to trash disposal management coupled with a 3D camera system and YOLOv4 algorithm to detect solid waste. However, the factors on their study point out that there is still much to be researched even if the outcomes are positive. Finally, proper waste management is highly imperative to the modern society but in order to address the issues at hand proffer adequate solutions need to be made. The study also notes that establishing separation techniques of recyclables like plastics have become 'a' increasingly elusive process in mature economies. Thus, it can be stated that there are unsuitable to apply manual sorting techniques for typical removal of large amounts of waste.

On the other hand, the efficiency of retrieval of wastes decreases and the input of operators is minimized in automatic sorting systems. Image training of the waste content image sets with background images was done using YOLOv4 algorithm training. This multi-functional equipment can sort the metal, paper, glass, plastic waste and other types of wastes. The Intel RealSense D415 3D camera is used as the core of this system to collect accurate 3D data for the integration and sorting of wastes. Analysis of the outcomes reveal that YOLOv4 has higher accuracy over YOLOv4-tiny in computational efficiency. In future expansion of the recycling facility work relevant robotics handling devices are expected to be installed in every recycling station where wastes placed on the belt will be sorted to the highest degree and this almost eliminate manual handling.

In the research of (Zailan et al., 2022) have identified that the issue with the accumulation of solid wastes in regional rivers has become quite common and for the effective maintenance of the riverine environment, owing to the process of urbanization, which is said to be capable of becoming a risk to both, ecosystems and human beings. In their study, they developed an automatic detection system which they used an improved YOLO model to target the problem with regard to floating debris detection on rivers. This system aims at helping in creating the required self-operating cleaning robots that will effectively control pollution in rivers. It can detect various types of garbage such as plastic bottles and cans, garbage bags, and containers under different illumination and complex background, which has not been done before.

In addition, the enhanced1214 YOLO networks benefits from CSPDarkNet53 backbone and Hard-Swish activation increases the efficiency of feature extraction while reducing computational need. The model used also achieved a 89%Mean Average Precision (mAP) so as to suggest the model has potential to outcompete previous approaches grossing an important efficiency on performance as preferred for the function of a reliable cleaning robot. In addition, the authors tune the hyperparameter and network structure for higher preds accuracy and enhanced generalization capacity of the detector. Techniques including the DIoU-NMS and transfer learning aid with regard to the model fitting as well as the response overfitting.

In the study (Panmuang and Rodmorn, 2024), used YOLO the You Only Look Once deep learning algorithm in solving the problem of urban garbage in Bangkok. The aim of the study is to improve waste collection arrangements based on a targeted detection and recognition of the identified images of the overfilled garbage bins using CCTV cameras. Therefore, a set of images consisting of 1,383 were collected and coding categories of each of them covered garbage and bin. The performance of the goal was measured by employing 4 realistic setups of the YOLO models namely YOLOv5n, YOLOv6n, YOLOv7 and YOLOv8n for categorization of the images. That is mean the YOLOv5n has the highest localization accuracy of 94.50%, while the YOLOv8n was the second-best localization accuracy of 93.80% and three last models that were YOLOv6n and YOLOv7 were less accurate localization of objects. The authors defended that applying the model for detection with the existing CCTVs can eliminate or at least reduce waste and pollution within cities.

In the study (P Unni et al., 2024) they applied the you only look once (YOLO) deep learning algorithm in dealing with the challenge of urban garbage in Bangkok. The idea of the study is the improvement of arrangements for waste collection by recognizing the specified images of the bins filled beyond their capacities with waste through CCTV. Therefore, 1,383 images were selected and coded of which garbage and bin were included in the coding category. The success of the goal was evaluated when possibility of using 4 practical variants of the YOLO models (YOLOv5n, YOLOv6n, YOLOv7, and YOLOv8n) was applied to image classification. The outputs demonstrated that among four versions of the model, YOLOv5n had the highest efficiency with an accuracy of 94.50% for the model and YOLOv8n was second efficient with 93.80%, while YOLOv6n and YOLOv7 had the lowest efficiency among all the models. The authors defended that applying the model for detection with the existing CCTVs can eliminate or at least reduce waste and pollutions within cities.

In the study (Ren et al., 2024) proposed the MRS-YOLO model because of the raising importance of searching for relevant methods of waste treatment together with the growth of household waste generation. Prevention from over-attachment to image is improved through the inclusion of Method of Specific Small Object Detection as well as the addition of outside context. The Key Component is composed of new addition including the SlideLoss_IOU Technique, application of RepVit in the structure of Transformer, and new feature extraction including multi-dimensional and dynamic convolution mutation. The improvement covers detection accuracy rate, speed, and resilience, which have been made available in prior on the YOLO models.

3 Methodology, Design and Implementation

The real-time Automated Street Waste Detection System uses YOLOv8 and comprises five main design phases: data collecting, data pre-processing, model training, and model evaluation. The architecture deploys machine learning techniques to classify and find urban garbage for smart cities. The following parts explain the critical research technique, as seen in Figure 1

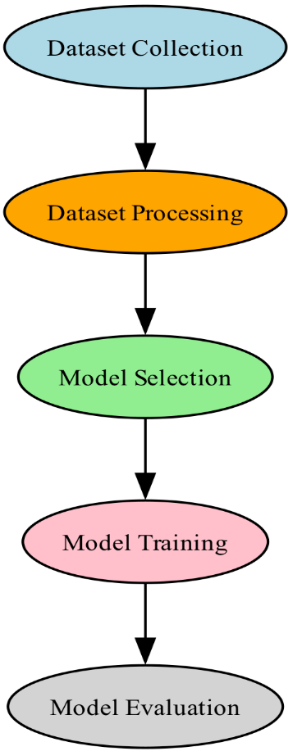

Figure 1 Methodology Design

3.1 Dataset collections

The dataset used in this study is taken from Roboflow and is named **Trash Detection Computer Vision Project** https://universe.roboflow.com/polygence-project/trashnet-a-set-of-annotated-images-of-trash-that-can-be-used-for-object-detection/dataset/20 (for Oriented Bounded Box, 2023) which is actually a combination of several images of urban wastes taken from the cams, streets, bins and dumping grounds. It is broad based because the dataset is made up of images taken at different settings, different lighting, and picture taken from different angles. These images were labelled according to kind of waste, the image committee as wastage in the forms of cardboard, glass, papers, plastics, metals and organic wastes was present in according to the label.

Balancing that into account, three different datasets – training dataset, validation dataset and test dataset – were compiled in order to have an adequate sample of the waste types. Table 1 shows that the training dataset holds 87% of the images, which are 5283 images; validation datasets have 8% of the images 499 images, while testing images consist of only 4%, which are 264 images. The goal here is to help load sufficient data samples into the model during training and picking the diversity.

Table 1: Dataset Division

Parameter	Description	Processing Applied
Augmentations	Flip, Rotate, Crop, Shear, Brightness, Noise	Generated 3 outputs per image
Total Images	6,046	Auto-Orient, Resize, Augment
Train Set	5283 images (87%)	Pre-processed and augmented
Validation Set	499 images (8%)	Pre-processed and augmented
Test Set	264 images (4%)	Pre-processed and un-augmented

3.2 Dataset Pre-processing

Data pre-processing is an essential process of data preparation for training, as well as for increasing efficiency of the YOLOv8 model. The dataset was processed using the following techniques:

3.2.1 Data Description

The dataset was containing 6 types of wastes including cardboard, glass, metal, paper, plastic, trash and total images of waste are 6049 has described above Figure 2. No classes were dropped. The input images were further scaled down to 416 x 416 pixels to correspond with the input requirements of our YOLOv8 model. Each of these examples created 3 augmented outputs.

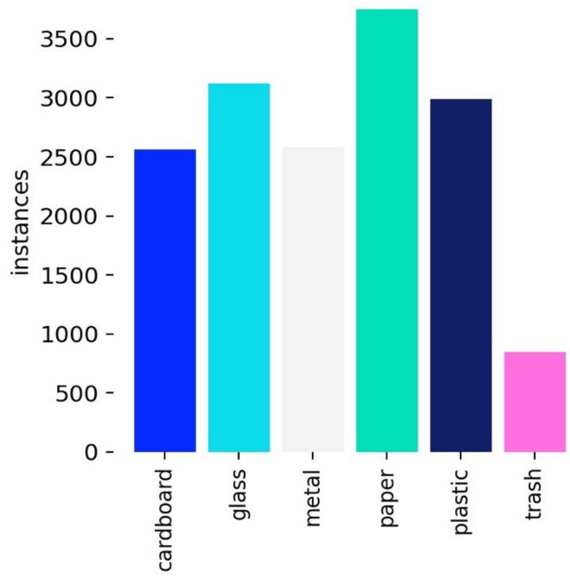

Figure 2: Total number of Images

3.2.2 Dataset Classes

The main dataset for the YOLOv8 based waste detection project focuses on the following classes distribution shown in Table 2, which are essential for urban waste categorization:

Table 2: Class Distribution

Class	Description
Plastic	Bottles, bags, and wrappers are considered to be general plastic waste.
Metal	These consists of tin canisters, crown cappers, bottle caps, and scrap iron.
Glass	Include glass bottles and jars bottles, broken pieces of glass vessels.
Paper	Material such as cardboard, normal paper, tissue and other associated paper materials.
Organic Waste	Food leftovers, garbage bags, and biodegradable materials.
Styrofoam	The cups, food carriers, and food packaging wrappings in the form of disposable mostly the Styrofoam.
Composite Waste	Cartons such as drinks carton and carded blister pack.
Electronics	Small electrical and electronic items such as batteries.

3.2.3 Pre-processing Steps

During data preparation in feature engineering, a few operations were accomplished to make the data set suitable for the model. The data was initially cleaned by auto-alignment to eliminate rotation issues produced by the camera's varied orientation angles during data gathering. Second, all photos were inscribed with a proportionate size ratio of 416 pixels in width and height as proposed by the YOLOv8 model creators. Category adjustments were created to reduce the dataset from 6 waste classes to better classes for better feature generalization after training. Normalization was employed to enhance the dataset and make the model more origin-resilient. The findings were compared on 90 rotated images (clockwise, counterclockwise, and upside down), 15° clockwise and counterclockwise normal rotations, and flipped x and y axes. Figure 3 Sample Image shows the tagged dataset photos processed for this submission.

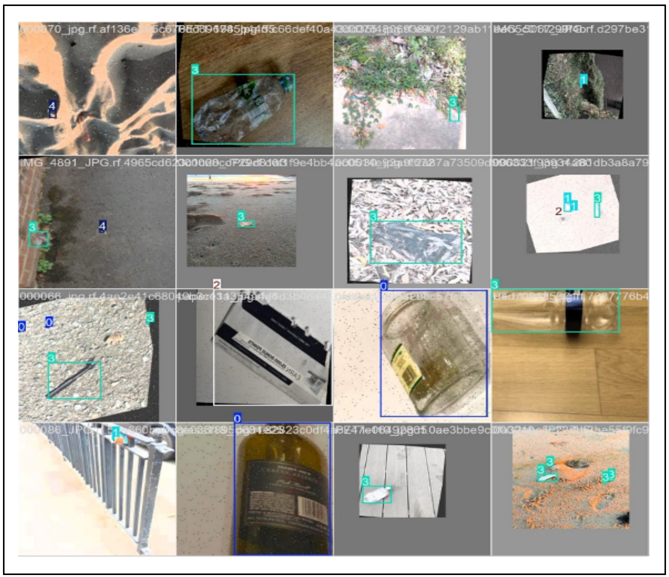

Figure 3: Sample Image

3.3 Model Architecture

Table 3, below, indicates YOLOv8 model architecture which describe some layer's output shape and some parameter. Initially, architecture is constructed with a Conv2d layer which has an output shape of (3, 608, 608) and a total of 1792 parameters. The second layer is a BatchNorm2d layer which has 128 parameters and its output shape that is equal 64*608*608. A parameter free Leaky ReLU activation layer is used afterwards, and the MaxPool2d layer is applied to reduce the spatial dimensions of the repeated blocks which are (64, 304, 304). The next is a BatchNorm2d layer that has 256 parameters, whereas, the following Conv2d layer with parameters 128, 304, 304 contains 73,856 parameters again, the block contains Leaky ReLU that followed by MaxPool2d layers and decreased the dimension to (128, 152, 152). ;zeń BatchNorm2d: 2048 ;A Conv2d : more than 2359296 ;Producing (1024, 76, 76) ;Conv2d more than 261375 parameters while Leaky ReLU is parameter-free; ;The Outputs of the Conv2d are (255, 76, 76) This summary also explains that YOLOv8 architecture is quite complicated.

Table 3: YOLOv8 Model Architecture in detail layer by layer

Layer	Output Shape	Number of Parameters
Conv2d	(3, 608, 608)	1,792
BatchNorm2d	(64, 608, 608)	128
LeakyReLU	(64, 608, 608)	0
MaxPool2d	(64, 304, 304)	0
Conv2d	(128, 304, 304)	73,856
BatchNorm2d	(128, 304, 304)	256
LeakyReLU	(128, 304, 304)	0
MaxPool2d	(128, 152, 152)	0
...	...	...
Conv2d	(1024, 76, 76)	2,359,296
BatchNorm2d	(1024, 76, 76)	2,048
LeakyReLU	(1024, 76, 76)	0
Conv2d	(255, 76, 76)	261,375

3.4 Model Selection

Two different versions of YOLO v8 were selected for the Automated Street Waste Detection system. These models achieve a reasonable compromise for accuracy and speed and therefore allow real-time implementation These models are YOLOv8l and YOLOv8s:

- **YOLOv8l (Large):** This version is more accurate and hence is more reliable. It has a bigger architecture which enables this variant to capture low-level essential features as well as high-order highly complex patterns present in the distribution data. It must be noted though that this variant is less efficient when high volumes of computation are maintain and also takes more time when performing operations. It captures dense or highly complex scenes effectively and thus is applicable in a diverse range of tasks including waste detection and classification in busy environments.

- **YOLOv8s (Small):** This type of model is characterized by the large structure as opposed to YOLOv8s this is the reason that YOLOv8s model is able to detect much more ranges of images considering the fact that large number of resources for computing are available. This model is particularly useful for image detection and recognition in cases where partial obstruction of the image occurs or the image has multiple waste objects that need to be scrutinized. Waste typologies are also discernible more effectively with YOLOv8s in general.

3.5 Model Training

The models were first trained on a dataset which had several labelled images where pre-learned weights were introduced so that the model could perform well and to achieve improved convergence and bearing in mind that the used dataset is not large. Therefore, in this case, a more plausible strategy was transferring learning where initial layers of the YOLOv8 models were frozen to retain the features learnt in the pre-trained model and the later layers were untrained for waste detection

- **Batch Size and Epochs** A batch size of 64 was set and the model was trained for 100 epochs. This duration allowed a good learning rate to be achieved without reaching overfitting monitored by performance on the validation set.
- **Loss Function:** The loss function that was applied in the course of the training was made of three portions, among them: classification loss – for the purpose of classification of the object the localization loss of bounding box prediction and the objectless loss of confidence scores of the designed object. These elements are trainable during the learning process in order to err less in classification as well as localization.
- **Optimizer and Learning Rate:** The use of Adam optimizer is well suited to this task because it is capable of learning rates adaptive to the situation. The learning rate also began with an Initial learning rate of 0.001. In this case, a way of decaying the learning rate with relation to the training epochs was adopted to enhance convergence.

3.6 Model Evaluation

After both models were developed, the experiment was performed on the test set which will determine the effectiveness of the model. Metrices such as precision, recall and the mean average precision were used as a guide in estimating the precision of waste detection. These metrics allow one compute over the model for assessing waste object detection capability such as correct waste positioning in images as well as the overall detection of all objects. capability of the model in relation to these other measures of detection.

- **Mean Average Precision (mAP):** Evaluates the model's precision across a range of Intersection over Union (IoU) thresholds, offering a comprehensive measure of detection performance. mAP provides a balanced assessment of the model's ability to detect waste objects accurately and consistently across different scenarios.
- **Precision:** Represents the proportion of correctly identified waste objects (true positives) to the total number of objects predicted by the model. It reflects the accuracy of the model's positive detections.

- **Recall:** Indicates the ratio of correctly identified waste objects (true positives) to the total number of actual waste objects in the dataset. It measures the model's ability to detect all relevant objects.

4 Results and Discussion

4.1 YOLOv8 Small Model (YOLOv8s) Performance Analysis

4.1.1 F1 Score

An advantage of the two F1-Confidence Curves as seen in Figures 4 and 5 is that from the graphs, it is clear that both the models have relatively better general performance compared with other models; the F1 score remains relatively high on the unseen classes such as the metal and the cardboard. Conversely however, the trash class remains a challenge to both the models, registering high fluctuations in the F1 scores at different levels of confidence switches. The YOLOv8s has the highest Average Precision for threshold value of 0.50. On the other hand, the YOLOv8l is most accurate at a 0.80 threshold to give the best performance. This may mean the first model is desirable when recall is important than precision and the second model where precision is important than recall. However, for an optimal model, a high recall and precision are not going to decide best model option since its application's needs are best served by a mix of recall or precision.

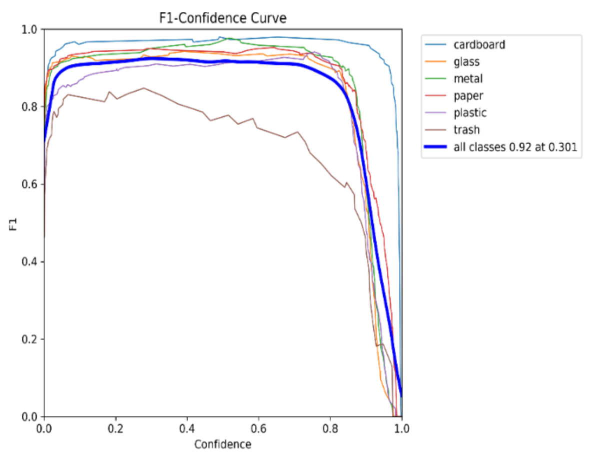

Figure 4: F1 Score of YOLOv8small

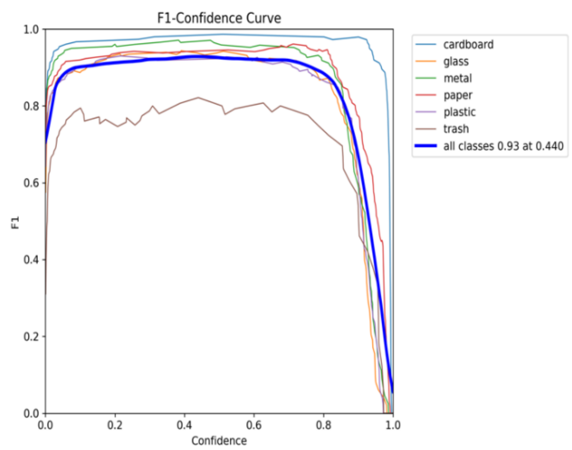

Figure 5: F1 Score of YOLOv8large

4.1.2 Precision

As shown in Figure 6, Precision of YOLOv8s, and Figure 7, Precision of YOLOv8l, stepped equal improvements in training's precision are illustrated. Firstly, the training of YOLOv8s could be described as rather pronounced at the early stages, with the maximum level of precision determined with the help of estimations that amounted to approx 0.88 and variable depending on the training process, On the other hand, the YOLOv8l model demonstrated significantly lower rates of growth, but with gradually increasing tendencies. At last, the increase in YOLOv8l was higher than the maximal of YOLOv8s in terms of performance. First, such performance proved that YOLOv8l had been developed to a greater extent; second, they actually presupposed a larger capacity of the model as well as a stronger factor for generalization. At the same time, however, it must be noted that with the help of YOLOv8s, the model has a fast-learning ability which is necessary in the process. However, recall and F1-score are also beneficial metrics to support the precision to analyze the model performance.

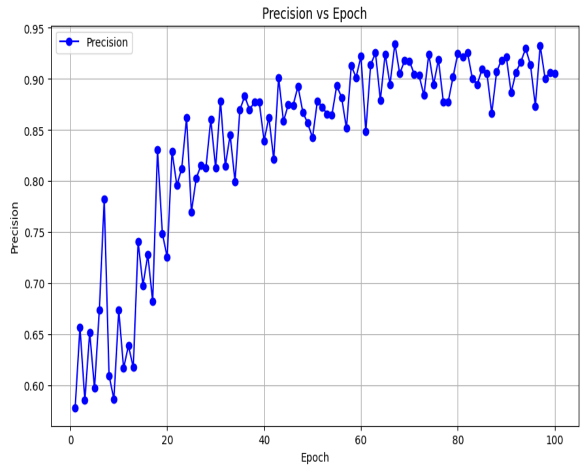

Figure 6: Precision for YOLOv8small

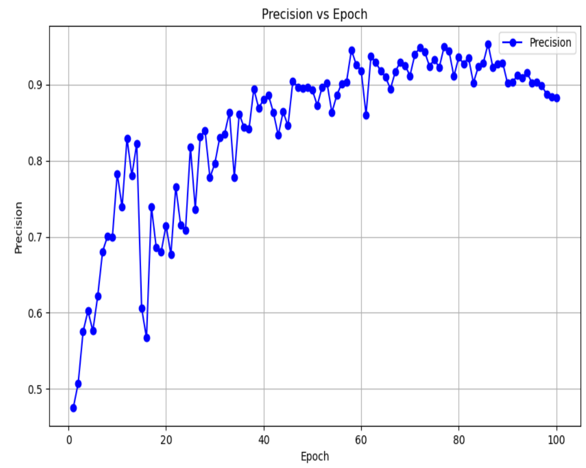

Figure 7: Precision for YOLOv8large

4.1.3 Recall

The overall recall for both YOLOv8s and YOLOv8l models in figure 8 and figure 9 improved gradually during the period of training. On the other hand, the recall increasing with, for example the initial peak of around YOLOv8s was slightly lower at 0.87 with steeper ascending rate and faster oscillations compared to the slower growing rate for YOLOv8l that asymptote around a slightly lower high recall rate of 0.86-0.87. From the fluctuations in figures obtained for peak recall and moderate oscillation in YOLOv8l it may be deduced that this particular model is better tuned for the work at hand and might well have a larger capacity and better generalization. However, one must not forget that learning capability of YOLOv8s also possesses a fast turnover time. It means that these models have different context depending on the particular application and the exact need for more precise model when it is necessary or more strict recall when it is urgent. For example, in a situation where the objective is to accurately enumerate each of the exposed objects at the specific of false positive objects such as in most of the YOLOv8s, then such situation is ideal for the use of the mentioned models. However, if such a situation occurs that a false positive is worse than low positive detection, then YOLOv8l may be used as it demonstrated equal detection rates with minor variabilities

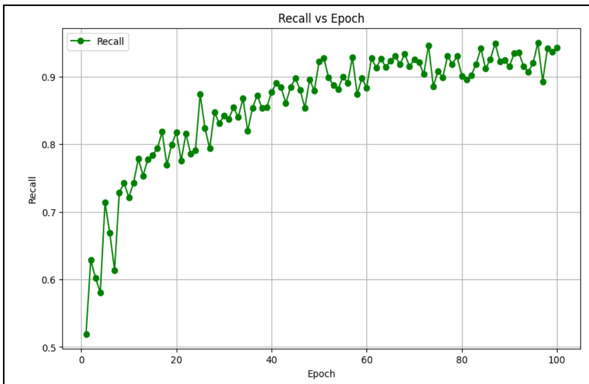

Figure 8: Recall for YOLOv8small

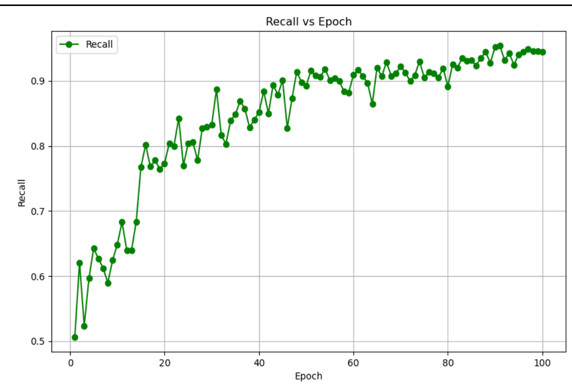

Figure 9: Recall for YOLOv8large

4.1.4 mAP@50

Interestingly, over the course of training, the YOLOv8s remain very sensitive to the initial mAP@50 score, which is normally considered to be an aspect of a model's accuracy and recall space in the detection tasks concerning the object's accuracy but during epoch 1 the mAP @ 50 is equal only to 0.6003 as per, Figure 10. The improvement over epochs brings a high 0.92957 at epoch 36 being 36 times better, performance is then tested at new epochs, in this case it would be this epoch. The increase in progress can be seen in the above Figure 10 which depicts the general evolution of the model improvement over several time intervals. The enhancement can be seen from 10th epoch which indicate the model is capable of handling prior forecast errors up to when the training gets to this level. The mentioned peak mAP@50 consistency can be described as practical use of the YOLOv8s model, where it is possible to track in real time objects that do not require high localization accuracy.
The YOLOv8l also continued to perform a steady increase of its expectations mAP@50 score that did not even get below 0.91 by epoch 36 as shown in the Figure 11. This is significant. The mAP@50 of this model also indicates that it is plausible as a detection model task that can be challenging like tracking of precision and recall. In the advanced specification of the above study, a linear mAP is indicated as well. This agreeing with this, there still existed areas that needed to be improved according to the current case which would be enduring unit needs.

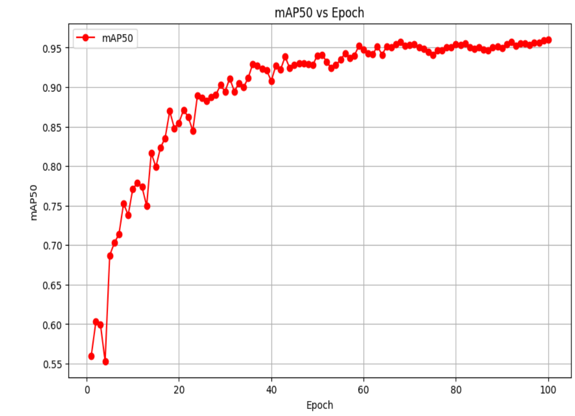

Figure 10: mAP@50 YOLOv8small

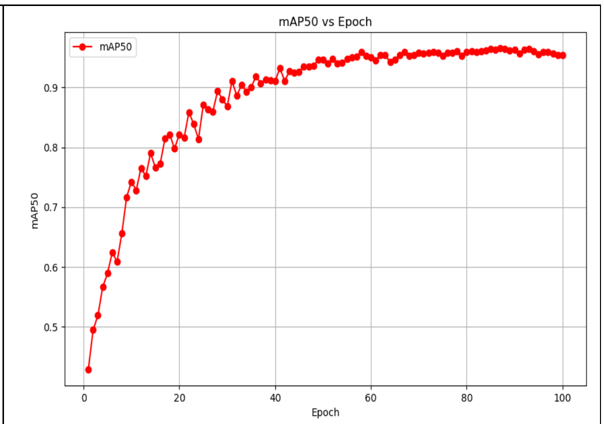

Figure 11: mAP@50 YOLOv8large

4.1.5 Confusion Matrix

With regards to the object classes, the YOLOv8s model had a high accuracy of detecting the most common classes of objects like, paper with 111/124, plastic with 90/103 or glass with 88/101 but struggled to determine which items were trash as can be seen in Figure 12. Therefore, the model achieved 100% true positive rate, in the sense that only 24 of 31 true trash instances were correctly classified. This means that the model may produce lower accuracy when it comes to discriminating trash from other categories that are less distinct, or visually associated; this has made this class to have low accuracy. However, the performance of the YOLOv8l model with regard to many categories was high: Cardboard (73/78), glass (90/94), metal (81/94), and paper (113/124) were also successfully detected as depicted in Figures 13. Despite being fairly robust in other tasks, this model struggled to sort the trash from the recyclable and did so with plastic and glass in particular. Among the 31 actual trash instances, the model recognized only 23 of them many of which it confused with plastic or paper bags. That's why the model works well with many other object classes, but as for the discrimination in terms of trash materials, it can be improved. In sum, the two models in question in the present study had a high level of expectation but at the same time to enable improvement, they also had certain drawbacks which can be worked on.

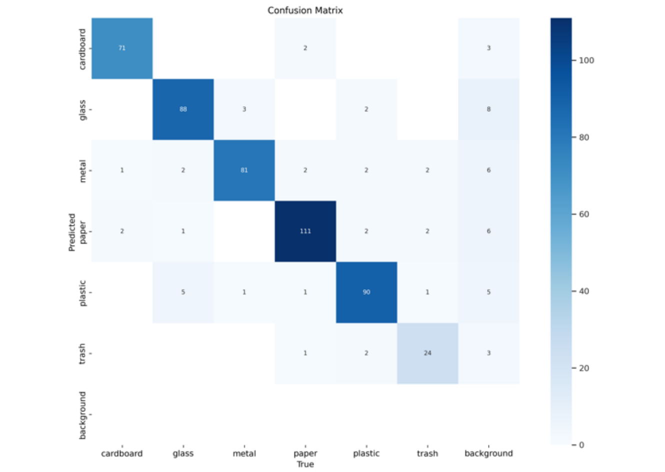

Figure 12: Confusion Matrix of YOLOv8small

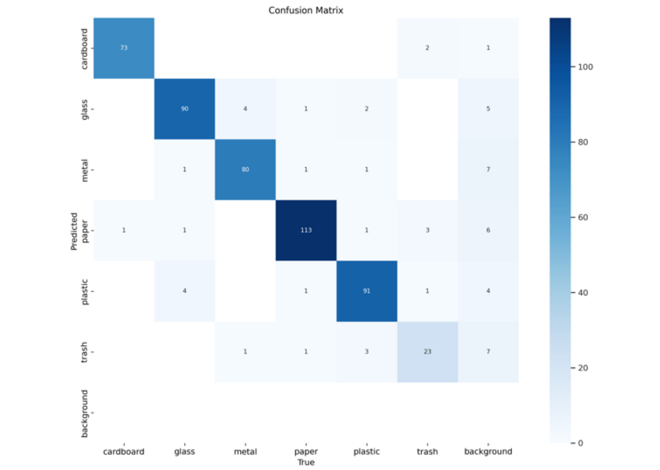

Figure 13: Confusion Matrix for YOLOv8large

4.1.6 Predicted result for YOLOv8small and YOLOv8large

From the results, we get YOLOv8s which is illustrated in Figure 14 well balanced and faster but with more weight and may sometimes fail to detect small objects, thus we can see that there is a compromise between speed and accuracy. YOLOv8l is more accurate in detection than the other modes, detecting all objects as seen in the figure below in Figure 15 due to the large network that the model encompasses and the computational power it holds. On the other hand, it also means if we have to choose a model size correctly, we must take into consideration our application needs for instance if we need a high real-time detection.

Figure 14: Prediction YOLOv8small

Figure 15: Prediction of YOLOv8large

Table 4 below is a comparison of performances of YOLOv8s and YOLOv8l in several parameters The efficiency parameters are as follows: Precision and recall values for YOLOv8l models are higher than for its counterpart equal to 0.93 and 0.92 respectively, whereas the f1 score is higher for YOLOv8l model. However, YOLOv8s yielded fewer false positives because it obtained a higher precision score than the rival methods with the score of 0.90. On recall both models were again equal with 0.94 meaning that both models have the same capacity of identifying occurrences of true positives. Anyway, what was even more amusing that in mAP@50 YOLOv8s scored higher 0.96) than YOLOv8l (0.94 which means overall better detection with the cutoff of 50% IOU. These results have demonstrated the relationship between precision and recall as well as between accuracy and precision.

Table 4: Performance Metrics

Performance Metric	YOLOv8s	YOLOv8l
F1 Score	**0.92**	**0.93**
Precision	**0.90**	**0.88**
Recall	**0.94**	**0.94**
mAP@50	**0.96**	**0.94**

4.2 Discussion

There is no qualitative difference between a YOLOv8s and a YOLOv8l, the only difference lies in the design parameters of the application. Similar to the previous models, YOLOv8s is optimized for the resource scarcest environment, it provides a good speed/accuracy trade off and is well suitable for real-time deployment on edge devices. specs, and the other is the accuracy and the precision, YOLOv8l is more accurate and precise than the YOLOv8s due to its high computational complexity in complex and densely crowded scenes. As put in the automated street waste detection system that we now implement, it is seen that the real time detection output are rather low. We hope to address this in the next phase where we hope to fine the model by tuning this to run more optimally, in terms of speed but no less accurate when consolidated to give good performance in both small scale and complex computation cases.

5 Conclusion

The study demonstrated the capabilities of YOLOv8 family models, particularly YOLOv8s and YOLOv8l, in enhancing systems aimed at automating street waste detection. Both models provided satisfactory results with each offering different benefits that could be useful for varying application requirements. Due to its smaller size and its minimally modified feature set, YOLOv8s facilitates the performance of real time operations in the constrained environments. Metrics like 0.90583 in precision, 0.94279 in recall, and 0.96042 in MAP@50 made YOLOv8s sufficiently reliable to manufacture for edge devices due to its accuracy and speed features. However, it may struggle to discriminate when target objects are small or overlap each other within complex environments. In contrast, YOLOv8l performs well when there is a need for high precision in detection, for example when the model is tasked with expanding its application in areas with smaller or overlapping objects such as busy urban districts. Its design facilitates continuous enhancement of the performance even as training phases shift and is also quite promising for resource-scarce areas. While requiring a considerable number of resources to function, YOLOv8l is very effective for uses where speed is not an issue. In both models, a poor ability to handle trash categorization as a relatively lesser distinct topic is said to be a problem. In order to rectify this drawback, model fine tuning, employing diverse waste images to augment data, and developing better classification methods can be pursued. As for the particularities of urban waste management cases, preference of YOLOv8s or YOLOv8l will depend on requirements of the deployment scenario, such as speed or precision) Integration of smart technologies into the wider urban waste management strategy is the goal as it will lead to more efficient solutions.

5.1 Future Work

The research provides a good base focusing on applying YOLOv8 models for waste monitoring but there are several more areas that need to be addressed in order to fully utilize them and improve on some of the limitations the models faced:

- **Model Optimization for Specific Scenarios**
  Although speed and performance of models are significantly improved across all applications, there is a need to further customize the models towards specific areas in urban waste management. This could be coming up with hybrid approaches that integrate faster version of the models together with the stronger version ordering to enable balanced performance across various applications. NAS and KD techniques are likely to play a central role in this task
- **Enhanced Data Augmentation and Dataset Expansion**
  It is well-established that the performance of an object detection model is directly related to its training data in terms of pool and diversity. It is most evitable that the construction of the dataset should ensure a broadened capture of waste images under different climates, seasons and regions. Further, use of advanced techniques for data augmentation like, synthetic data generation and adversarial training could help the models in gaining robustness towards lighting, occlusion and object overlap
- **Improved Classification for Ambiguous Categories**
  Each of the models still suffers from problems of distinguishing less easily differentiable categories for instance while labeling "trash." Further work should look into either adding more division on top of the categorization or using additional networks which mitigate such conflicts. It may also be useful to apply CATMA for greater recognition of different but vaguely the same categories of garbage classification – LINDO, or add technologies of semantic segmentation
- **Integration with IoT and Smart City Infrastructure**
  For effective usage of YOLOv8, particularly in urban waste management, it has to be coupled to edge IoT devices and urban frameworks. Such future systems may see the installation of YOL08s on inefficient IoT devices to enable a frequent monitoring of trash and usage of YOL08l to enhance trash monitoring accuracy. Such development may allow changes in the model's operational techniques on waste collection and resource distribution to be made on the fly
- **Energy Efficiency and Green AI**
  As computational needs become more sophisticated, energy efficiency measures during training and deployment of models become paramount. Further research should be directed towards minimizing energy costs incurred by the puncturing vines of YOLOv8 models through performing quantization and pruning techniques, and maintaining target accuracy as well
- **Long-Term Scalability and Deployment**
  Lastly, further research should investigate these systems' long-term convenience with respect to deployment in expansive urban areas. This also entails looking into issues of model upkeep, retraining of the model with new information, and actual usage of the model in the different urban landscapes

  With the above-mentioned measures, it will become easier to target the further enhancement of YOLOv8 models, enabling the creation of more efficient and environmentally friendly solutions for waste management in urban areas, which are supported by international standards

References

1. Erin, K., Bingöl, B. and Boru, B. (2022) 'YOLO – Based Waste Detection', *Journal of Smart Systems Research*, 3(2), pp. 120–127.

2. Sharma, P., Srinivasan, K., Azizan, M. M., Hasikin, K., Salwa, A., & Khairuddin, M. (2022). An automated solid waste detection using the optimized YOLO model for riverine management.

3. Panmuang, M., & Rodmorn, C. (2024). Garbage Detection using YOLO Algorithm for Urban Management in Bangkok. *WSEAS Transactions on Computer Research, 12, 236–243*. https://doi.org/10.37394/232018.2024.12.23

4. Wahyutama, A. B., & Hwang, M. (2022). YOLO-Based Object Detection for Separate Collection of Recyclables and Capacity Monitoring of Trash Bins. *Electronics (Switzerland), 11(9)*. https://doi.org/10.3390/electronics11091323

5. Sun, Q., Zhang, X., Li, Y., & Wang, J. (2023). YOLOv5-OCDS: An Improved Garbage Detection Model Based on YOLOv5. *Electronics (Switzerland),* 12(16). https://doi.org/10.3390/electronics12163403

6. Suprapto, B. Y., Kelvin, Kurniawan, M. A., Ardela, M. K., Hikmarika, H., Husin, Z., & Dwijayanti, S. (2021). Identification of garbage in the river based on the YOLO algorithm. *International Journal of Electronics and Telecommunications*, 67(4), 727–733.

7. Unni, A. P., A, A. N., Krishna A, S. N., Livera, S., & Professor, A. (2024). *Waste Segregation using YOLO v8 based Object Detection and Robotics* (Vol. 11). www.jetir.org

8. Ren, Y., Li, Y., Gao, X., & Gao, · Xinya. (2024). MRS-YOLO: A High-Precision Model for EEcient Waste Detection and Classiication MRS-YOLO: *A High-Precision Model for Efficient Waste Detection and Classification*. https://doi.org/10.21203/rs.3.rs-4485704/v1

9. Cook, E., Silva de Souza Lima Cano, N., & Velis, C. A. (2024). Informal recycling sector contribution to plastic pollution mitigation: A systematic scoping review and quantitative analysis of prevalence and productivity. In Resources, Conservation and Recycling (Vol. 206). Elsevier B.V. https://doi.org/10.1016/j.resconrec.2024.107588

10. Erin, K., Bingöl, B., & Boru, B. (2022). YOLO – Based Waste Detection. In Research Article Journal of Smart Systems Research (Vol. 3, Issue 2). http://biyak.subu.edu.tr

11. Kumar, S., Smith, S. R., Fowler, G., Velis, C., Kumar, S. J., Arya, S., Rena, Kumar, R., & Cheeseman, C. (2017). Challenges and opportunities associated with waste management in India. In Royal Society Open Science (Vol. 4, Issue 3). Royal Society. https://doi.org/10.1098/rsos.160764

12. Mishra, A., Ghosh, N., & Jena, P. (2019). Internet of Things based Waste Management System for Smart Cities: A real time route optimization for waste collection vehicles. INTERNATIONAL JOURNAL OF COMPUTER SCIENCES AND ENGINEERING, 7, 541–548. https://doi.org/10.26438/ijcse/v7i4.541548

13. Mitra, A., & Li, Y. (2020). Detection of Waste Materials Using Deep Learning and Image Processing.

14. Tamakloe, C.-N., & Rosca, E. (2020). Smart Systems and the Internet of Things (IOT) For Waste Management. 1–6. https://doi.org/10.1109/CIVEMSA48639.2020.9132968

15. Wang, H., Wang, L., Chen, H., Li, X., Zhang, X., & Zhou, Y. (2024). Waste-YOLO: towards high accuracy real-time abnormal waste detection in waste-to-energy power plant for production safety. Measurement Science and Technology, 35(1). https://doi.org/10.1088/1361-6501/ad042a