The Eurasia Proceedings of Science, Technology,
Engineering & Mathematics (EPSTEM)
ISSN: 2602-3199

The Eurasia Proceedings of Science, Technology, Engineering & Mathematics (EPSTEM), 2023

Volume 24, Pages 55-62

IConTech 2023: International Conference on Technology

Detecting Litter in Street Sweepers Using Deep Learning

Suheda Gokbudak
Koluman Automotive Industry

Emir Enes Tas
Koluman Automotive Industry

Onur Ozer
Koluman Automotive Industry

Veysel Tilegi
Koluman Automotive Industry

Abstract: Street sweeping vehicles are essential equipment in our daily lives designed to clean streets and roads. With numerous mechanical components, they play a significant role in collecting all types of waste and contributing to environmental cleanliness. These vehicles typically consist of rotating brushes, collecting belts, and components involving water or air currents. Among these parts, brushes and vacuums are the most energy-consuming elements in street sweepers. Moreover, they are often operated in full power mode due to semi-automatic control systems, leaving the remaining control to the driver. However, this practice results in energy wastage and noise pollution. The aim of this study is to adjust vacuum suction in street sweepers according to the size of waste using image processing and deep learning techniques, thus achieving energy conservation. In this research, the YOLOv7 model and OpenCV are employed to train artificial intelligence for waste detection in street sweepers and accordingly regulate vacuum suction.

Keywords: Waste detection, Artificial intelligence, OpenCV, YOLOv7, Street sweeper.

Introduction

Road sweepers are typically vehicles used by city and municipal cleaning teams. These vehicles are designed to sweep and collect garbage in areas such as streets, roads, and sidewalks (Min et al., 2019). Road sweepers have a vacuum nozzle located at the lower part of the vehicle's center, and side brushes push the garbage toward the vacuum nozzle, allowing the garbage to be collected. However, when there is a large piece of debris, it needs to be manually adjusted for removal. In other words, the suction power of the vacuum needs to be adjusted according to the size of the debris. The constant adjustment of vacuum suction is the responsibility of the driver, which can distract the driver's attention while operating the vehicle. If these machines are constantly operated at full power, it leads to energy waste and noise pollution.

To address these issues, artificial intelligence and image processing can be used. The necessary equipment for such a system includes a camera, a PC board, and a GPU. The images captured by the camera are processed using artificial intelligence to detect the size of the garbage and adjust the vacuum suction accordingly. In this study, the YOLOv7 model, the latest version of the YOLO system, is used for real-time object detection and dimensioning. Training is conducted to teach the AI system to recognize garbage and adjust the vacuum suction, making the system autonomous. This not only leads to energy savings but also prevents the vacuum from operating at full power, reducing energy waste and noise pollution (Uzun & Karaca, 2022).

Related Studies

In the present day, there are various methods for waste detection, and studies related to waste detection using YOLOv7 have also been conducted. Uzun and Karaca (2022) conducted a study on deep learning-based waste detection for autonomous waste collection vehicles. In this study, 500 images of paper cups and 250 images of other types of waste were classified using deep learning algorithms, and a new dataset was proposed. To classify paper cups, pre-trained networks such as SquenzeNet, VGG-19, and GoogleNet were used, and it was concluded that SquenzeNet achieved the highest classification accuracy at 97.77%.

Donati et al. (2020) carried out a study on waste detection using deep learning to achieve energy savings in street sweepers. In this study, a known and simple neural network called U-Net was used for semantic segmentation. The U-Net architecture was modified and improved to work with less training data and provide more precise segmentation. Such networks are primarily used in defense and biomedical fields. The left side of the model represents the encoder, where the model learns what the image contains, and the right side represents the decoder, where the model learns where the image is located. These studies demonstrate the effective utilization of artificial intelligence and deep learning techniques in waste detection and waste management.

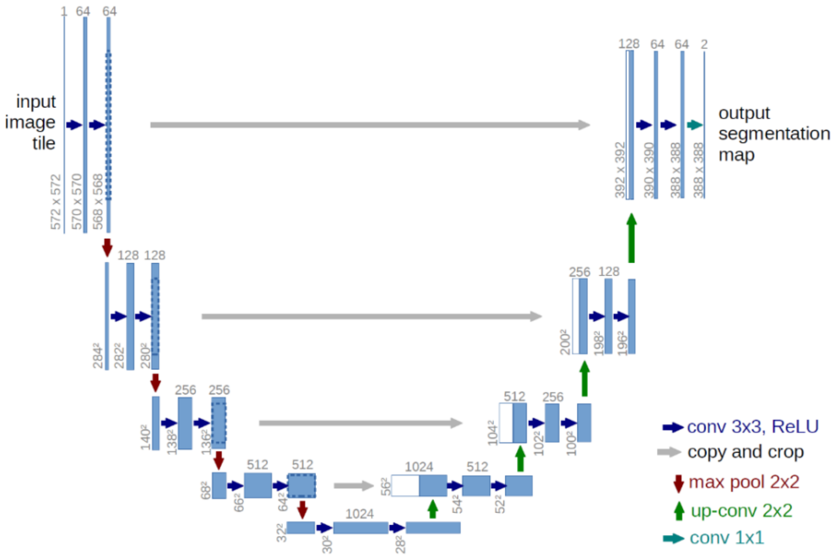

Figure 1.U-Net architecture

Normally, the U-Net architecture is not real-time. However, a downsized version of U-Net, known as U-Net 2020, is capable of pixel-wise segmentation of a 1-megapixel image on a mobile GPU in less than one second, making it suitable for real-time use. Furthermore, the study selected the DALSA Genie Nano C1920 camera model with a resolution of 1920x1200 for image processing. For the computer used to process the images, an Nvidia Jetson TX2 was chosen. In the experimental study conducted, a dataset comprising approximately 400 images and 80 test images was used, and the system demonstrated a sensitivity of 94%.

TABLE I

TRACKING RESULTS ON THE MOT16 CHALLENGE. WE COMPARE THE TRACKING PERFORMANCE OF YOLOV7-DEEPSORT AND YOLOV5(S/M/L)-DEEPSORT.

Model	YOLOv5s	YOLOv5m	YOLOv5l	YOLOv7
MOTA	39.60	39.01	40.77	40.82
MOTP	80.85	81.87	81.96	82.01
IDF1	52.39	51.56	52.43	53.65
IDs	432	432	547	514
ML	39.65%	33.27%	31.92%	32.11%
MT	15.45%	17.41%	20.70%	20.12%
FP	5375	7612	7853	7940
FN	60882	59297	56990	57434

Figure 2. Experimental results Yang (2022)

In the article by Yang. (2022), it was observed that incorporating DeepSort into the YOLOv7 method has resulted in more successful multi-object tracking in target tracking. This was achieved by conducting comparisons with YOLOv5 and its derivatives, leading to the obtained results.

Min et al. (2019) employed the Faster R-CNN model to classify waste types. This novel approach for object detection is built upon the Fast R-CNN method, with the addition of a Region Proposal Network (RPN) responsible for generating region proposals

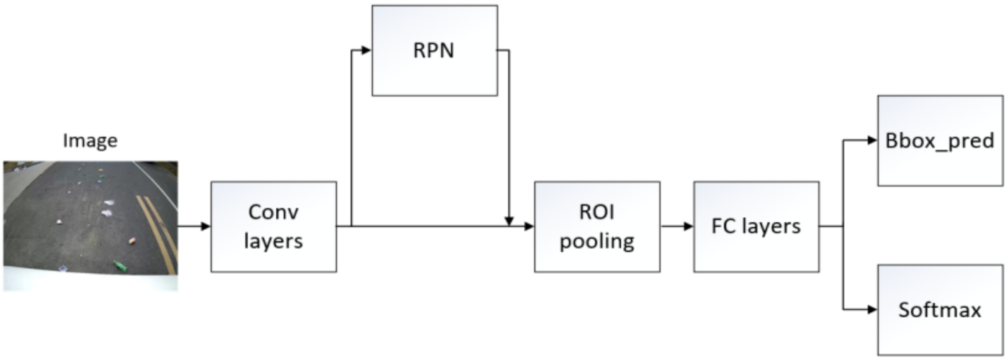

Figure 3. Faster-RCNN flow chart

In the article by Zheng et al. (2022), a segmentation model for waste detection in road sweepers is described. The model utilizes an encoder-decoder structure similar to SegNet. In this approach, feature extraction is applied to input images by the convolutional blocks in the encoder part. Simultaneously, pixel-wise segmentation is generated by gradually recovering the spatial dimensions of feature maps in the decoder part. In addition to the convolutional blocks, the model adopts two modules in the decoder structure, namely, FPG and SG, along with down-sampling and up-sampling blocks.

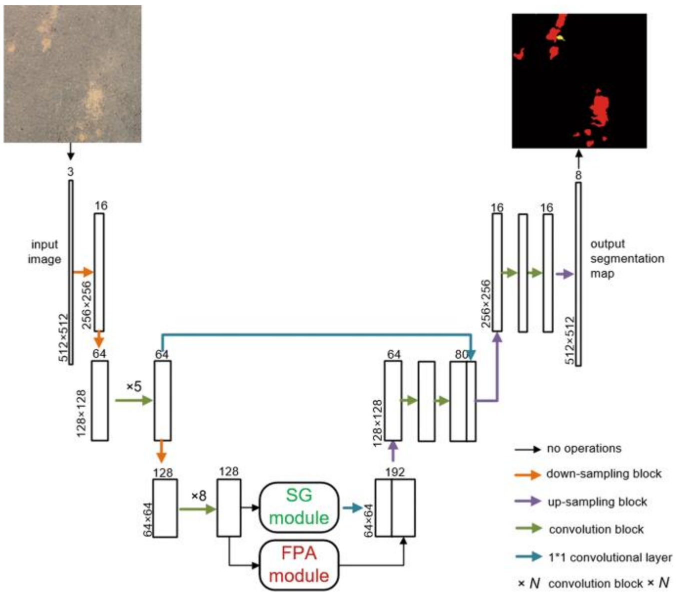

Figure 4. Overall framework of the proposed method

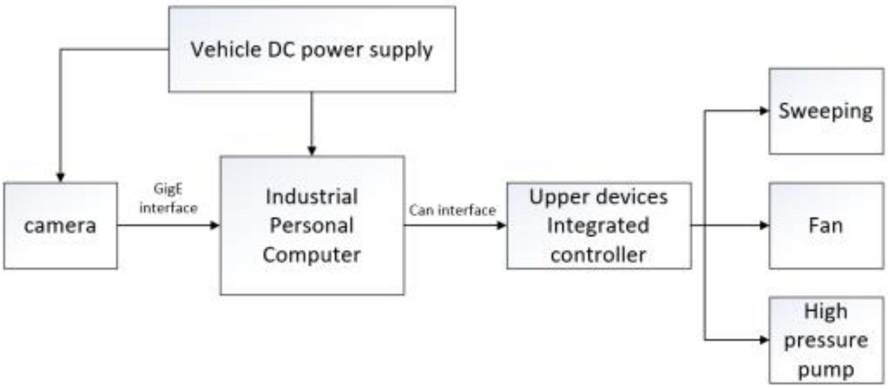

Figure 5. The power control system of intelligent road sweeper operation

In the article by Min et al. (2019c), a power control system is proposed for road sweeper operation, and the design of its fundamental technologies is addressed. Visual recognition and detection algorithms have been developed to recognize essential features such as waste type and quantity. An intelligent power control system is created by integrating an industrial computer, Nuvo-5095GC, a vehicle DC power supply, and a cleaning device controller. The camera is connected to the industrial computer via a GigE interface. The vehicle's DC power supply provides power to both the vehicle's industrial computer and the camera. The in-vehicle industrial computer is connected to the integrated controller through a CAN interface. The cleaning unit's integrated controller is connected to the sweeping disk, fan, and high-pressure water pump via a digital I/O interface.

Applied Methods

YOLO (You Only Look Once)

YOLO, which stands for "You Only Look Once," is a widely used algorithm in object tracking. The YOLO algorithm was introduced by Joseph Redmon (2015).(Nasrullah & Diker, 2021). It is an extremely fast object detection algorithm known for its capabilities in object recognition and detection. The YOLO model can process an image of 416x416 pixels in just 22 milliseconds, making it an ideal algorithm for real-time applications due to its speed. Furthermore, YOLO can detect a large number of objects and classes simultaneously, making it a highly versatile approach for object classification and detection. When evaluated in terms of accuracy and speed, it is much faster than the Single Shot Detection (SSD) method and provides similar accuracy (Ozel et al., 2021). Additionally, YOLO's network architecture is more effective than other methods and is easily scalable, allowing for a balance between speed and accuracy by adjusting the model's size. Its most superior feature compared to other algorithms is that it does not require retraining for artificial intelligence.

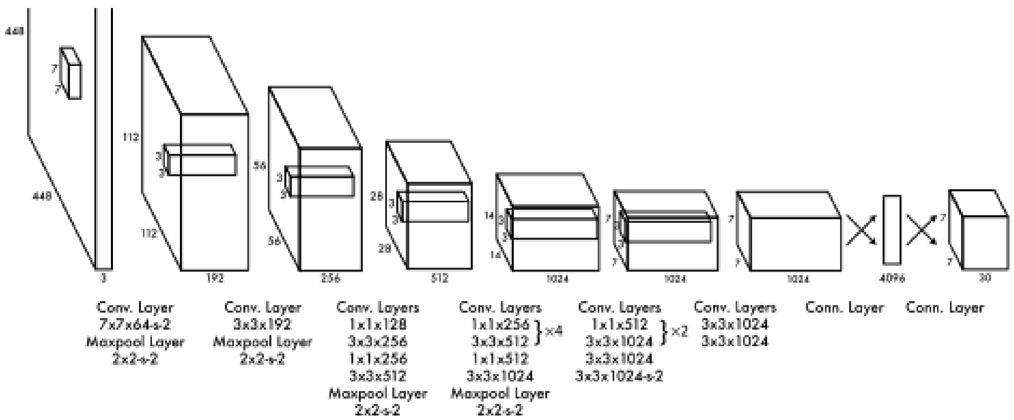

Figure 6. YOLO architecture

Equation (1) provides the mathematical model of the YOLO algorithm.

$$\lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} \mathbb{1}_{ij}^{\text{obj}} \left[(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 \right]$$
$$+ \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} \mathbb{1}_{ij}^{\text{obj}} \left[\left(\sqrt{w_i} - \sqrt{\hat{w}_i} \right)^2 + \left(\sqrt{h_i} - \sqrt{\hat{h}_i} \right)^2 \right]$$
$$+ \sum_{i=0}^{S^2} \sum_{j=0}^{B} \mathbb{1}_{ij}^{\text{obj}} \left(C_i - \hat{C}_i \right)^2$$
$$+ \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} \mathbb{1}_{ij}^{\text{noobj}} \left(C_i - \hat{C}_i \right)^2$$
$$+ \sum_{i=0}^{S^2} \mathbb{1}_{i}^{\text{obj}} \sum_{c \in \text{classes}} (p_i(c) - \hat{p}_i(c))^2 \quad (1)$$

YOLO uses certain terms to determine the bounding box of detected objects:

Bx: x-coordinate of the center point of the object.
By: y-coordinate of the center point of the object.
Bw: Width of the object.
Bh: Height of the object.
c: Represents the class of the object.

The score resulting from the calculation of Equation (1) indicates whether an object is detected within a valid region.

YOLOv7

YOLOv7 (You Only Look Once version 7) is a prominent algorithm that stands out as a fast and accurate real-time object detection model for computer vision tasks. This model was introduced as an official research paper published by Wang, et al. (2022) (Wang, 2023). Furthermore, the source code of YOLOv7 has been released as open-source software and is available for free under the GPL-3.0 license. This allows developers and researchers to use YOLOv7 in their own projects, providing them with the opportunity to enhance and contribute to the YOLOv7 name and their personal expertise (Yang, 2022).

Data Set

For training purposes, a dataset was prepared using existing datasets and Google Images. This dataset was organized in such a way that each waste group, such as biological waste, metal, plastic, and paper, had separate files. This was done with the aim of ensuring more accurate inferences. After creating the dataset, the data labeling process needs to be carried out. The performance of a deep learning model depends on how well the data is labeled. After the data labeling process, the resulting TXT files are added to the dataset.

Figure 6. Images from the prepared data set

Data augmentation is a technique employed to artificially increase the number of training examples (Shorten et al., 2021). This is commonly carried out on the existing input dataset either before or during the training process. Essentially, it involves applying transformations to the original data without altering the original images to generate new images that may be different but belong to the same classes as the original data. This is a fundamental step in addressing scalability issues, particularly in fully supervised datasets. The original dataset is utilized to create additional variations and construct a larger training dataset by automatically rotating, zooming, and flipping the images with random parameters (Perez, 2017).

Real-Time Vehicle Control

After the completion of dataset preparation, labeling, and data augmentation processes, the final stage is to ensure that the sweeper vacuums work effectively with the model. This is achieved by being able to control the actuators of the sweeper via CanBus. The sweeper vehicle in consideration has a vacuum on the left and right sides, each equipped with brushes that can collect debris when lifted.

In this design, one critical aspect is the y-coordinates of the debris (the distance of debris from the vehicle). The reason for its significance is that it directly affects the delay before activating the vacuum based on this distance. Assuming the vehicle moves in a straight line, if we frame the debris at a y-coordinate between d_{near} and d_{far} at time t_0, the brushes and vacuums should operate at time $t_1 = \frac{v_{avg}}{y}$

Here, v_{avg} represents the average vehicle speed. However, this situation can be risky when the vehicle speed is high (Donati et al., 2020).

System Design

Camera

The camera system constitutes one of the most critical components of the system. The camera to be used should have certain features, including low power consumption, adequate resolution (for accurate debris detection), fast and high-bandwidth data communication with the computer, and resistance to vibration and electromagnetic fields. To find a camera that possesses these features, it is necessary to conduct product research on cameras designed for industrial applications. The DALSA Genie Nano C1920 (@1920×1200) camera could be a suitable choice for this system, but there are many options available as well.

Figure 7. DALSA genie nano C1920 (@1920×1200)

Computer Selection

In computer selection, another essential feature to consider is low power consumption. Additionally, having a fast GPU, a CanBus interface, and resistance to vibrations are critical factors in the selection process. Based on these features, the Nvidia Jetson TX2 model has been chosen.

Figure 8. Nvidia Jetson TX2

Software

In this system, the YOLOv7 algorithm, commonly used in image processing, has been employed. The operational flowchart of the intelligent control system is provided in Figure 1.1 below. Furthermore, artificial intelligence training has been carried out using a dataset to achieve the automation of debris detection.

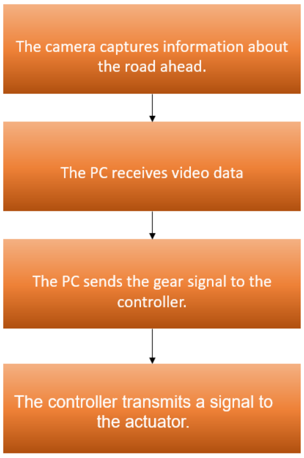

Figure 9. Intelligent control process

Conclusion

The developed system aims to achieve autonomous debris detection in road sweepers through image processing and artificial intelligence, while maintaining cost-effectiveness. Although it remains theoretical at this stage, accuracy will be substantiated through the necessary experiments in the future.

Scientific Ethics Declaration

The authors declare that the scientific ethical and legal responsibility of this article published in EPSTEM. journal belongs to the authors.

Acknowledgements or Notes

* This article was presented as an oral presentation at the International Conference on Technology (www.icontechno.net) held in Antalya/Turkey on November 16-19, 2023.

References

Donati, L. L., Fontanini, T., Tagliaferri, F., & Prati, A. (2020). An energy saving road sweeper using deep vision for garbage detection. *Applied Sciences, 10*(22), 8146.

Min, H., Zhu, X., Yan, B., & Yu, Y. (2019). Research on Visual Algorithm of road garbage based on intelligent control of road sweeper. *Journal of Physics*, *1302*(3)

Nasrullah, M., & Diker, A. (2021). Derin ogrenme ve YOLO algoritmaları ile nesne tespiti ve serit belirleme. *2 nd International Conference on Intelligent Transportation Systems*. Balıkesir, Turkey.

Ozel, M. A., Baysal, S. S., & Sahin, M. E. (2021). Derin ogrenme algoritması (YOLO) ile dinamik test suresince suspansiyon parcalarında catlak tespiti. *Europan Journal of Science and Technology*, *26*(1-5).

Perez, L. (2017, December 13). The effectiveness of data augmentation in image classification using deep learning. Retrieved from https://arxiv.org/abs/1712.04621

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, *6*(1).

Shorten, C., Khoshgoftaar, T. M., & Furht, B. (2021). Text data augmentation for deep learning. *Journal of Big Data*, *8*(1), 101.

Uzun, S., & Karaca, D. (2022). Deep learning based garbage detection for autonomous garbage collection vehicles. *Europan Journal of Science and Technology,32,* 1194-1198.

Wang, C. (2023). YOLOV7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *Computer Vision Foundation*, 1-12.

Yang, F. (2022). Video object tracking based on YOLOv7 and DeepSORT.

Zheng, C., Cao, D., & Hu, C. (2022). A similarity-guided segmentation model for garbage detection under road scene. *Frontiers of Optoelectronics, 15*(1), 22.

Author Information

Suheda Gokbudak
Koluman Automotive Industry
Mersin, Turkey
Contact e-mail: suheda.gokbudak@koluman.com

Emir Enes Tas
Koluman Automotive Industry
Mersin, Turkey

Onur Ozer
Koluman Automotive Industry
Mersin, Turkey

Veysel Tilegi
Koluman Automotive Industry
Mersin, Turkey

To cite this article:

Gokbudak, S., Tas, E.E., Ozer, O., & Tilegi, V. (2023). Detecting litter in street sweepers using deep learning. *The Eurasia Proceedings of Science, Technology, Engineering & Mathematics (EPSTEM), 24,* 55-62.