Battery Waste Detection Using Selective Search with Deep Learning Classification

Allan Magno E. Pecundo
Department of Information Systems & Computer Science
Ateneo de Manila University
Quezon City, Philippines
allan.pecundo@obf.ateneo.edu

Patricia Angela R. Abu, PhD
Department of Information Systems & Computer Science
Ateneo de Manila University
Quezon City, Philippines
pabu@ateneo.edu

Vidal Wyatt Lopez
Department of Information Systems & Computer Science
Ateneo de Manila University
Quezon City, Philippines

***Abstract*— Waste generation remains an issue in the Philippines with large amount of waste and lack of consistent waste management. E-wastes, like batteries, lack widespread waste management infrastructure despite their hazardous nature when improperly disposed of. Currently, there have been detection models built to aid waste management, mostly based on modern deep-learning techniques like two-stage or single stage object detectors. However, no research has been done on e-waste detection and there are no publicly available e-waste detection datasets that can be used to train more modern object detection models due to the requirement of bounding box annotations. This research aimed to build an object detector for batteries as an e-waste using only a classification dataset for the training. The model built was able to detect batteries at 8% mAP@.50 but performance increased to 24% mAP@0.30; this was due to the limitation of the unsupervised region proposal method, Selective Search. The model built may be applied to intelligent waste management systems to help manage the segregation and address improper disposal of batteries which require extra care compared to more common trash. This study may also act as a stepping stone for other researchers to further improve the object detection model by taking the steps necessary to use better models such as creating datasets that may be used directly by more modern object detection models.**

Keywords— Object Detection, Deep Learning, Waste Management, E-Waste

I. Introduction

In the Philippines, waste generation remains a dominant problem that affects communities, especially those within urban areas. According to a Senate report, the Philippines has produced approximately 40,087.45 Tons of waste daily for the past year of 2016, an increase from the previous year's average of 39,422.46; with the National Capital Region (NCR) alone contributing to approximately 9,212.92 Tons of waste or 23% of the total country's waste [1]. The amount of waste produced is also commonly mishandled with multiple areas around the cities in the Philippines have garbage lying around various areas from public streets, to residential areas, or even coastal areas; all of which pose threats from the garbage disposed around the area.

Even with government policies at either the national level or local level formed to try and mitigate the waste problem, such as plastic ban or stronger push for recycling systems, the problem still persists as seen in the forms of multiple forms of trash found littered on the streets and even bodies of water. The lack of compliance seen in following the recycling system poses a challenge to the implementation of the country's waste management system. The landfills of local governments are also seen to contain massive volumes of unsegregated garbage. This problem becomes even larger for less common forms of trash that are not part of the usual "recyclable" category.

There is this growing field of AI application in the form of intelligent waste management systems, wherein AI models are developed to help tackle the problems of waste management systems. One example is the development of waste recycling robots to aid waste management. Greyparrot AI is one company that specializes in Computer Vision to help automate recycling as well as provide analytics regarding the state of the current garbage system such as trash composition [2]. Some countries have even adapted the use of some form of AI to aid in their waste management systems. In the US, robots have been installed in a recycling plant in Sarasota, Florida with the task of helping identify and segregate the different types of trash using Computer Vision [3].

One area of waste management which is both important yet challenging to implement is the management of electronic waste (e-waste). Battery is a common example of an item that would often end up as electronic waste given the multiple devices in a household that would normally require this. Certain batteries in particular are hazardous when left to degrade in the environment such as rechargeable batteries which may contain cobalt, cadmium, and lead; heavy metals which may degrade surrounding air, topsoil, and groundwater if present [4]. These batteries require more special treatment in terms of proper waste management due to their potential risk, especially if improperly disposed. Lithium batteries for example pose as fire risk and are combustible when damaged or mishandled [5]. Waste management infrastructure in the Philippines geared towards e-waste is not as common or publicly accessible which poses a risk in terms of mishandling of battery as a particular waste.

With this in mind, the research area of AI application, specifically Computer Vision, in this area of waste management will be explored. The study aims to tackle research on object detection applied to the context of detecting battery as an e-waste. Tackling this application of AI into e-waste can provide future solutions to the challenges of battery waste management in the Philippines, given the lack of consistent infrastructure that allows for proper disposal of this type of waste. The objectives of the research will be as follows:

- To create an object detection model, specifically for the detection of single-cell batteries

II. Review of Related Literature

A. OBJECT DETECTION METHODS

Object detection has been a field of Computer Vision that is growing with popularity in terms of its various applications. A review on the different development of object detection models have categorized most models into 2 types: (1) Traditional Models, and (2) Deep Learning-Based Models [6].

Commonly used Traditional forms of object detection mentioned in the review [6] are (1) Scale-Invariant Feature Transform or SIFT, (2) Histogram of Gradients or HOG, (3) Speeded Up Robust Features or SURF, and (4) Oriented FAST and rotated Brief or ORB. Most of the traditional forms of garbage detection use these methods as some form of pre-processing or transformation of the images which aid in being able to detect the image; such as how HOG identifies the object of interest by identifying where the object is present represented by the different intensities of gradients in the image. One study in where this form of object detection was used was in a study by Zhou and Yu, wherein the built a pedestrian detection model using am SVM classifier with HOG as a form of feature extraction to detect the object [7]. These traditional forms of object detection however are usually paired with region search algorithms, such as sliding windows, to enable the location of the object of interest before classification happens so that region boundaries are proposed by the model.

Deep Learning-Based forms of object detection are commonly classified into 2 types [6]: (1) Two-stage classifiers such as R-CNN, and (2) One-stage classifiers such as YOLO and SSD. Two-stage classifiers detect object through the identification of the regions of interest or bounding boxes where the object may be, and the classification of the object found within that identified region, done as separate steps. Convolutional Neural Networks are usually used to come up this. One-stage classifiers detect the object by identifying the position of the object through estimation of the bounding box of the object and the identification of the class all in a single step. One-stage classifiers are found to usually outperform the two-stage classifiers in terms of speed.

B. STUDIES ON GARBAGE DETECTION

On the application to garbage detection, there have been studies wherein datasets with images of garbage have been used in the context of either classification or detection within images. Most studies on garbage detection or classification found have employed the use of Deep Learning-Based methods.

One study by Patel et. al utilized a mix of images from SpotGarbangeGini Github repository and Google Images to create garbage detection models using 5 different object detection models namely; (1) EfficientDet-D1, (2) SSD ResNet-50 V1, (3) Faster R-CNN ResNet-101 V1, (4) CenterNet ResNet-101, and (5) YOLOv5M and compare which of the models yielded the best performance in terms of detecting garbage on the dataset used. YOLOv5M ended up with the best performance yielding mean average precision (mAP) of 0.613 [8].

Another study by Li et al, utilized a one-stage classifier in YOLOv3 to detect garbage on water surfaces. This model was implemented on a capture robot which was able to detect and clean the garbage found floating on the water it was deployed in. The model was able to yield a high mAP of 91.43 [9].

A study by Han also made use of YOLOv3 system, with an improved mobilenetV3-large backbone called mobilenetV3-Lite, for garbage detection. In this study, the garbage was classified into 4 groups, namely (1) recyclable garbage, (2) kitchen garbage, (3) hazardous garbage, and (4) other garbage wherein the final model was able to get a mean average precision (mAP) of 89.89. It was also deployed on a Jetson Nano device wherein it was able to accurately classify the garbage type upto 94.56% while the quantity of garbage detected was as high as 90.91% in terms of accuracy [10].

A study by Arai et al developed a model for detecting and classifying garbage, either classified as (1) paper trash, (2) plastic bottle, or (3) lunchbox trash. The YOLOv2 model was used for deployment where multiple experiments were conducted; one experiment focused on testing the accuracy of the garbage detection in Haneda Innovation City and another

Authors	Objective	Model Approach	Specific Method
Patel et al [8]	Garbage Detection (4 Classes: street waste, household waste, medical waste, other waste)	Deep-Learning Based	EfficientDet-D1, SSD ResNet-50 V1, Faster R-CNN ResNet-101 V1, CenterNet ResNet-101, YOLOv5M
Li et al [9]	Garbage Detection (3 Classes: bottles, bags, styrofoam)	Deep-Learning Based	YOLOv3
Han [10]	Garbage Detection (4 Classes: recyclable waste, kitchen waste, hazardous waste, other waste)	Deep-Learning Based	YOLOv3 with mobilenetV3-Lite backbone
Arai et al [11]	Garbage Detection (3 Classes: can, PET, lunchbox)	Deep-Learning Based	YOLOv2 with benchmarking against YOLOv3, SSD, DSSD, and Faster R-CNN
Karthikeyan et al [12]	Garbage Detection (2 Classes: biodegradable, non-biodegradable)	Deep-Learning Based	SSD with Augmented Clustering NMS
Anjum et al [13]	Garbage Localization	Deep-Learning Based	CNN

Figure 1. Summary of Reviewed Literature

experiment where the model was benchmarked against other object detection models. In the experiment at Haneda Innovation City, the main YOLOv2 model yielded an accuracy of 84.5%. In the other experiment where it was compared to other object detection models, the YOLOv2 model yielded an accuracy of 93.0% although the DSSD model yielded the highest accuracy with 97.8% [11].

A study by Karthikeyan et al used an SSD model with augmented clustering non-max suppression (NMS) to create a garbage detector that would be able to identify the bounding box of the garbage object and also classify whether the detected garbage was biodegradable or non-biodegradable for the purpose of aiding waste management. The model yielded a performance of 0.965 mAP [12].

Another study by Anjum et al, used CNN to localize areas of an image wherein the garbage was found. Unlike garbage detection studies where the objective is to identify bounding boxes where the items of interest are found as well as classify the garbage type, the study was only done only with the objective of identifying which pixels in an image contained the garbage. The model was trained on labelled images where there were areas of an image annotated which pixels had garbage and which ones were non-garbage and the model was tested on separate non-labelled images. Performance was measured based on a survey of respondents who would rate how much overlap there was between the localization of garbage identified by the model and the actual area where garbage was found on the picture [13].

III. METHODOLOGY

A. DATASET PREPARATION

There are very few open-source or public object detection datasets related to the area of waste management such as TACO-master [14] and Trash-ICRA19 [15]; with very few of the observations that are within the area of electric waste, such as batteries. With this in mind, the study will make use of labelled images of batteries to train the object detector with the limitation being that the methodology will be restricted to a detector built on a separate classifier model paired with a region proposal or object segmentation model due to the lack of datasets with the "battery" class having the standard format required for more modern deep-learning object detection algorithms such as R-CNN, YOLO, and SSD which require both class and bounding box ground-truth labels.

A publicly available dataset was found on Kaggle [16] wherein different types of images of items that are commonly found as waste were sorted into folders based on their labels. In total there are 12 different class labels identified: battery, biological, brown-glass, carboard, clothes, green-glass, metal, paper, plastic, shoes, trash, and white-glass. One of the labels of interest collected was the battery label hence this dataset being feasible in building an object classifier which can be later used for object detection when coupled with other techniques such as object segmentation or region proposal.

For building the model, the dataset used will make use of only 2 classes ("battery" as the main class and "not battery" as balance classes for the classifier model) from the original dataset.

The "not battery" class was created from taking random images of other classes in the available dataset, mostly from the "plastic" and "cardboard" class.

To allow the model to capture detect batteries that have been improperly disposed (e.g., left on the streets, ground, landfills, and etc.), the "battery" class of data set that will be used for training will be augmented to allow for the model to detect different orientations in which the objects that are lying around might be found at.

In total, the class distribution of the final dataset after compilation and augmentation is as follows:

- battery (653 images)
- not battery (1756 images)

To validate the performance of the model, a separate test dataset composed of 12 images scraped from Google Image Search, containing batteries. Annotations of labels and bounding boxes were manually created by the researcher using LabelIMG [17]. The setting of the batteries in the images varied from those placed on table surface, held by a hand, and those found disposed in the ground.

B. BUILDING THE MODEL

The object detector model was built using 2 separate models: a region proposal detector, and a classifier that would predict whether the proposed region contained the classes of interest or not. This model is proposed given the current dataset which does not contain ground-truth bounding box labels of the object of interest. The proposed method allows the use of a trained classifier model in the context of object detection with the inclusion of a region proposal detector that does not require prior training.

1) SELECTIVE SEARCH ALGORITHM [18]

To identify the possible areas where the object of interest (battery) is located, the Selective Search algorithm was used. Most modern deep-learning detection methods, both two-stage and one-stage, have built in region proposal in the model. These models however require the use of ground-truth region annotations; which are not present in the current dataset.

Selective Search algorithm identifies regions of interest where the object may be present similar to unsupervised learning. The algorithm starts with selecting an initial set of regions proposed using graph-based segmentation. Then, among all the regions proposed, the algorithm then identifies which proposed regions are similar and combines them to form a larger region; these combined regions are then added into the list of proposed regions as larger regions. This process is then repeated until the algorithm ends up with only 1 large region remaining. The output of this algorithm will then be a set of region proposals of different scales from small ones (from the initial regions proposed) to larger ones (the combined regions based on similarity). The similarity criteria that the algorithm can consider when determining what regions to combine are color, fill, size, and texture .

Figure 2. Illustration of Selective Search region proposals

2) RESNET-50 CONVOLUTIONAL NETWORK [18]

To classify whether the proposed object region is a battery or not, the ResNet-50 model was used.

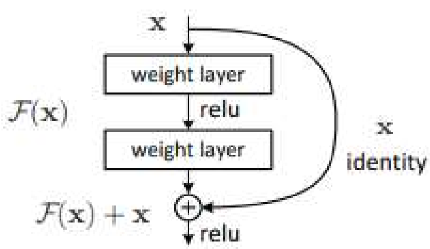

Figure 3. Illustration of Residual building block [19]

The ResNet-50 architecture makes use of Residual blocks, illustrated in Figure 3, which are combinations of connect weight layers, similar to plain Neural Network architectures, except with the incorporation of skip connections aside from the standard direct feedforward connections. The skip connects, use an identity mapping which means that certain layers in the model will take into account inputs from the previous connected layer as well as inputs from the skip connect a few layers prior. This type of architecture allows the model to contain the benefit of being able to fit more complex functions that is allowed in deeper network architectures but reduce the problem of vanishing gradients that is present in training models with more actual number of layers. Based on classification performance on the ImageNet 2012 dataset, the ResNet-50 architecture was found to outperform the VGG-16 network architecture.

3) OBJECT DETECTION MODEL FLOW

In the study, the use of Selective Search implementation in OpenCV coupled with Non-Max Suppression was used to be able to predict the region proposals while ResNet-50, with modifications done at the last layer to allow the model to return classification based on the 2 classes of interest (battery, not battery), was used as the classifier model. The framework is very similar to how the R-CNN object detector works without the fine-tuning of region proposals based on comparison with ground-truth bounding box labels.

The flow of the object detection model built is as follows:

1. Compile dataset of images of containing the classes of interest; "battery", "not-battery"

2. Use data augmentation techniques such as rotation, shearing, etc. to create larger dataset account for different possible orientations and conditions of the objects found on the image

3. Modify ResNet-50 classifier layer to account for number of classes and train on training dataset

4. Compile separate object detection test set composed of images with the classes of interest and annotate (class labels and bounding box ground-truth)

5. Using OpenCV selective search with non-max suppression, identify final region proposals (bounding boxes) that might contain objects

6. Using trained model, predict object class of interest on the different proposed bounding boxes

7. Filter final detections to "battery" object class

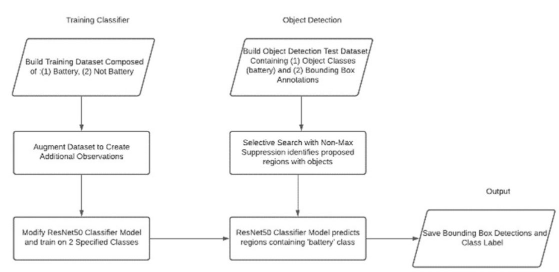

Figure 4. Flow of Object Detection Model

IV. RESULTS AND DISCUSSIONS

The detection of the model was measured with mean average precision (mAP) as the metric for performance, similar to what was used in the mentioned literature of object detection. Given the limitation of using a method that does not train the bounding boxes for detection, the mAP was measured using 3 different Intersection-Over-Union (IOU) thresholds: at 50%, 30%, and 20%. This was measured using the mAP tool developed by Cartucho [20].

The following table, Figure 5, summarizes the results of the object detection result of the model.

Metric	mAP@.50	mAP@.30	mAP@.20
Selective Search with ResNet-50	0.089	0.229	0.260

Figure 5. Summary of model performance

The resulting model tends to have difficulty in identifying the battery location if there are many batteries in the image and with the different batteries very close to each other.

Figure 6. Sample Image of Detection on Test Set (mAP@0.50)

In the sample image in Figure 6, the image was able to detect the 3 out of the 4 batteries correctly as represented by the green boxes which represent correct detection (e.g., detection box intersects with ground truth box, represented in blue). However, in the sample images there are also 2 false positive images seen, represented by the red boxes, while there is 1 remaining battery that was not detected correctly, represented by the box in pink.

Another challenge seen in the detection is the determination of the acceptable bounding boxes for the image. This is seen in how the model performance improves greatly with the increase of allowable IOU from 0.50 to 0.30.

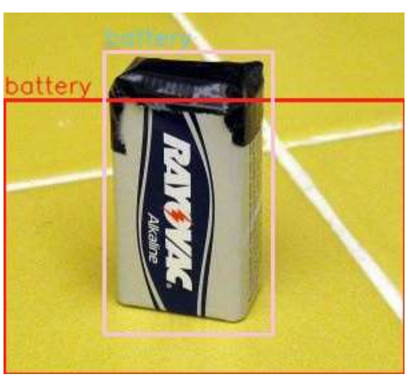

Figure 7. Sample Image of Single Battery Detection on Test Set (mAP@0.50)

Figure 8. Sample Image of Single Battery Detection on Test Set (mAP@0.30)

In Figure 7, we see that the model prediction bounding box includes a large area beyond the battery, which is not classified as a correct detection under the threshold of 0.50 IOU. Whereas in Figure 8, we see that under the threshold of 0.30 IOU, the detection is considered to be correct, illustrated by the green box color of the detection vs the red box in the previous figure representing incorrect detection.

The current model dependent on Selective Search for region proposal contains some limitations in determining the best bounding box for the prediction as it is mostly dependent on the ResNet-50 classifier model for determining the confidence of prediction for the proposed region.

The detection errors also get larger when the image either contains more batteries in the picture or have more textured surfaces such as grass or sand.

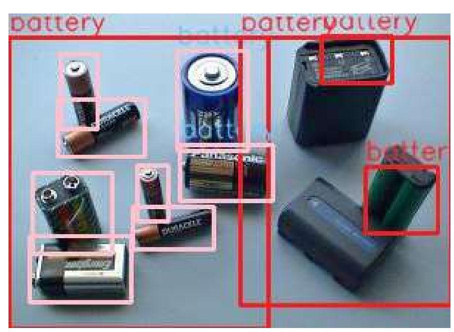

Figure 9. Sample Image of Multiply Battery Detection on Test Set (mAP@.50)

In Figure 9, the model will detect multiple batteries as one large battery causing a failed detection due to the small overlap of IOU driven by the single predicted bounding box having an area which covers all the batteries in the image, whereas each of the batteries within the large predicted bounding box have their own individually smaller ground-truth bounding box.

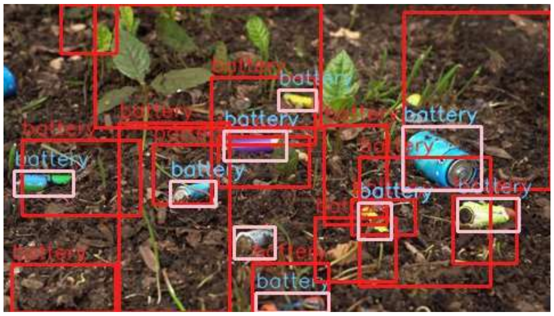

Figure 10. Sample Image of Multiply Battery Detection on Test Set with soil background (mAP@.50)

Figure 11. Sample Image of Multiply Battery Detection on Test Set with grass background (mAP@.50)

In Figure 10 and Figure 11, we see that the there aside from the large bounding box predictions mentioned in the previous example, there are cases wherein multiple false positive predictions made on the textured grass or soil. These multiple false positive predictions also contribute to the low mAP performance of the model at different metrics.

Overall, the model is still able to detect batteries although challenges with getting more accurate bounding boxes are present hence expectations on the IOU standard for considering detection should be lower given the limitation of the applicable methodology and training data.

V. Conclusion and Recommendations

While the model built was able to recognize batteries, the model was still not able to perform at the same level of the more modern deep learning models that will require the training data to contain ground-truth annotations of the locations of the object aside from classes. While this type of data is not as readily available, especially in the context of garbage detection, specifically for electronic waste like batteries.

It may still be the case that in order to create a much better performing model, more modern deep learning methods should be tested to see if the reduction of the false positive cases, as well as incorrect bounding box sizes get to be reduced. To be able to conduct this, future researchers should consider the additional step of annotating the training data with the ground-truth bounding boxes which are required by the deep-learning based object detection algorithms mentioned, as it is likely that the model will be able to learn how to predict the bounding boxes better rather than rely only on the suggestion bounding boxes generated by the Selective Search algorithm.

REFERENCES

[1] "Philippine Solid Wastes at a Glance," *Senate of the Philippines*, Nov-2017. [Online]. Available: http://legacy.senate.gov.ph/publications/SEPO/AAG_Philippine Solid Wastes_Nov2017.pdf. [Accessed: 11-Oct-2021].

[2] C. Hackl, "How 4 companies are using AI to solve waste issues on earth and in space," *Forbes*, 21-Jul-2020. [Online]. Available: https://www.forbes.com/sites/cathyhackl/2020/07/18/how-4-companies-are-using-ai-to-solve-waste-issues-on-earth--in-space/?sh=7bd0720035fa. [Accessed: 11-Oct-2021].

[3] M. Petrova, "America is drowning in garbage. Now robots are being put on duty to help solve the recycling crisis," *CNBC*, 27-Jul-2019. [Online]. Available: https://www.cnbc.com/2019/07/26/meet-the-robots-being-used-to-help-solve-americas-recycling-crisis.html. [Accessed: 11-Oct-2021].

[4] "Battery Recycling FAQs," ENVIRONMENTAL HEALTH & SAFETY - UC SANTA CRUZ. [Online]. Available: https://ehs.ucsc.edu/programs/waste-management/recycling-disposal/batteries-faq.html. [Accessed: 10-Dec-2021].

[5] "Used Household Batteries," United States - Environmental Protection Agency. [Online]. Available: https://www.epa.gov/recycle/used-household-batteries. [Accessed: 10-Dec-2021].

[6] K. Li and L. Cao, "A Review of Object Detection Techniques," 2020 5th International Conference on Electromechanical Control Technology and Transportation (ICECTT), 2020, pp. 385-390, doi: 10.1109/ICECTT50890.2020.00091.

[7] H. Zhou and G. Yu, "Research on pedestrian detection technology based on the SVM classifier trained by HOG and LTP features," *Future Generation Computer Systems*, vol. 125, pp. 604–615, 2021.*A review of object detection techniques*

[8] D. Patel, F. Patel, S. Patel, N. Patel, D. Shah and V. Patel, "Garbage Detection using Advanced Object Detection Techniques," 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), 2021, pp. 526-531, doi: 10.1109/ICAIS50930.2021.9395916.

[9] X. Li, M. Tian, S. Kong, L. Wu, and J. Yu, "A modified YOLOv3 detection method for vision-based water surface garbage capture robot," *International Journal of Advanced Robotic Systems*, vol. 17, no. 3, 2020.

[10] W. Han, "A YOLOV3 System for Garbage Detection Based on MobileNetV3_Lite as Backbone," 2021 International Conference on Electronics, Circuits and Information Engineering (ECIE), 2021, pp. 254-258, doi: 10.1109/ECIE52353.2021.00061.

[11] Y. Arai, R. Miyagusuku and K. Ozaki, "Development and Testing of Garbage Detection for Autonomous Robots in Outdoor Environments," 2021 IEEE/SICE International Symposium on System Integration (SII), 2021, pp. 138-143, doi: 10.1109/IEEECONF49454.2021.9382646.

[12] M. Karthikeyan, T. S. Subashini, and R. Jebakumar, "SSD based waste separation in smart garbage using augmented clustering NMS," *Automated Software Engineering*, vol. 28, no. 2, Aug. 2021.

[13] M. Anjum and M. S. Umar, "Garbage localization based on weakly supervised learning in Deep Convolutional Neural Network," 2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), 2018, pp. 1108-1113, doi: 10.1109/ICACCCN.2018.8748568.

[14] P. F. Proença and P. Simões, "TACO: Trash Annotations in Context for Litter Detection," arXiv.org.

[15] M. Fulton, J. Hong, M. J. Islam and J. Sattar, "Robotic Detection of Marine Litter Using Deep Visual Detection Models," 2019 International Conference on Robotics and Automation (ICRA), 2019, pp. 5752-5758, doi: 10.1109/ICRA.2019.8793975.

[16] M. Mohamed, "Garbage classification (12 classes)," Kaggle, 24-Jan-2021. [Online]. Available: https://www.kaggle.com/mostafaabla/garbage-classification. [Accessed: 08-Nov-2021].

[17] Tzutalin. LabelImg. Git code (2015). https://github.com/tzutalin/labelImg

[18] J. R. Uijlings, K. E. van de Sande, T. Gevers, and A. W. Smeulders, "Selective search for object recognition," International Journal of Computer Vision, vol. 104, no. 2, pp. 154–171, 2013.

[19] K. He, X. Zhang, S. Ren, en J. Sun, "Deep Residual Learning for Image Recognition", arXiv [cs.CV]. 2015.

[20] J. Cartucho, R. Ventura and M. Veloso, "Robust Object Recognition Through Symbiotic Deep Learning In Mobile Robots," 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2018, pp. 2336-2341, doi: 10.1109/IROS.2018.8594067.