

Greenlock: An approach to Automatic garbage detection framework designing using CNN

Vyom Yadav (Researcher)

vyom.yadav20@gmail.com

Department of Information Technology, Manipal university Jaipur, Jaipur, Rajasthan

Vasu Bajaj (Researcher)

vasubajaj09@gmail.com

Department of Information Technology, Manipal university Jaipur, Jaipur, Rajasthan

Utkarsh Dutt (Research Associate)

utkarsh.dutt@gmail.com

Department of Information Technology, Manipal university Jaipur, Jaipur, Rajasthan

Akhilesh Kumar Sharma (Mentor)

akhileshkumar.sharma@jaipur.manipal.edu

Department of Information Technology, Manipal university Jaipur, Jaipur, Rajasthan

January 9, 2020.

Abstract-This paper proposes a system for automatic detection of litter and garbage dumps in CCTV feeds with the help of deep learning implementations. The designed system named Greenlock scans and identifies entities that resemble an accumulation of garbage or a garbage dump in real time and alerts the respective authorities to deal with the issue by locating the point of origin. The entity is labelled as garbage if it passes a certain similarity threshold. ResNet-50 has been used for the training purpose alongside TensorFlow for mathematical operations for the neural network. Combined with a preexisting CCTV surveillance system, this system has the capability to hugely minimize garbage management costs via the prevention of formation of big dumps. The automatic detection also saves the manpower required in manual surveillance and contributes towards healthy neighborhoods and cleaner cities.

Keywords: Garbage Detection, Resnet, Tensorflow.

Introduction

India generates 62 million tons of waste every year, of which less than 60% is collected i.e. around 25 million tons of waste remains in the open and forever stagnating. The country is moving towards a technological revolution, work on smart cities and sanitation have moved in to the limelight. Still, the method of collection of waste products is ancient and inefficient. Due to mismanagement of waste in urban and rural areas, garbage dumps tend to pile up around corners of vicinities creating a source of bad odour and a breeding ground for pests and diseases. Often times, it is very difficult to find these accumulations of waste and eliminate them before they cause harm to a colony, neighborhood, locality or a city. The efficiency and effectiveness of pre-existing waste control measures needs to be improved to save on monetary resources utilized in these procedures and preserve the aesthetic value of cities. An even more pressing concern is the deterioration of healthy living conditions due to these accumulations. The idea was to create a product that uses the live feed of CCTV cameras on the streets to identify the areas where garbage has begun to accumulate. This will assist the garbage collection and waste control authorities like the

Nagar Nigam in cleaning the city in a less time consuming and focused fashion.

Urban garbage monitoring, currently, is done entirely with the help of human resources. In present times all major cities have CCTV cameras in place nowadays which have been put in for ensuring security, detecting and catching traffic violators and perpetrators of criminal activity. Progress has already started to be seen in the direction of automating the processing of these video feeds with computer vision instead of manual monitoring. In better developed cities in the country, automatic vehicular number plate capturing and image capturing have been devised to catch traffic offenders exceeding the speed limit without the efforts of an on-field traffic police officer. Similarly, this proposition is to viably use these camera feeds for detecting garbage dumps with the help of machine learning techniques based in computer vision. This idea is very feasible as it does not require changes in the existing infrastructure, only nominal additions to deal with the processing needs of the software to analyze the video feed, making it not only easier to implement but economically viable as well.

Large advancements have already been made in the field of using machine learning techniques using computer vision with the concepts of deep learning and neural networks (CNN). Frameworks based on these technologies are making automation easier as they make the need for human intervention redundant in day to day tasks. Neural networks are used to mimic the human mind to act in place of individuals and focus human resources on much more important tasks. Convolutional Neural Networks are used to classify images or parts of images into differentiable categories by detecting various features of the image like edges, colors and backgrounds to classify an object present in it.

Problem Statement

To provide a smart solution for garbage collection that analyses live video feeds and highlights garbage present in frames.

Literature

As per Huang, Kevin Murphy & wu et. al. [3], their work provided a model for selecting a detection architecture that achieves the right speed/memory/accuracy balance for a given application and platform. Various ways were investigated to trade accuracy for speed and memory usage in modern convolutional object detection systems

As per Anitya, Kumar & wu et. al. [4], their paper introduces a Quick Locale based Convolutional System strategy (Quick R-CNN) for question recognition. Quick R-CNN expands on past work to effectively characterize object recommendations utilizing profound convolutional systems. Quick R-CNN trains the VGG16 arrange nine times more efficiently than R-CNN, is 213 times faster at test-time, and accomplishes a higher Guide on PASCAL VOC 2012.

As per Nurminen, Jukka & wu et. al. [5], In their research work modern machine learning based approaches are used, in particular, the Yolo neural network system, to detect high-level objects, e.g. pumps or valves, in diagrams which can be scanned from paper archives or stored in vector or pixel form. In this concept, a simulator is used to automatically generate labeled training material to the system. A previously trained network is retrained to detect the components of interest.

As per Chen, Chunlin & wu et. al. [6], their research works to solve the problem of detecting objects, especially small objects, in complex scenes, the authors of this paper have proposed a novel module named as Adaptive Convolution Block (ACB), which adjusts the parameters of convolutional filters corresponding to the current feature maps and then filter these feature maps with the obtained convolutional filters to generate boosted features. Due to such adaptive convolution, the boosted features can pay more emphasis on the exercised objects, suppressing

the background information caused by irrelevant surroundings and improve the detection accuracy.

As per Huang, Lili & Wu et. al. [7], their research focuses on instruction-guided object detection, i.e., predicting the objects associated with the implementation of a specific instruction for intelligent robots. An amendment to the current detection paradigm is proposed by incorporating semantic instruction description effectively. To address the challenges related to picking out instruction-related objects from the detection results of a general object detector, a flexible dataset is introduced that can well adapt to the variation of the instruction set and only annotates instruction-related object samples. An

amendment to the current detection paradigm is proposed by incorporating semantic instruction description effectively.

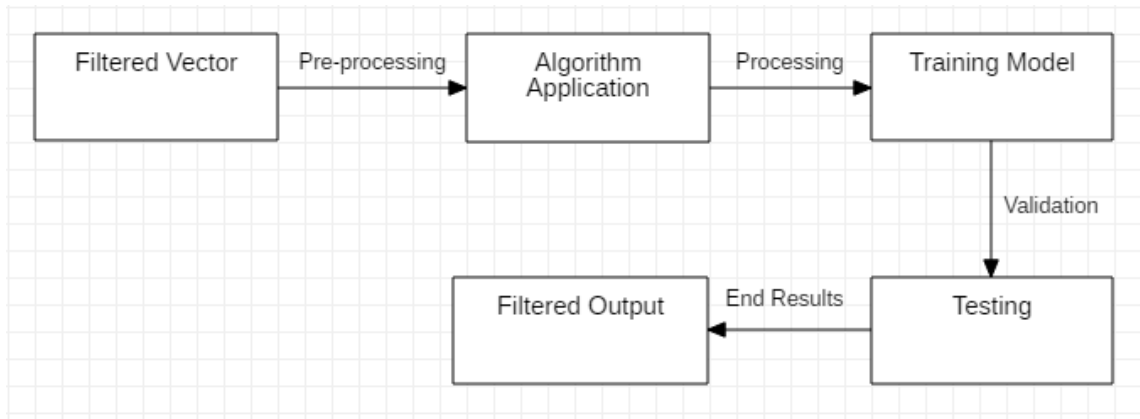


Figure 1: Flowchart depicting the methodology

CNN Algorithms	Confidence %
Standard Tensorflow Algo	75
Inception v1.0	85
Faster r-CNN	90
ResNet50	95

Table 1: CNN Algorithms Comparison Table

As per Kharinov, Mikhail & Wu et. al. [8], their paper focuses on the problem of automatized object detection in color images. The explication basing on the classic pixel clustering methods is discussed and advanced. The parameter for the heterogeneity of image areas is introduced. New methods for improving the quality of an image with automatically produced object names are suggested.

As per Liu, Ying & Wu et. al. [9], their paper proposes a garbage detection system based on deep learning and narrowband IOT concepts. The system automatically identifies garbage directly in embedded monitoring module, and manages thousands of monitoring front-ends via background server and narrow-band Internet of Things. The improved YOLOv2 network model is adopted to do garbage detection and recognition, in the front-end embedded module of the system

As per Huang, Jun & Wu et. al. [10], their research proposes a method for object detection by receiving a user input that specifies one or more first regions and one or more second regions in a template image. The other regions include furthermore objects of interest. The method further constitutes each of the one or more first regions recognizing a third region in an image under detection comparable to the first region in the template image by matching the image from the point of espial with the template image. The method further constitutes computing a transformation function based on the similarities from each of the one or more first regions to its comparable third region. The method further constitutes the computed transformation function to the one or more second regions to localize one or more fourth regions in the image for the object detection.

Methodology

1.1 Diagram:

Fig 1 shows the general progression of the research discussed in this paper.

Table 1 shows the comparisons between CNN Algorithms which were implemented until ResNet50 was finalized as the final framework.

1.2 Algorithm used:

Learning algorithm – Learning algorithm of ResNet is main back-propagation algorithm furthermore, continuous and online learning algorithm response is applied.

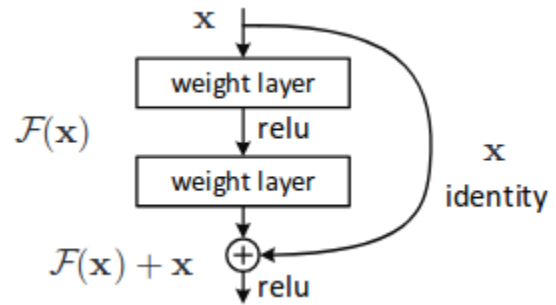


Figure 2: Residual learning

- The identity shortcuts (x) can be directly used when the input and output are of the same dimensions.

$$Y = F(X, \{W_i\}) + X$$

- When the dimensions change, A) The shortcut still performs identity mapping, with extra zero entries padded with the increased dimension. B) The projection shortcut is used to match the dimension (done by 1×1 conv) using the following formula

$$Y = F(X, \{W_i\}) + W_s X$$

- The first case adds no extra parameters, the second one adds in the form of $W_{\{s\}}$
- ResNet50 is 3 layers deep

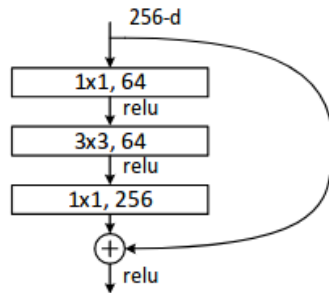


Figure 3: ResNet50 layer architecture

Evaluation Method:

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

In this formula, TP, FP, and FN are the number of true cases, false positive cases, and false negative cases respectively.

1.3 Methodology

The solution proposed in this research work boasts a minimalistic architecture thus does not have many technical requirements. Anyone can run the software to its maximum efficiency as it requires a live video source as its primary and only external component.

The research work began with intensive research into the necessary fields with the help of different research papers dealing with detection for custom

objects, and several blogs detailing similar projects ideas to find the most efficient technology stack to be used. Initially work started with Tensorflow to run our basic object detection code with its library of trained objects, but the confidence threshold of the resultant matched objects was never above 50%. To circumvent this issue a private python code was used which allowed for the use of ResNet50 (A convolutional neural network) model increasing the confidence percentage for more pixelated images. Running the code on these packages gave positive results.

Next moving forward the primary focus was on creating and training the main detection model with a custom dataset. The dataset is formed of 20000 images of various forms of garbage heaps and dumps found all over the world with plans to update the source code to implement a self-learning agent which learns from the outside environment and makes better predictions over time. A Nvidia graphics card along with tensorflow-gpu library were used for training. The trained dataset was then tested for image files. The threshold in the start was 85% but the restrains made the threshold percentage to go up to 95%. Finally, Custom designed python modules derived from the *raccon dataset by datitran* were used for running the webcam live feed from video source devices, using the OpenCV package of python. Multiple webcams feed can work in cohesion to create a community network to prevent illicit garbage collection throughout a whole smart community. Via the output feed, the target location will automatically get pinged on the software's hub

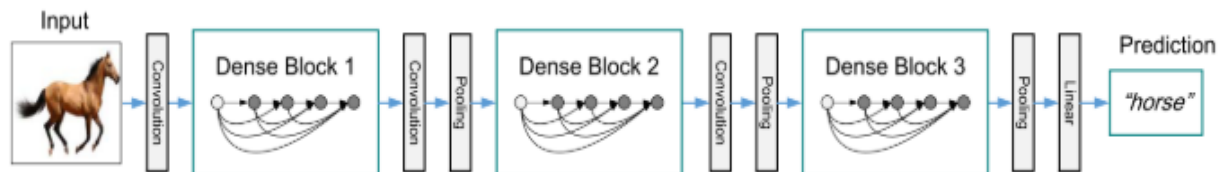


Figure 4: A Standard ResNet50 Network

portal and all active organizations can be informed.

The whole process was done in an Anaconda environment running Python 3.5 and Tensorflow 1.12.0.

Technology used –

ResNet-50 is a convolutional neural network that is trained on more than a hundred million images from the ImageNet database. This network is 50 layers deep and can classify images into 1000

object categories. As a result, ResNet-50 has learned rich feature representations for a wide range of images and has an image input size of 224-by-224.

TensorFlow is a free and open-source software library for dataflow and differentiable programming across a range of task and a symbolic math library. It is used for machine learning applications such as neural networks. Experience in TensorFlow is a standard expectation of the machine learning industry. Several python packages are being used here, such as SciPy, NumPy, OpenCV3, Pillow, Matplotlib and Keras.

Modules

Identification of correct Technology/Tools – The project started with a search for the correct tools required to materialize the project’s vision. Tensorflow was the first library to be selected for the training models followed by several python packages like OpenCV, Keras, Pillow.

Testing Models-Several models like Inceptionv3 and R-CNN were tried before settling for ResNet50 framework for the training of the module. ResNet50, when used with Tensorflow was able to provide faster training speed as well as better accuracy.

Collection of Garbage Datasets-Initially the project requirement included a large dataset consisting mainly of Indian/local street garbage and open dumps. The final dataset consists of images taken from SpotGarbage Dataset, some public datasets from Kaggle as well a large number of personal inclusions. The dataset contains over 15000 training images and over 5000 testing images.

Training the Dataset-To train a robust detection classifier the images in the dataset were individually labelled. The image .xml data, containing information such as borders and dimensions of objects in the images, was used to create .csv files containing all the data for the train and test images in tabular structures for labelling. With the images labeled, the TFRecords were generated that served as input data to the TensorFlow training model. Xml_to_csv.py and generate_tfrecord.py scripts from Dat Tran’s Raccoon Detector dataset were used with slight modifications to the original work to include our custom directory structure. Then a label map was created and editing of the training configuration file was done. Then the training code was executed and it took about 8 hours and 2 million steps for the model to be successfully trained.

Coding image/video modules & module for live feeds – The coding of the different modules was done in python programming language mainly utilizing the OpenCV package and the trained dataset in an anaconda environment running via a command line prompt.

Testing of the Dataset- The trained module can be utilized to detect garbage in an image, video or a live webcam feed.

Number of test pictures	Test Speed	Average Detection Accuracy	Detection Accuracy Variance
5000	0.71	85.07	0.089

Table 2: Summary of Test results

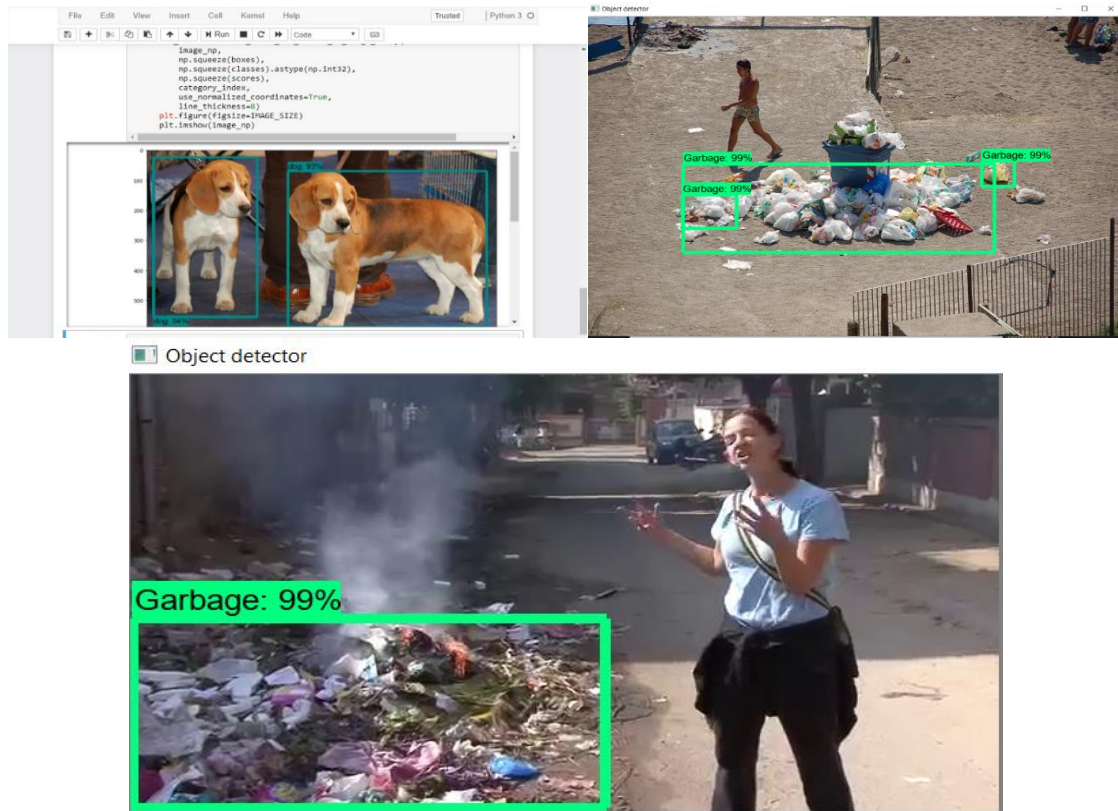


Figure 5: Prototype 1, 2 & 3

Prototype 1- The first prototype worked with object detection for all preexisting libraries of Tensorflow only on images. It can be seen in Fig 5.1 that both the objects are being detected as dogs along with a confidence percentage.

Prototype 2 -The second prototype worked with the designed custom garbage database running classifier on images, videos. It can be seen in Fig 5.2 that separate dumps of garbage are also shown as separate labels rather than just one.

Prototype 3- The third prototype worked with the designed custom garbage database running classifier on live feed. This is the finalized product, as it can be seen below in Fig 5.3, the garbage dump present in the frame running on a live footage can be segmented very precisely and is detected with a very high certainty level.

Graphs:

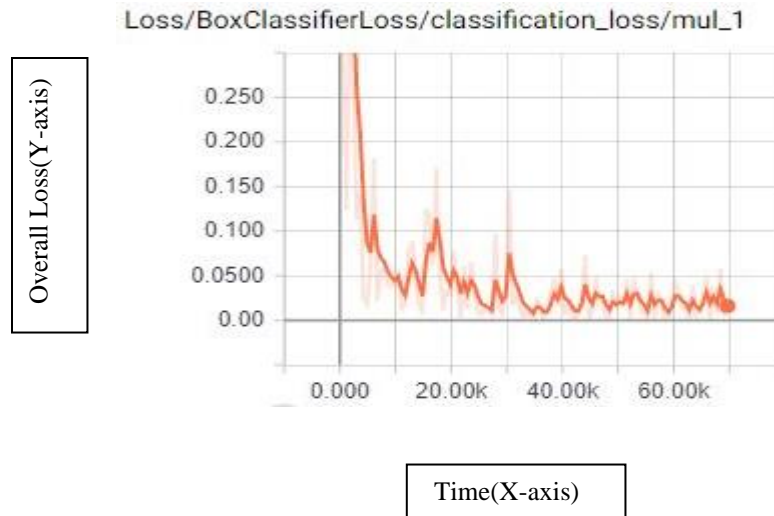


Figure 6: Total loss during training

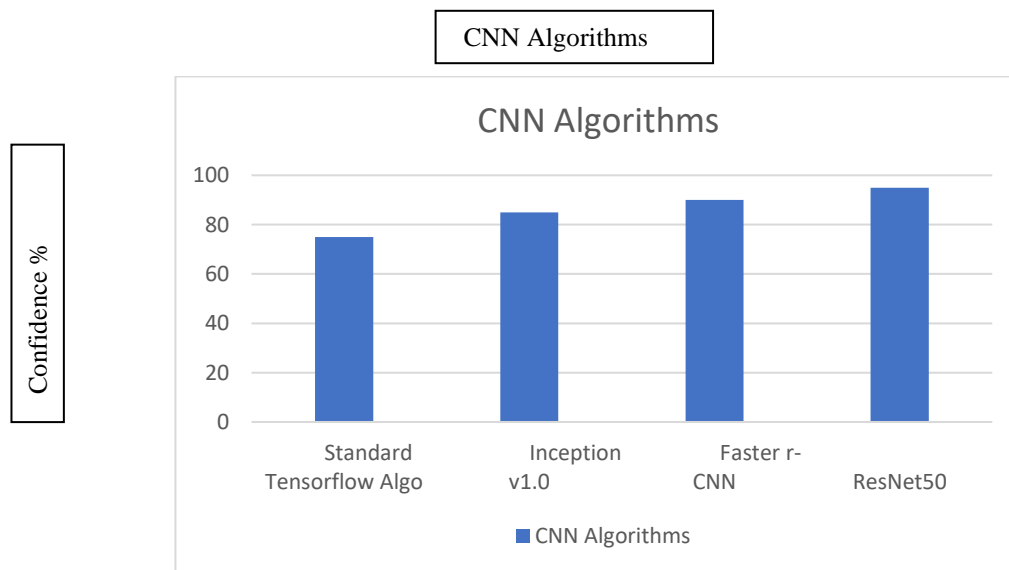


Figure 7: CNN Algorithms Comparison Chart

This paper proposes a live garbage detection product which can be used by private as well as public authorities alike. It can be developed further to provide a wide array of services like distinguishing between biodegradable, non-biodegradable and toxic wastes, classification of waste like glass, paper, polythene etc., even determining the best and most economical process for different types of waste segregated from the dumps.

Conclusion

This research work proposes a solution to the garbage collection problem, faced by many in the modern communities. It is an automated garbage detection solution which can be used by garbage collection authorities as well societies alike. It entails live video analysis (like CCTVs) for identification of open garbage dumps and

unofficial land dumps in streets, societies etc. For garbage detection several frameworks were utilized like the ResNet50 model and the Tensorflow technology to custom train a model. OpenCV3 along with other packages of python were implemented to run the live video feed. The implementation still needs to map every frame faster which can be done using better models and a larger dataset for training, so as to decrease the lag suffered during video testing. The dataset can also be implemented in a improved manner by specifying the dataset towards a state or city for better accuracy.

In the future versions of the project there are plans to implement-

- Classification of garbage into various categories like Biodegradable, Non-Biodegradable, Hazardous and Recyclable wastes.
- Train the model for a large dataset and differentiate waste into different classes like cardboard, glass, paper, plastic, tin/metals etc.
- Suggest the best methods to treat the various types garbage identified by the software.
- Integrate the machine learning model and deploy it on mobile and various IOT devices.
- Shift the model on a cloud-based system for an even faster detection and better accuracy.

This venture can help in cleaning the streets of the country also increase public awareness over the subject as well. Some quality of life improvements are still needed in the project, but with the lack of other similar service in the market it's a step forward in the right direction.

References:

- [1] Ying Liu and Zhishan Ge, "Research on automatic Garbage Detection System Based on Deep Learning and Narrowband Internet of Things", 2018.
- [2] Zhong-Qiu Zhao and Peng Zheng, "Object Detection with Deep Learning: A Review", 2017.
- [3] Huang, Jonathan & Rathod, Vivek & Sun, Chen & Zhu,, "Speed/Accuracy Trade-Offs for Modern Convolutional Object Detectors", 2017
- [4] Nurminen, Jukka & Rainio, Kari & Numminen, Jukka-Pekka, "Object Detection in Design Diagrams with Machine Learning.", 2018.
- [5] Anitya & Kumar, Akhilesh & Bhushan, Vinayak , "World of intelligence defense object detection "machine learning", 2018.
- [6] Chen, Chunlin & Ling, Qiang," Adaptive Convolution for Object Detection", 2019.
- [7] Huang, Lili & Wu, Hefeng & Li, Guanbin & Wang, Qing, "Instruction-guided object detection.", 2019.
- [8] Kharinov, Mikhail & Buslavsky, A, "Object Detection in Color Image", 2019.
- [9] Liu, Ying & Ge, Zhishan & Lv, Guoyun & Wang, Shikai., "Research on Automatic Garbage Detection System Based on Deep Learning and Narrowband Internet of Things.", 2018.
- [10] Huang, Jingchang & Yan, Chi & Zhu, Jun , "Object detection" , 2019.
- [11] Adrian Rosebrook, "Object detection with deep learning and OpenCV"
- [12] TensorFlow-Object-Detection-API, GitHub.com
- [13] Raccon Dataset, GitHub.com