



Research Article

Deep learning-based object detection for smart solid waste management system

Meron Desta^{1*}, Tagel Aboneh² and Bisrat Derebssa³

¹Addis Ababa Science and Technology University, Addis Ababa, Ethiopia

²Addis Ababa Science and Technology University, High- Performance Computing and Big Data Analytics, Addis Ababa Science and Technology University, Addis Ababa, Ethiopia

³School of Electrical and Computer Engineering, Addis Ababa Institute of Technology, Addis Ababa University, Addis Ababa, Ethiopia

Received: 07 September, 2022

Accepted: 30 August, 2023

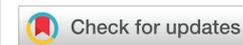
Published: 31 August, 2023

***Corresponding author:** Meron Desta, Addis Ababa Science and Technology University, Addis Ababa, Ethiopia, E-mail: merond308@gmail.com

Keywords: Deep Learning, Object Detection, YOLOv4, YOLOv4-tiny, Waste

Copyright License: © 2023 Desta M, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

<https://www.peertechzpublications.org>



Abstract

Currently in Ethiopia, pollution and environmental damage brought on by waste increased along with industrialization, urbanization, and global population levels. Waste sorting, which is still done improperly from the household level to the final disposal site, is a prevalent issue. Real-time and accurate waste detection in image and video data is a crucial and difficult task in the intelligent waste management system. Accurately locating and classifying these wastes is challenging, particularly when there are various types of waste present. So, a single-stage YOLOv4-waste deep neural network model is proposed. In this study, a deep learning algorithm for object detection using YOLOv4 and YOLOv4-tiny is trained and evaluated. A total of 3529 waste images are divided into 7 classes, which include, cardboard, glass, metal, organic, paper, plastic, and trash. Each model uses three various inputs throughout the testing phase, including input images, videos, and webcams. Experiments with hyper-parameters on subdivision values and mosaic data augmentation were also done in the YOLOv4-tiny model. The outcome demonstrates that YOLOv4 performs better than YOLOv4-tiny for object detection specifically for waste detection. The outcome shows that YOLOv4 performs better than YOLOv4-tiny for object detection, even if YOLOv4-tiny's scores are higher in terms of computing speed. The best results from the YOLOv4 model reach mAP 91.25%, precision 0.91, recall 0.88, F1-score 0.89, and Average IoU 81.55%, while the best YOLOv4-tiny results are mAP 82.02%, precision 0.75, recall 0.76, F1-score 0.75, and Average IoU 63.59%. This research also proves that the models with smaller subdivision values and using a mosaic have optimal performance.

Introduction

Solid waste is defined as any type of garbage, trash, refuse, or discarded material. It can be classified according to where the waste is generated, such as municipal solid waste, medical waste, and e-waste [1].

Approximately 62 million tons per year are produced in sub-Saharan Africa [2]. Even while production is rising, garbage collection rates in developing countries are usually less than 70%. Over 50% of the waste that is collected is frequently dumped in uncontrolled landfills, and only 15% is recycled safely and ethically. In African nations, household waste accounts for the majority. The amount of garbage generated per person each day in Addis Ababa is estimated to be between 0.4 and 1.23 lit, 0.11 to 0.25 kilograms, and 205 to 370 kg/m³ [3]. Waste generation is rising despite the city's inadequate solid waste collection and disposal system [4].

Both the Federal Democratic Republic of Ethiopia (FDRE) and the Oromia National Regional State Government have their headquarters in Addis Ababa. Addis Ababa serves as both a diplomatic hub and the headquarters for numerous organizations. Addis Ababa spans a land area of 540 km² and is geographically situated between latitudes 8055 and 9005 in the north and 38040 and 38050 in the east. With a population of roughly 3.5 million and a growth rate of 8% per year, there are 99 Kebeles and 10 sub-cities (Kifle Ketema) with a density of 5936.2/km² [2].

Habitually, without any segregation, the majority of the solid trash is treated by Traditional garbage disposal techniques like burial or landfill, incineration, chemical corrosion, etc. which will significantly pollute the soil and air while also being quite expensive. Landfilling is one of the commonly used waste disposal management. When doing that, the most serious

worry is plastic garbage, which is the most common and causes the most long-term environmental harm [5]. Depending on the substance and structure, plastics can decompose in 20 to 500 years. The fact that organic waste decomposes anaerobically in landfills, producing methane rather than using up resources, is another issue with this method of disposal. Methane produces a stronger greenhouse gas effect than carbon dioxide when it is released into the atmosphere. When exposed to oxygen, it frequently starts an uncontrolled fire in the landfill. However, if managed differently from other waste types, organic waste could be turned into a renewable energy source. Anaerobic digestion of organic waste results in the production of biogas, a fuel that is high in methane. By replacing fossil fuels with renewable sources of energy like methane, greenhouse gas emissions can be reduced and global warming can be slowed.

Only 65% of the garbage produced in Addis Ababa is collected and disposed of, 5% is recycled, 5% is composted, and 25% is not collected and deposited in places that are not permitted [6]. 71% of solid waste is generated by households, with the remaining 26% coming from businesses, which are then divided as follows: hotels (3%), hospitals (1%), commercial centers (9%), and street cleaning (10%). It is inferred that tight cooperation between the government and the household is required to manage solid waste appropriately and effectively from its source. The generated municipal solid trash is delivered to Koshe (Reppi), an unmanaged landfill that is currently situated in the city's core and provides a serious health risk to the surrounding neighborhoods [3].

The majority of the solid waste that is collected in Addis Ababa is frequently discarded in the Reppi open dumping site without any sorting. Open dumping, which in this context refers to the unplanned disposal of waste without the involvement of environmental protection measures, is, in accordance with van Niekerk and Wegmann [7] by far it is the most common practice in Africa. As a result, this disposal approach has a negative effect on both the community and the ecology. Many African cities only have one official landfill site, which is frequently overflowing and poses a major threat to public health and safety [8]. In a similar vein,

Addis Ababa's only dumping site since 1964 has been Reppi

In the waste dump, there is no segregation, which encourages a lot of salvagers (waste pickers) to enter and search the area for recyclable and reusable items. The landfill has also contributed to a number of societal concerns, including odor and environmental challenges. Cholera, typhoid, and amoebic infections, which make up nearly half of all illnesses in the country, are more prevalent due to the current lack of cooperation in waste collection and disposal. Reproductive, dermatological, and visual problems are among other detrimental effects on health. Significant health risks include dermatitis, noise pollution, diarrhea, and, most importantly, the prevalence of children under 10 playing with condoms and other abandoned medical equipment such as syringes and needles [9] (Figure 1).



Figure 1: Waste pickers picking waste from general waste [9].

Existing system used to manage waste in Ethiopia - formal waste management process

There are 10 sub-cities in Addis Ababa (the smallest administrative unit in FDRE). Additionally, each *Kebele* may have 7500 – 8500 families. The existing waste management procedure has two components: formal and informal. The procedure's formal waste management comprises two stages of collection and disposal at a dumpsite, both of which are done entirely by government workers. The informal segments feature a large cast of actors. Efforts were made on one's initiative to collect various wastes and sell them at a spot called "Menallesh Tera" in the largest open market in the nation, known as Merkato. Individuals and other industrial players visit *Menallesh Tera* to obtain the supplies they need [10].

Containers are positioned in common locations near the main roadways in each *Kebele*. For various families, the distance to these bins could be varied. Some people may live just nearby, while others may live one or more kilometers away. Employees use trolleys to transport sacks of garbage to the containers according to schedules from the *Kebele*. At this moment, the collection process is at its core. People must take their rubbish to the common area on their own if they live far from the waste collection containers or if they are unable to pay the costs. The containers are yellow/green in hue and 8 m³ in capacity. The government vehicles then return to Reppi after emptying the containers [3].

The garbage is eventually discharged at Reppi. It was believed to be far enough away when Reppi was formed in 1964 to not be a concern, but due to the city's quick expansion, towns have already been constructed all around the dump site. Leachate or gas cannot be effectively collected at the landfill (Figure 2).

The exact boundaries of the dumpsite are unknown because there isn't even an appropriate fence surrounding it, but the amount of garbage there occupies 25 acres. The garbage is not covered by topsoil or anything on top of it. It experiences two months of midsummer heavy rain in addition to bright sunlight throughout the year. At the scene, tens of thousands of vultures and scavengers are at work. It's impossible to determine how much harm has been done because there hasn't been a reliable record of the rubbish thrown onsite and no system in place to collect leachate or emissions (Figure 3)[6].



Figure 2: Waste collection types (a) Primary collection and (b) accumulation from households and small commercial centers [9].

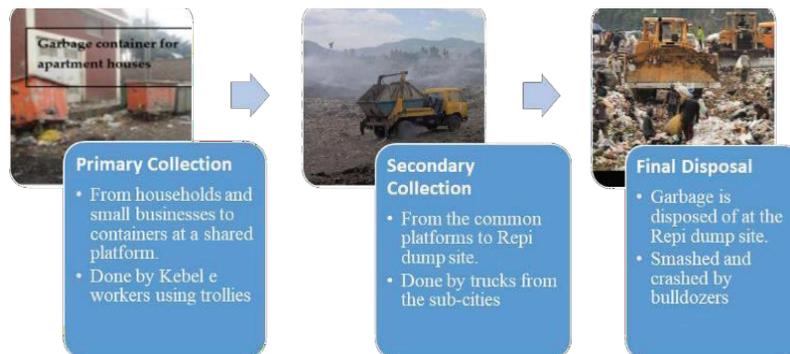


Figure 3: Formal waste management process.

At a great distance, the dumpsite's revolting odor calls for observing it is also pretty unpleasant. The dumpsite poses a substantial threat to the health and life of many people who reside nearby and is seriously harming the environment.

Challenges in waste collection, transportation, and disposal stages

Since the wastes are not divided into their components, every stage of the waste management process involves significant difficulties and, in general, health concerns for the participants during the primary and secondary waste collection stages. Health issues can affect the eyes, skin, or lungs, abandoned medical supplies like syringes and needles by children are another severe health risk, and other health risks like respiratory issues, dermatitis, and vision issues are among the risks that are experienced by waste collectors. Because there is no segregation in the waste dump, a lot of salvages (waste pickers) including children come in and look about for recyclable and reusable items.

The landfill has also contributed to a number of societal problems, including odor and environmental challenges. There is no efficient method for gathering leachate or gas at the dump. There isn't even an appropriate fence surrounding the dumpsite, so it's impossible to determine its exact boundaries. The garbage is not buried beneath anything or covered by dirt. It receives two months of intense midsummer rain in addition to bright sunlight throughout the entire year. At the scene, there are countless vultures and scavengers at work who will be exposed to series health problems [6].

Related works

Even if computer vision-based trash segregation hasn't been used in our country, there have been many attempts at it worldwide. However, each of these efforts has its setbacks with regard to how well it can execute the task. Most of them use two-stage detectors which are bulky to be deployed/used in IOT and mobile devices. These detectors require more inference time than single-stage detectors. And almost all of them are modeled in which they detect a single waste at a time.

There is currently no automatic waste segregation system at the residential level in Ethiopia, making the creation of a practical, affordable, and eco-friendly classification model for urban households urgent.

The effectiveness of computer processing of images has significantly increased as a result of the significant increase in computer operating speed. CNN (Convolutional Neural Network) based deep learning models have started to take center stage in the area of image recognition and classification. The process of separating waste into its many components is one of the most crucial parts of waste management, and it is typically carried out manually by hand-picking.

The process of separating the waste into its many components is one of the most crucial parts of waste management, and it is typically carried out manually by hand-picking.

So, with the help of computer vision, we can make the process efficient and resilient through image segmentation

and classification as waste segregation become a significant concern in our lives. These systems' increasing demand for accurate and effective segmentation and recognition methods ties up with modern computer architectures' increasing processing power and improved image recognition algorithms.

An intelligent garbage classifier; that analyzes images from the camera, the robot arm, and the conveyor belt for visual classification is used. It employed watershed to separate an overlapping waste and K-NN for classification, with the shape being the most significant characteristic they took into consideration. Nevertheless, they omitted to mention the classifier's accuracy. Since the same class of garbage might come in a variety of sizes and shapes, using merely shape alone to identify objects is insufficient.

Mittal, et al. [11], used a Convolution Neural Network (CNN) which is a machine learning algorithm, was utilized as the model in this study and was applied to a dataset of images of trash. This study classifies diverse waste images into the appropriate categories and continues to provide training accuracy and test accuracy at 91% and 81% respectively [11].

The output of medical waste is outpacing the demand for health in a progressive way as demand rises. Gyawali et al. [12], made a Comparative Analysis of Multiple Deep CNN Models for Waste. They suggest a deep learning method for classifying and identifying medical waste. They suggest a deep learning-based classification strategy [12]. In this case, with ResNeXt serving as a suitable deep neural network for actual implementation, and then they suggest transfer learning techniques to enhance the classification outcomes. Using the method on 3480 photos, they were able to identify 8 different types of medical waste with 97.2 percent accuracy; the average F1-score of five-fold cross-validation was 97.2 percent. This study offered a deep learning-based technique for high accuracy and average precision automatic detection and classification of 8 types of medical waste [13]. An automated system based on a deep learning approach and conventional techniques by aims for the accurate separation of waste into recycling categories in order to reduce the damage caused by improper garbage disposal, more specifically residential. Glass, metal, paper, and plastic were among the four garbage categories taken into consideration. They get an accuracy of 80% using SVM and 88% when using KNN. Results indicate that the computational cost of CNN algorithms is typically higher than that of conventional techniques, necessitating more powerful computing facilities [14].

An image processing-based intelligent garbage sorting system that is hardware and software was integrated to classify data with an overall accuracy of 83.38% on the issue of solid waste separation systems using the SURF-BOW feature extraction technique and multiclass SVM. The difficulty with the classical approach is having to choose which components of a given image are essential. As there are more classes to categorize, feature extraction becomes more challenging. For each feature definition, the CV engineer must also carefully modify a huge number of parameters. The engineer's judgment, based on much trial and error, must be used to identify which attributes best define different classifications of objects [15].

A transfer learning-based DenseNet169 waste image classification model to increase the speed and precision of waste categorization processing was also utilized. DenseNet169 model that is appropriate for their experimental dataset based on the deep learning network DenseNet169's pre-trained model. According to the experimental findings, the DenseNet169 model's classification accuracy after transfer learning is above 82%, which is higher than that of conventional image classification algorithms. But the DenseNet169 suffers from duplicated gradient flow throughout the layers which adversely affects the accuracy of the model. And their accuracy can be modified using different techniques [16].

A significant, renewable source of energy is municipal solid trash. For image categorization, convolutional neural networks are employed. These wastes are divided into many divisions using equipment constructed in the shape of a trashcan. The study would introduce automation in the field of waste management and save valuable time if such waste materials weren't separated by humans. The ResNet18 Network was used, and the best validation accuracy was discovered to be 87.8%. An important constituent of household waste classes is not considered in their target classes and the selected model network performs the detection process in two stages it's not applicable for real-time detection [13].

A transfer learning-based DenseNet169 waste image classification model to increase the speed and precision of waste categorization processing was also utilized. They were able to create a DenseNet169 model that is appropriate for their experimental dataset based on the deep learning network DenseNet169's pre-trained model. According to the experimental findings, the DenseNet169 model's classification accuracy after transfer learning is above 82%, which is higher than that of conventional image classification algorithms. But the DenseNet169 suffers from duplicated gradient flow throughout the layers which adversely affects the accuracy of the model. Their accuracy can be modified using different techniques [11].

Materials and methods

This work makes use of a variety of software. For deep learning-based waste object detection, Python programming language of version 3.8 with Anaconda IDE Jupyter notebook, TensorFlow library V2.1.2, and Open-cv module is used. Labeling is done by using the LABELIMG tool. Which is `ImgAnnotationLab_V4.1.0.0`. That is a free, open-source tool that can graphically label images. Training is done using Google COLAB which is a web-based Python editor that allows anyone to write and run arbitrary Python code. It's notably useful for machine learning, data analysis, and education. The collection of digital images of different waste images is done by using a TCL t766S mobile phone camera with a Resolution 720*1600 and a Logi Techc-720 USB camera is used to take images for the test dataset in real-time.

Proposed architecture

Can a machine accurately explain the content of an image or video the same way a person could? A machine's ability

to accurately describe the contents of an image or video is subjected to the Turing test in computer vision. In order to determine the answer to this issue, the development of a deep learning algorithm for image classification is examined in this work. Deep learning has greatly increased the accuracy rate of many computer vision tasks. YOLO which is the state of an art single-stage real-time object detection algorithm is proposed. It is based on the Convolutional Neural Network [17]. This algorithm can identify objects in images using webcam input in real-time, video input, and image input. Here, in this thesis, Yolov4 and Yolov4-tiny (compressed version of the original model with fewer parameters. This model is often referred to as a lightweight version of the original Yolov4 model and it can be deployed on various edge devices) model is used.

Architecture: YOLO is a state-of-the-art real-time object detection algorithm based on a Convolutional Neural Network. It was developed by Joseph Redmon in 2016. This technique can identify objects in images using webcam input in real-time, video input, and image input. YOLOv4 uses an Artificial Neural Network approach to find objects in images. This network segments the image into regions and forecasts the probability and bounding box of each region. The bounding box is then compared to each anticipated probability after each. When many bounding boxes are found for the same object, Non-Max Suppression is employed to make a determination [17].

The Prediction (head network), Backbone network, and Neck network are the three main divisions of the YOLO network architecture. The backbone network is primarily in charge of extracting image features, however as deep learning has advanced, it has been shown that while the number of layers in the network increases, so does the amount of extracted feature data and thus increases training costs. Instead, its training impact will diminish after a certain number of levels. The Neck network can enhance shallow features derived from the backbone network, process and refine those characteristics, and blend shallow and deep features to boost network robustness and produce more useful features. The Head network classifies and regresses the features obtained by the backbone and neck networks (Figure 4) [18]. We will see each network in detail.

The main network of YOLOV4 is CSPDarknet53. A convolutional neural network with 53 layers is called DarkNet-53. Cross-StagePartial-Connection is referred to as CSP. DenseNet and CSP are used by CSPDarknet53 to increase

convolutional networks' capacity for learning, reduce memory and computation requirements for network models, and maintain accuracy. The input feature map channel layer is split in half before each residual network in Darknet's five residual modules, and CSP is added after each large residual module [20]. The CSPDarknet53 backbone network was built using the Darknet53 development. The basic residual module was added with the CSP structure shown below (Figure 5).

The CSP structure divides the original residual module into two parts, one of which is directly connected to the residual network and the other of which is connected via the residual network. The output of the weak network is merged. With this approach, fewer variables and less computation are required while achieving great accuracy [22].

The Mish function is basically the activation function in CSPDarknet53. Mish, a novel self-regularized non-monotonic activation function [23].

Mish is bounded below and unbounded above, and its range is [0,31]. Mish reduced the conditions for the Dying ReLU phenomenon intentionally in order to save a small bit of unfavorable information. The ReLU function can become saturated as a result of a significant negative bias, and this can prevent the weights from being updated during the backpropagation phase, making the neurons useless for prediction.

Mish properties promote improved communication and expressivity. Being unbounded above, Mish avoids saturation, which normally causes training to slow down owing to near-zero gradients substantially. Being confined below is also advantageous since it results in a substantial regularization effect [23].

The model will be better able to identify items at multiple scales thanks to the feature pyramid, enabling it to identify the same thing at varied sizes and scales. A feature extractor known as a Feature Pyramid Network, or FPN, produces proportionally scaled feature maps at several levels in a completely convolutional manner from a single-scale image of any size. In order to be employed in applications like object detection, it serves as a general method for creating feature pyramids inside deep convolutional networks (Figure 6) [24].

YOLOV4's Neck network makes use of PANetP and SSP

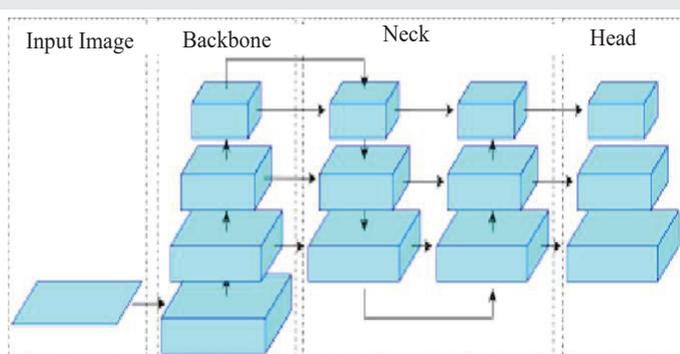


Figure 4: Network structure for yolov4 [19].

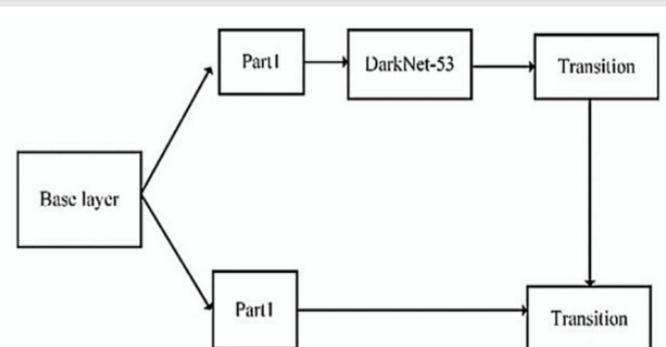


Figure 5: CSPDarkNet53 for yolov4 [21].

(Spatial Pyramid Pooling). In order to obtain feature maps with the same dimensions, SSP first convolves the candidate pictures with sliding kernels of four different sizes: 1, 5, 9, and 13 [26]. The spatial size of each candidate map may be preserved via SSP (Figure 7).

Before joining feature maps with different core sizes as output, SSP maintains the spatial size of each candidate map, resulting in a fixed-size feature map. The development of PANet is based on FPN and Mask RCNN. PANet presents a more flexible ROI Pooling (Region of Interest Pooling) that can extract and integrate features at different sizes, whereas FPN exclusively extracts data from high-level feature layers.

After all, by using all the fused features in the neck network, most prediction work is done during the detection stage. The head's function in a one-stage detector is to do dense prediction. The dense prediction, which includes the label, the prediction's confidence score, and a vector containing the center, height, and breadth of the anticipated bounding box, is the final prediction.

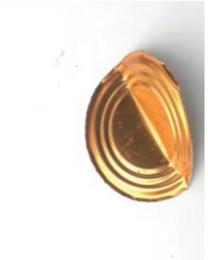
Data collection: In this research, 2529 waste are taken from The Stanford TrashNet Dataset [28], And 1,000 images of waste from Repi dump sites as well from households and common platforms in which the town's wastes are collected using mobile phones. This dataset includes seven waste types: glass, metal, cardboard, organic paper, plastic, and trash. The dataset is separated into three sections: one for training, one for testing, and one for validation. The dataset is a .jpg image

with a resolution of 416×416 pixels. So, number of images used in this study is 3529. Sample datasets are shown in Table 1.

The dataset is constructed and separated into seven groups. It includes real-time wastes that are intermingled with each other. This is the unique feature of the research that the model is trained as it can detect more than one waste group at a time.

Data labelling: After constructing our dataset, labeling will be the next task to be performed. Labeling every image with a tool that produces a.txt file containing image data, such as LabelImg. It is done by using the LABELIMG tool. LabelImg is a free, open-source program for marking images graphically labeled or annotated (Figure 8).

Table 1 Single and mixed type of waste sample dataset.

Name	Image	Name	Image
Cardboard		Metal	
Glass		Organic	
Paper		Trash	
Plastic		Mixed	
Mixed		Mixed	

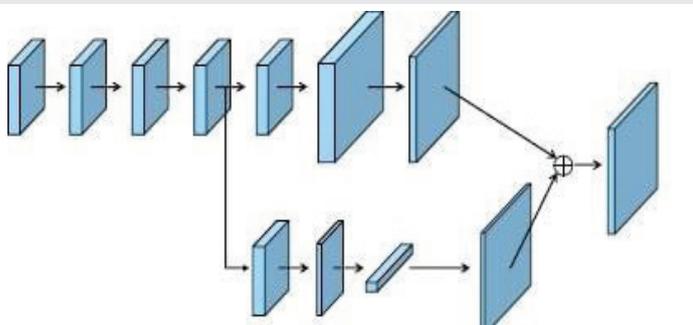


Figure 6: Path aggregate network structure for yolov4 [25].

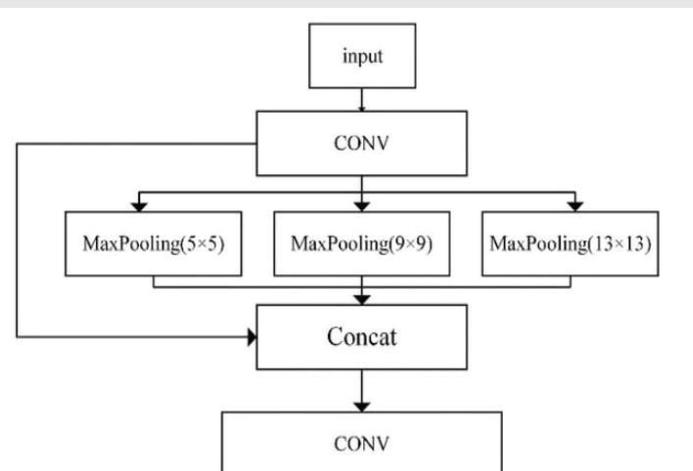


Figure 7: Spatial pyramid pooling for yolov4 [27].

Data augmentation: To create variances within the data so that it can accurately generalize the unknown data data-augmentation is a technique used. Data Augmentation is a technique used for manipulating/modifying data without losing its essence (Figure 9).

Mosaic Data augmentation was performed in this study to replicate the training data and to increase the context information that can be found in a single image so that it can increase the learning ability of the model.

Experiment

Subdivision and Mosaic Augmentation Parameters Tuning

The test is conducted by changing the subdivision value and mosaic data augmentation technique as tuned parameters in the YOLOv4tiny model. Because it is needed to change the variable hyperparameters in this case subdivisions and the mosaic augmentation effects accordingly to match the GPU-RAM performance of Colab. The experiment setup for subdivision and mosaic data augmentation is shown in Table 2.

As we said earlier in this section sub-division and mosaic



Figure 8: Labeling the Datasets.



Figure 9: Mosaic data augmentation.

Table 2 experiment setup for subdivision and mosaic data augmentation.

Tuning/experiment	Sub-division	mosaic
1	16	With mosaic
2	8	Without mosaic
3	16	With mosaic
4	8	Without mosaic

data augmentation are taken as a parameter for evaluating the performance of the model for the given dataset.

Results

Training result for YOLOv4 model with Sub-division value 16 and mosaic data augmentation

The best weight values that result from the YOLOv4 model are yolov4-custom_best weights, which are the outcomes of the training and validation process with a remarkable AP value for each waste class. It has an AP value of 95.28% for cardboard, 94.48% for Glass, 93.28% for Metal, 83.28% for Organic, 95.33% for Paper, 88.09% for Plastic, and 89.00% for Trash class. The model works best for all classes; however, it has little trouble detecting organic wastes from stacks of veggies that were placed in the organic class. It will be mistakenly placed in the trash classes. This happens when stacks of vegetables look to be mingled together with each other or especially with the trash class of waste. Even human sight would make this mistake.

The blue curve represents training loss basically the CIoU loss (Complete Intersection-Over-Union) on the training dataset. As the curve progressively gets dropped after each iteration, it can be said that the model learns well during the training process. The red curve represents the mean average precision of the model which depicts if the model generalizes effectively on a *never-before-seen dataset* or *validation set*. The YOLOv4 model has the best mAP of 91.25 % and an average loss of 0.4679. It takes 16 hours to complete 14,000 iterations. As it can be said from Figure 10, the continuous drop in loss value after each iteration shows YOLOv4's effective learning process. The YOLOv4 loss function-generated curve is remarkably stable after 8400 epochs. This training is carried out by applying mosaic data augmentation on the training dataset.

Training result for YOLOv4-tiny model with sub-division value 8 and mosaic data augmentation

The yolov4-tiny_custom_best weight is the best weight value which is the outcome of the training of the YOLOv4-tiny model which results in a very good AP value for each class except for the organic class. The model scored Cardboard 81.96%, Glass 90.00%, Metal 91.23%, Organic 51.88%, Paper 93.37%, Plastic 91.86%, and Trash 72.95% AP values.

The blue curve represents training loss basically the CIoU loss (Complete Intersection-Over-Union) on the training dataset similar to the above Figure 11. The red curve represents the mean average precision of the model which depicts if the model generalizes effectively on a *never-before-seen dataset* or *validation set*. The YOLOv4-tiny model has the best mAP of 82.02 % and an average loss of 0.1059. It takes 3 hours (13 hours

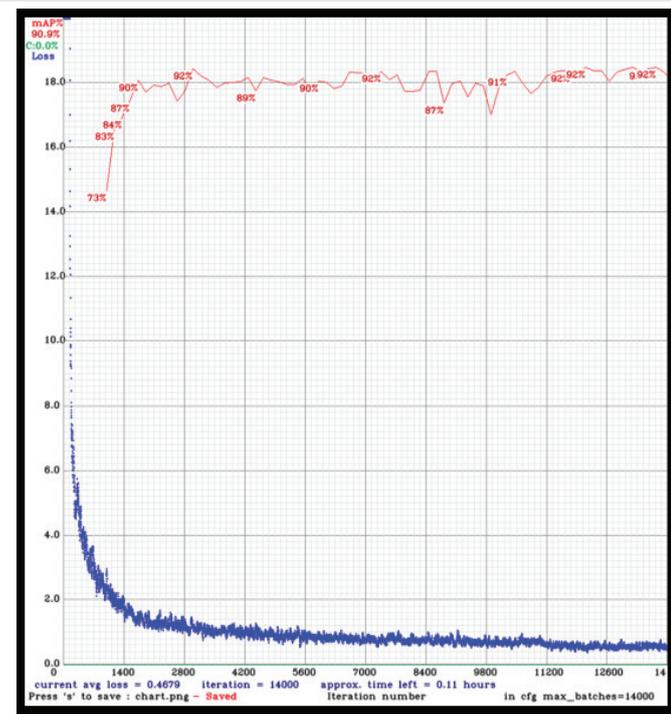


Figure 10: Average Loss and mAP for YOLOv4 iteration 1000 – 14,000 with mosaic data augmentation.

faster than the YOLOv4 model inference time) to complete 14,000 iterations.

This training is carried out by applying mosaic data augmentation on the training dataset and taking the subdivision value to be 08. After 7000 iterations, the curve produced by the loss function YOLOv4-tiny is quite stable.

The YOLOv4-tiny models have a lower AP value than that of the YOLOv4 model. As stated earlier, the YOLOv4 model scores a higher mean average value than the YOLOv4-tiny. The tiny version of the YOLOv4 model identifies well all classes, except the organic class. However, for organic wastes of piles of vegetables, it face difficulties in identifying organic wastes that were categorized under the class organic. It will misclassify with the trash classes.

Figure 11 depicts the Average Loss and mAP of the YOLOv4 Tiny model.

Subdivision and mosaic augmentation parameters tuning

The test is conducted by changing the subdivision value as well as mosaic data augmentation technique as tuned parameters in the model which has fewer parameters than the original YOLOv4 model that means in the YOLOv4-tiny model. Because it is needed to change the variable hyper-parameters in this case subdivisions and the mosaic augmentation effects accordingly to match the GPU RAM performance of Colab.

According to Table 3, using the subdivision value of 8 gives a mAP score that is about 2.4% higher than using the subdivision 16 value. The mAP value is 2.2% higher with mosaic data augmentation than it would be without it. This demonstrates that utilizing a mosaic data augmentation and smaller

subdivision values (8) improves the model’s performance. The subdivision and data mosaic augmentation settings not only impact the mAP value but also the computation speed. The table demonstrates that computation time decreases with decreasing subdivision value. In contrast, the model using the data mosaic augmentation takes more time than the model not using the data mosaic augmentation.

Discussion

According to Table 3, using the subdivision 8 value results in a mAP value that is about 2.4% higher than using the subdivision 16 value. The mAP value is 2.2% higher with mosaic data augmentation than it would be without it. This demonstrates that utilizing a mosaic data augmentation and smaller subdivision values (8) improves the model’s performance. The subdivision and data mosaic augmentation settings not only impact the mAP value but also the computation speed. The table demonstrates that computation time decreases with decreasing subdivision value. In contrast, the model using the data mosaic augmentation takes more time than the model not using the data mosaic augmentation.

When compared to prior studies, particularly those that used the TrashNet Dataset (12) with only four classes of waste using Resnet18 and Using this approach 81% testing accuracy is attained. The accuracy of the YOLOv4 model created in

Table 3 Effect of Mosaic Data Augmentation and Subdivision Parameters.

Subdivision	Mosaic	mAP	Time
16	Without mosaic	77.35	150 minutes
16	With mosaic	79.86	180 minutes
8	Without mosaic	80.06	120 minutes
8	With mosaic	82.02	185 minutes

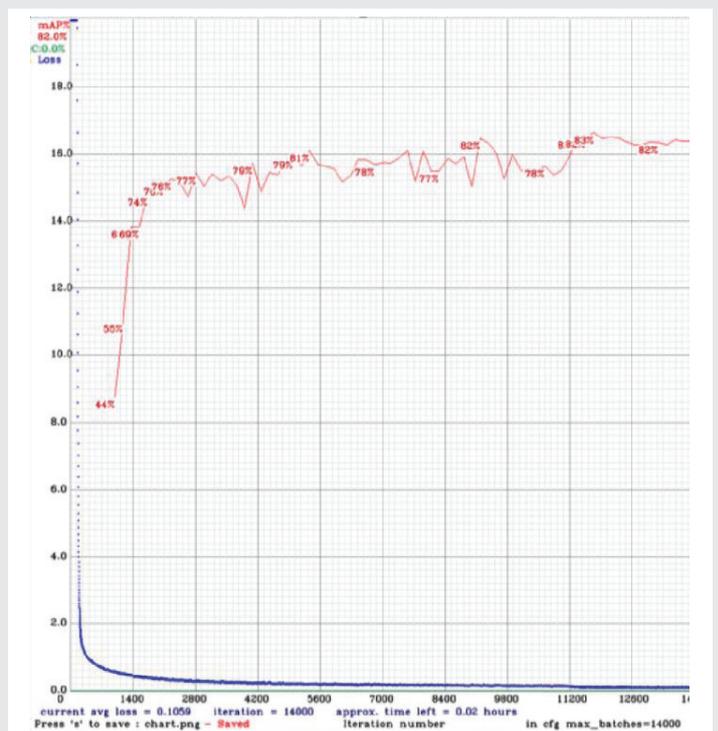


Figure 11: Average Loss and mAP for YOLOv4-tiny iteration 1000 – 14,000 with mosaic data augmentation.



this study is higher. Another study [28] did research CNN for waste classification achieving 76% testing accuracy. Which is lower than the model in this thesis that uses Mosaic data augmentation.

The algorithm outperforms all the reviewed approaches for waste detection. Besides the model performs well on the jumbled set of wastes in real life. On the contrary, the papers on waste detection using a deep learning approach are not tested in real-life data. This Study additionally shows computation speed / projected time results that were not shown in the prior two studies. Real-time waste segregation becomes reality using the yolov4-waste model with an average fps (frame per second) of 52.4. The model can be used to detect waste in real-time. The comparisons can be summarized as follows

Conclusion

In this paper, a classification model and real-time object detection that can show the bounding box and prediction probability of objects in each image are done. The model can be run on Google Colab, a local PC, or a laptop. The model can also be run with three various inputs, including images, videos, and a webcam. The findings reveal that the YOLOv4-tiny architecture, which is simpler than the YOLOv4 design, influences the algorithm's performance in both prediction probability and prediction time. Where YOLOv4 outperforms in prediction probability. The mAP value is 91.25 percent, the precision is 0.9 % the recall is 0.88 %, F1-Score is 0.89, and the average IOU is 80.47%. Meanwhile, as compared to the YOLOv4 model, YOLOv4-tiny has a faster computing speed. Although YOLOv4-tiny has good computing efficiency, it does not have the same detection capabilities or prediction probability as the YOLOv4 model. The mAP value of the YOLOv4-tiny model is 82.02 %, the precision is 0.75 %, the recall is 0.76 %, F1-Score is 0.75%, and the average IOU is 63.59%.

Acknowledgments

Department of Electromechanical Engineering, Department of Software Engineering College of Electrical and Mechanical Engineering, High-Performance Computing and Big Data Analytics, Addis Ababa Science and Technology University, Addis Ababa, Ethiopia.

References

1. WHO. COMPENDIUM OF WHO AND OTHER UN GUIDANCE ON HEALTH AND ENVIRONMENT: 2022 UPDATE. WHO FACT SHEET. 2022; 2019(DECEMBER):5. [HTTPS://WWW.WHO.INT/EN/NEWS-ROOM/FACT-SHEETS/DETAIL/ARSENIC](https://www.who.int/en/news-room/fact-sheets/detail/arsenic)
2. NAHMAN A, YONLI AH, FISSEHA GG, JAMIDU HY, KIDANE GG, OSMAN MOHAMED ABDELNAEM M, AHMED MT, MOHAMMED MA, LOUTFY NM, OSIBANJO O, OELOFSE SR. AFRICA WASTE MANAGEMENT OUTLOOK. UN ENVIRONMENT PROGRAMME, NAIROBI, KENYA. VOL. 123, PROGRAMME ONU ENVIRONNEMENT. 2018; 123.
3. WEGHMANN V, NIEKERK S VAN. MUNICIPAL SOLID WASTE MANAGEMENT SERVICES IN AFRICA AND ARAB COUNTRIES. MUNIC WORK SERV MUNIC SERV MUNIC. 2018; 1-68.
4. WORLD HEALTH ORGANIZATION. REGIONAL OFFICE FOR THE WESTERN PACIFIC. GUIDES FOR MUNICIPAL SOLID WASTE MANAGEMENT IN PACIFIC ISLAND COUNTRIES. HEAL CITIES - HEAL ISLANDS DOC SER. 1996;

- 6: 183; 1965). (CONSTRUCTED IN. OPENLITTERMAP. COM – OPEN DATA ON PLASTIC POLLUTION WITH BLOCKCHAIN REWARDS (LITTERCOIN). 2018.
5. PAANANEN-PORKKA M. SOLID WASTE MANAGEMENT IN ADDIS ABABA. 2016.
6. MOHAMMED A, ELIAS E. DOMESTIC SOLID WASTE MANAGEMENT ENVIRONMENTAL IMPACTS IN ADDIS ABABA CITY. J ENVIRON WASTE MANAG. 2017; 4(1):194-203.
7. BJERKLI CL. CAMILLA LOUISE BJERKLI URBAN SERVICES AND GOVERNANCE. 2013.
8. O'MAHONY N, CAMPBELL S, CARVALHO A, HARAPANAHALLI S, HERNANDEZ GV, KRPAKOVÁ L. DEEP LEARNING VS. TRADITIONAL COMPUTER VISION. ADV INTEL SYST COMPUT. 2020; 943: 128-44.
9. GELAN E. MUNICIPAL SOLID WASTE MANAGEMENT PRACTICES FOR ACHIEVING GREEN ARCHITECTURE CONCEPTS IN ADDIS ABABA, ETHIOPIA. TECHNOLOGIES. 2021; 9(3):48.
10. SALMADOR A, CID JP, NOVELLE IR. INTELLIGENT GARBAGE CLASSIFIER. INT J INTERACT MULTIMED ARTIF INTEL. 2008; 1(1):31-6.
11. MITTAL I, TIWARI A, RANA B, AND SINGH P. TRASH CLASSIFICATION: CLASSIFYING GARBAGE USING DEEP LEARNING. 2020; 11 WWW.JESPUBLICATION.COM
12. GYAWALI D, REGMI A, SHAKYA A, GAUTAM A, SHRESTHA S. COMPARATIVE ANALYSIS OF MULTIPLE DEEP CNN MODELS FOR WASTE. 2019.
13. COSTA BS, BERNARDES ACS, PEREIRA JVA, ZAMPA VH, PEREIRA VA, MATOS GF. ARTIFICIAL INTELLIGENCE IN AUTOMATED SORTING IN TRASH RECYCLING. 2019; 198-205.
14. LIU Y, FUNG K-C, DING W, GUO H, QU T, XIAO C. NOVEL SMART WASTE SORTING SYSTEM BASED ON IMAGE PROCESSING ALGORITHMS: SURF-BoW AND MULTI-CLASS SVM. COMPUT INF SCI. 2018; 11(3):35.
15. ZHANG Q, YANG Q, ZHANG X, BAO Q, SU J, LIU X. WASTE IMAGE CLASSIFICATION BASED ON TRANSFER LEARNING AND CONVOLUTIONAL NEURAL NETWORK. WASTE MANAG. 2021 NOV; 135:150-157. DOI: 10.1016/j.wasman.2021.08.038. EPUB 2021 SEP 8. PMID: 34509053.
16. REDMON J, DIVVALA S, GIRSHICK R, FARHADI A. YOU ONLY LOOK ONCE: UNIFIED, REAL-TIME OBJECT DETECTION.
17. HUANG Y, ZHENG J, SUN S, YANG C, LIU J. APPLIED SCIENCES OPTIMIZED YOLOv3 ALGORITHM AND ITS APPLICATION IN TRAFFIC FLOW DETECTIONS. 2020.
18. WU L, MA J, ZHAO Y, LIU H. APPLE DETECTION IN THE COMPLEX SCENE USING THE IMPROVED YOLOv4 MODEL. AGRONOMY. 2021; 11(3).
19. HUANG G, LIU Z, VAN DER MAATEN L, WEINBERGER KQ. DENSELY CONNECTED CONVOLUTIONAL NETWORKS. PROC - 30TH IEEE CONF COMPUT VIS PATTERN RECOGNITION, CVPR 2017; 2261-9.
20. WANG CY, MARK LIAO HY, WU YH, CHEN PY, HSIEH JW, YEH IH. CSPNET: A NEW BACKBONE THAT CAN ENHANCE THE LEARNING CAPABILITY OF CNN. IEEE COMPUT SOC CONF COMPUT VIS PATTERN RECOGNIT WORK. 2020; 1571-80.
21. LI X, DUAN C, ZHI Y, AND YIN P. WAFER CRACK DETECTION BASED ON YOLOv4 TARGET DETECTION METHOD. IOP CONF SER EARTH ENVIRON SCI. 2021; 1802:2.
22. MISRA D. 1908.08681V2. MISH A SELF REGUL NON-MONOTONIC NEURAL ACT FUNCT. 2019; 1.
23. LI X, LAI T, WANG S, CHEN Q, YANG C, CHEN R. WEIGHTED FEATURE PYRAMID NETWORKS FOR OBJECT DETECTION. PROC - 2019 IEEE INTL CONF PARALLEL DISTRIB PROCESS WITH APPL BIG DATA CLOUD COMPUT SUSTAIN COMPUT COMMUN SOC COMPUT NETWORKING, ISPA/BDCloud/SUSTAINCOM/SOCIALCOM 2019. 2019; 1500-4.
24. LIU S, QI L, QIN H, SHI J, JIA J. PATH AGGREGATION NETWORK FOR INSTANCE SEGMENTATION. PROC IEEE COMPUT SOC CONF COMPUT VIS PATTERN RECOGNIT. 2018; 8759-68.
25. ZAIDI SSA, ANSARI MS, ASLAM A, KANWAL N, ASGHAR M, LEE B. A SURVEY OF MODERN DEEP LEARNING BASED OBJECT DETECTION MODELS. 2021; 1-18. [HTTP://ARXIV.ORG/ABS/2104.11892](http://arxiv.org/abs/2104.11892)
26. HE K, ZHANG X, REN S, SUN J. SPATIAL PYRAMID POOLING IN DEEP CONVOLUTIONAL NETWORKS FOR VISUAL RECOGNITION. 1-14.
27. YANG M, THUNG G. CLASSIFICATION OF TRASH FOR RECYCLABILITY STATUS. 1-6.
28. SIDHARTH R, ROHIT P, VISHAGAN S, KARTHIKA R, GANESAN M. DEEP LEARNING BASED SMART GARBAGE CLASSIFIER FOR EFFECTIVE WASTE MANAGEMENT. PROC 5TH INT CONF COMMUN ELECTRON SYST ICCES. 2020; 1086-9.