The 10th International Workshop on Simulation for Energy, Sustainable Development & Environment (SESDE), 012
19th International Multidisciplinary Modeling & Simulation Multiconference

2724-0061 © 2022 The Authors.
doi: 10.46354/i3m.2022.sesde.012

UGV Digital Twin for Supporting Innovative Environmental Treatments of Large Areas by using AI

Antonio Giovannetti[1,2,*], Mehrnoosh Mashayekhizadeh[2], Daniele Cefaliello[2], Farshad Shamlu[2], Hinda Taib[2]

[1]University of Genoa, Via Opera Pia 15, Genoa, 16145, Italy
[2]Simulation Team, Via Cadorna 2, Savona, 17100, Italy
[*]Corresponding author. Email address: antonio.giovannetti@simulationteam.com

Abstract

The efforts to maintain low level of pollution in the environment need to meet sustainable and innovative solutions in order to be effective in the modern and complex world. Indeed, currently we are facing critical issues in terms of sustainability and new technologies are crucial to identify new solutions: recycling the abandoned wastes is a critical issue that needs to be considered and implemented beside their collection and processing. Due to these reasons, the purpose of this research is to develop an intelligent cleaning complex system based on UGV (Unmanned Ground Vehicle), which focus is to enhance the ability of strategic decisions making and achieve satisfactory trash detection in the wild for deployment. The system stands on AI (Artificial Intelligence) based on deep neural networks, which behaviour is simulated in a 3D Unity synthetic environment, through the usage of a digital twin, that simulates the desired solution of detecting and collecting garbage. The developed virtual environment chosen by the authors is a contaminated beach, in which an armed and autonomous robot manages to collect and recycle the wastes; the scenario is flexible to changes and further ones can be developed for training and testing of the intelligent system.

Keywords: Digital Twin, Artificial Intelligence (AI), Modelign & Simulation (M&S), Sustainability, Visual Analytics

1. Introduction

Cleaning up contaminated ecosystems is one part of sustainable environmental engineering. Although a renewable design includes minimum waste, usually this is not the case. Moreover, the recycling process is a crucial step, and still it is not applied to the tons of glass, metal, paper and in particular plastic (L. Parker, 2018) garbage that is left out in the environment.
Collecting garbage from the environment is tiresome work and needs to be done routinely, however it is not faced properly enough. In addition, environmental contaminants themselves could be relevantly dangerous to humans affecting health and causing diseases (U. S. Environmental Protecting Agency (EPA), 2020), since polluted environments are a perfect habitat for bacteria and their spread through air and animals, which have become a major problem considering the overpopulated world and the growth of the Coastal Population and Towns; in facts on sea shores near to industrial areas and large towns, there is a very big impact due to port, industrial and touristic activities that create serious environmental problems; this is hard to be addressed in sustainable way from economic, environmental and social point of view. Within this context, it finds place the following research of developing an intelligent cleaning UGV for improving sustainability and quality of the environment.

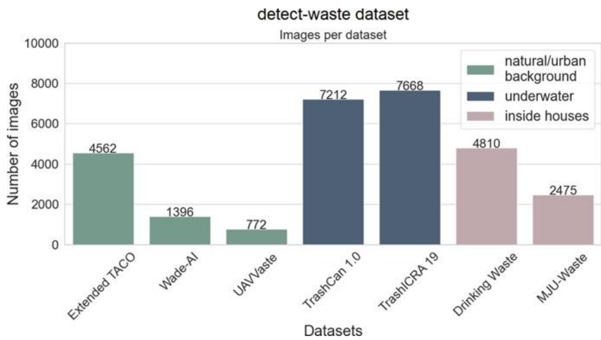

Figure 1. Public benchmark for waste detection datasets

The system focuses on two main aspects, that will be deepen in the course of the document: the Intelligence Core based on AI (Artificial Intelligence) that should be able to detect and classify the garbage within complex environments through vision segmentation and detection for grasping and placement of the wastes in the appropriate recycling container, as well as the development of a Virtual Framework able to simulate the use of the new autonomous systems to validate them and to test their capabilities, thanks to a digital twin, under different employable and flexible conditions through the use of simulation itself. In this paper it is proposed an innovative approach devoted to address wide areas on sea shores where, both industrial, port and tourists activities are impacting; obviously this approach could be effectively used also in other scenarios with proper training.

2. State of the Art

In order to address the proposed problem, it was decided to develop a solution based on innovative light and inexpensive UGVs (Unmanned Guided Vehicle) devoted to serve the area.

The focus of this paper is on the specific UGV solution and especially its intelligence for making it able to operate autonomously on the field. Obviously to address this problem it was necessary to investigate AI solutions that could be implemented on a in innovative way.

So, the paper focuses on visual analytics to be used to identify the garbage and the navigation intelligence to move on the area and find it avoiding problems with obstacles and people.

In this paper, two state-of-the-art deep neural network architectures have been tested, YOLOv5 and DeepLabv3+, considered to be the intelligent vision recognition component of a cleaning and recycling robot, further endowed with an autonomous navigation system based on the A* pathfinding algorithm, simulated through a digital twin in a 3D virtual environment. YOLOv5 is a family of object detection architectures capable of, and commonly used as well for, semantic segmentation (Jocher et al., 2020). However, as the founder of the model reported in the GitHub community and it will be shown as well in the following, that the efficiency of this latter process, in terms of precision, recall and f-1 score, is lower if compared to other architectures specifically designed for it, such as DeepLabv3+, considered the most performing for this purpose (Chen et al., 2018). Nevertheless, similar researches have been conducted for the garbage detection and semantic segmentation through YOLOv5 providing promising results in terms of the efficiency on the litter recognition (Liu and Li, 2021; Wu et al., 2021), as well as a research for the application of such architecture on a fixed robotic arm aiming for recycling, that has shown interesting outputs in terms of potentiality (Lei et al., 2021).

This study aims to extend and improve these results, providing a more efficient image segmentation thanks to the further deep learning architecture named DeepLabv3+. Furthermore, TACO dataset, for the garbage labelled images, is considered in this study for a further train and test, being reported as the most complete one in the outdoor background, as shown in Figure 1 (Majchrowska et al., 2022).

In addition, a digital twin simulated in a virtual scenario is considered to validate and test the outputs in order to enhance the decision-making process and the flexibility of the employed solutions. The use of simulation for training new Robotic Systems and related AI is proven to be extremely effective both in terms of costs and in terms of its flexibility for different uses and applications (Bruzzone et al., 2016; 2019b). The complexity of this project requires the merged utilization of heterogenous components, deepen in the next paragraph, in order to achieve its purpose. Indeed, there is a variety of tools engineered to excel in some specific tasks, such as simulation or machine vision.

3. Materials and Method

The developed cleaning system is split into two main components: the Intelligence Core and the Virtual Framework, based on different methods that will be explored in the following.

3.1. Intelligence Core

The first component of the intelligent cleaning system, the Intelligent Core, relies upon three modules, each guaranteeing a step for the needed behaviour: object detection, object tracking and semantic segmentation. The object detection uses the YOLOv5 model, a Deep Neural Network based on PyTorch framework, pre-trained on the well-known COCO dataset and successively fine-tuned by the authors on the TACO one, which includes a considerable number of labelled wastes images.

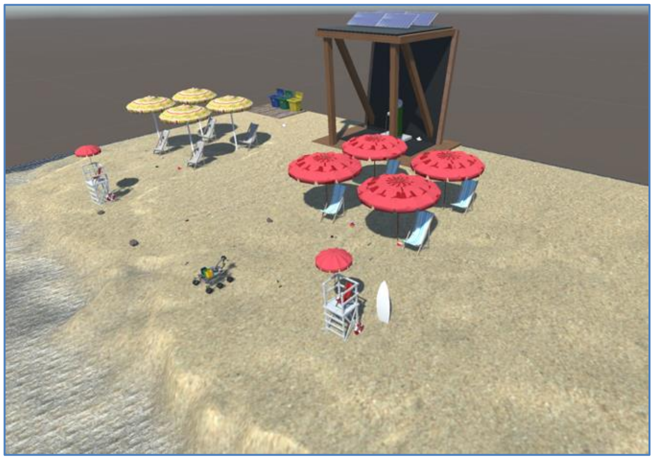

Figure 2. A snapshot of the contaminated coastline environment developed on Unity3D, including the recycling area and a power station for the UGV's batteries recharge.

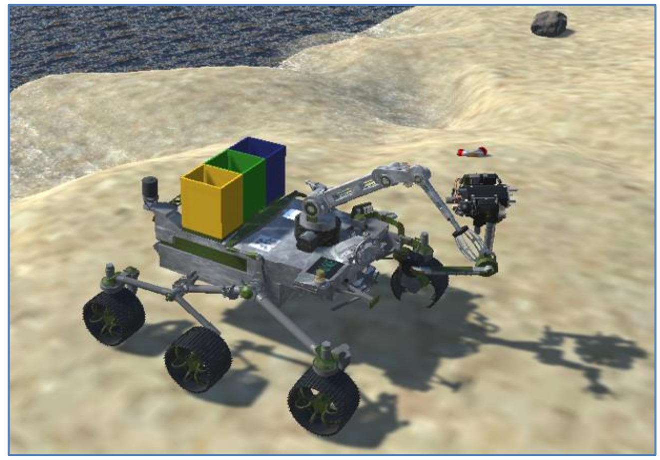

Figure 3. A magnification of the cleaning robot 3D model, reworked from the Perseverance rover template as demonstrative purpose.

Concerning the object tracking, an A* algorithm is considered for the shortest path within the environment-network consisting in the garbage and also obstacles (non-garbage), which real-time focus and collision avoidance needs to be considered thanks to the continuous research path method. A* algorithm can be considered an extension of the well-known Dijkstra's, but it includes also forward-looking component, which is an estimate of the length to complete the path to the destination from a specific node (Zeng and Church, 2007), for this reason it has been selected.

The semantic segmentation, that determines the garbage perimeter and class (waste type), is conducted by the DeepLabv3+ Deep Neural Network that, thanks to its decoder module, enhance the segmentation results; as the first step, this algorithm has been pre-trained on the COCO dataset and then fine-tuned by the authors on TACO.

3.2. Virtual Framework

Concerning the Virtual Framework, it has been developed by the authors within the Unity3D platform. It consists of a detailed and contaminated coastline environment with randomly generated garbage and obstacles such as the elements of a common bathing beach and rocks (Figure 2). The choice of the selected environment is demonstrative and it is meant to be flexibly modified for desired tests. Pictures of the wastes belonging to, and delved into, this environment have been used for training and testing the AI vision software, which results will be shown in the next paragraph, while the usage of a 3D model for the armed UGV is used as a digital twin to simulate the intelligent cleaning system (Figure 3), which movement behaviour is coded in C# language. The configuration of the UGV in this case is used as example therefore more basic solutions could be adopted to guarantee low costs and flexibility in these scenarios.

3.3. General Architecture

Advanced tools, such as simulation and machine vision, are often not directly interoperable. PyTorch and Keras frameworks are optimized for Python, while a simulation environment, such as Unity, requires the logic written in C#. In such case, a problem of software integration or even of its interoperability arises. Based on specific combination of applications and chosen general architecture, various approaches could be used to address the issue of making system parts interact one with another. For instance, it could be done by using High Level Architecture (HLA), which is the standard for interoperable simulation (Bruzzone et al. 2014), by implementation of RESTful APIs, (Representational State Transfer, Application Programming Interface), connection through WebSocket, RPC (Remote Procedure Call) or even utilization of Blockchain (Bruzzone et al. 2020; 2019a). In the case of interest, the main challenge is indeed the data exchange between Machine Vision algorithms in a dedicated Python framework and the simulation in Unity environment. Furthermore, the target software architecture must be capable to work freely not only with the virtual environment, but also potentially with physical Unmanned Vehicles. Considering this, it is decided to handle video transfer from simulation or camera to the Machine Vision application by means of MJPEG frames and network socket. In this case, it is relatively easy to obtain screencast from the virtual environment, while the required codecs are also supported by variety of hardware cameras. Vice versa, the Python application returns definition and spatial state of recognized targets. The basic architecture is summarized in Figure 4. It is evident that the Machine Vision Application could be subjected to multiple phases for progressive elaboration including preprocessing related to light and contrast levels especially considering that external environment could be subjected to quite different conditions.

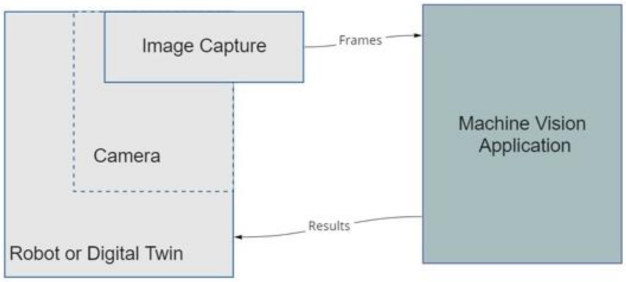

Figure 4. Modules interoperability architecture

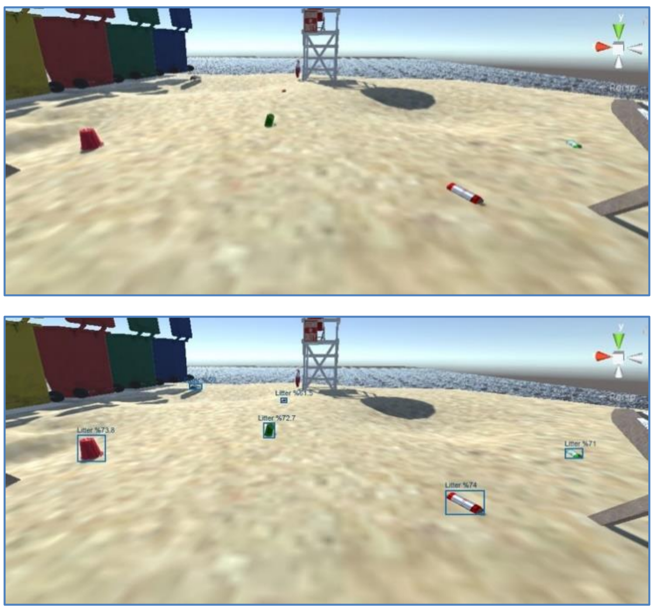

Figure 5. The picture of generic litter 3D-objects belonging to the Virtual Framework (above) properly detected by the YOLOv5 object detection model (below)

4. Results and Discussion

In the following paragraphs are reported the results, in terms of performance evaluation parameters, concerning the Intelligence Core, in particular for the detection and segmentation processes, by YOLOv5 and DeepLabv3+ models respectively.

The mentioned Deep Neural Networks have been tested on the TACO dataset, furthermore snapshots of the 3D litter objects belonging to the Virtual Framework have been used to simulate and test the behaviour of the digital twin.

4.1. Object Detection evaluation

The success of the analyzed deep neural network approach for the object detection has been evaluated in terms of Average Precision at a 50% threshold (AP50). The metrics corresponds to the average precision (AP) and average recall (AR), considering the average of 10 precision-recall pairs computed through altering the IoU (Intersection over Union, that is the intersection of the prediction and ground-truth regions of the c-th class over the union of the total) value from 50% to 95%, at steps of 5%. In waste object detection, the use of AP (Average Precision) in measuring accuracy is quite popular.

Average precision defines the average precision value for recall value over 0 to 1. Precision is the ratio of true positives (TP) to the cases that are predicted as positive i.e., true positive and false positive (FP).

In the following, it is possible to analyze the Table 1 that proposes the effectiveness of the deep neural networks evaluated over the the Taco dataset used for the experimentation.

Table 1. Summary of the Average Precision at 50% (AP50), Average Recall (AR) and f-1 measure for the detection of the most remarkable classes of garbage within the TACO test dataset, by YOLOv5

Class	AP50	AR	F1
Garbage	0.85	0.82	0.76
Non-Garbage	0.79	0.76	0.68
Background	0.73	0.75	0.73
Glass	**0.65**	**0.56**	**0.53**
Metal	**0.73**	**0.66**	**0.67**

As it can be seen, YOLOv5 reached promising results on this dataset, with 85% of AP50 on the garbage detection. In the Virtual Framework the litter is well recognized with a discrete and decreasing accuracy with the increase of the distance from the camera, as shown in Figure 5, that simulates the behaviour of the robot image vision.

4.2. Image Segmentation evaluation

The image segmentation performance is evaluated in terms of precision and recall, referred as the positive predictive value and the true positive ratio respectively, as well as the and f-1 score, that is the harmonic mean of the latters. The results for the different 7 classes included in TACO dataset are shown in Table 2, in which a heterogenous performance is obtained for the different classes, among which glass and metals&plastic appear to be the most well recognized, result that is supported in the Virtual Framework as shown in Figure 6, that provides an example of classification for a metal can delved into the sand.

Table 2. Summary of precision, recall and f-1 score for the segmentation (classification) of the most remarkable classes of garbage within the TACO test dataset, by DeepLabv3+

Class	Precision	Recall	F-1 Score
Background	0.97	0.97	0.97
Paper	0.62	0.76	0.68
Unknown litter	0.52	0.65	0.58
Glass	**0.83**	**0.82**	**0.83**
Metals & plastic	**0.87**	**0.70**	**0.78**
Bio	0.62	0.51	0.56
Non-Recyclable	0.52	0.65	0.58

Figure 6 The picture of a metal can 3D-object belonging to the Virtual Framework (left) properly classified by the DeepLabv3+ image segmentation software (right), that has successfully recognized the litter class with an accuracy of 81%

The high recall rating for a specific litter class indicates that it was infrequently confused with any of the others, i.e. correctly distinguished. An important comparison needs to be done between the values related to Glass and Metals, here reported, with the respective values inside Table 1, considerably lower, to show the inefficiency of YOLOv5 in segmentation process. In addition, the classes of non-recyclable, bio, and unknown litter caused the most confusion. This is likely caused by the partially decomposition that affects these types of garbage. It is important to note that the background was successfully separated from the remaining garbage with a precision and recall rate of 97% for both of them. The performance was enhanced and, as would be expected, the number of false positives was decreased by adding a distinct class for background. A further information can be obtained by the Intersection over Union (IoU) parameter. The values reported for the segmentation of the test set of TACO are shown in Figure 7, enhancing the ability to resolutely recognize the garbage within a discrete size, implying a poor classification, i.e. a probable confusion, for small and big garbage.

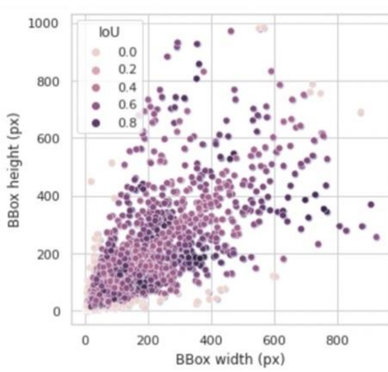

Figure 7 IOU reported values by DeepLabv3+ for the segmentation process tested on TACO dataset

5. Conclusions

The present research has analyzed the case of an innovative solution devoted to address environmental treatment of wide areas with special attention to seashores affected by industrial, port and touristic impacts. Within this case study it was studied the development of a simple low cost UGV able to act as operator within this innovative system and to serve as autonomous garbage collection robotic system. Therefore, it is evident that, this is just a preliminary part of this research for sustainable solutions devoted to clean and recycle garbage over complex environments such as a coastline. This robotic system therefore in order to convenient and sustainable should be able to provide reliable and efficient services that are relying on the AI solutions adopted to direct its actions and this paper focuses on these specific aspects by creating a digital twin to be used to test and validate its design. Indeed, this study has allowed designing, validating, and testing the UGV capabilities of performing the activities of detection, localization and classification for an automatic motion control, which simulation in a closed virtual environment enables further extensions and possibilities for testing under different conditions, that can lead to different monitorable behaviours. Furthermore, it has been demonstrated how the TACO dataset and tools may be effectively utilized to detect trash, providing efficient performances in the development of the AI. Although TACO dataset is a solid starting point, it is certainly needed an improvement in the amount of tagged photos in the collection of the training and test sets, that is periodically supported. Furthermore, it has been realized that the identification results on small objects (such as cigarettes) in the sand with the presented network configuration are still not sufficient, affecting the overall AP (Average Precision) greatly. Future work should develop better models and ways to fully use the high resolution of the garbage image dataset. Indeed, it is possible to increase the input resolution, but this dramatically increases the memory footprint, increasing requirements for computational power for the autonomous systems. Alternatively, Mask R-CNN might be performed in a sliding window fashion and then predictions fused; however, this approach would remove context from the surrounding windows and introduce other criticalities. As a result, an approach that is both efficient and lossless is necessary to be obtained by further researches. The authors are working on extending this work by developing also the part related to the swarm intelligence and coordination system to be used to support operations over a wide area as well as additional services for the UGV devoted to edutainment of people on the area to improve the sustainability culture and improve acceptance and social sustainability of this approach.

References

A.G. Bruzzone et al. (2014). Simulation exploration experience: providing effective surveillance and defense for a moon base against threats from outer space. *18th International Symposium on Distributed Simulation and Real Time Applications (pp. 121-126). IEEE/ACM.*

A.G. Bruzzone et al. (2016). Autonomous Systems for operations in critical environments. *Simulation Series, 48(5) 17-24*

A.G. Bruzzone et al. (2019a). Application of blockchain in interoperable simulation for strategic decision making. *Summer Simulation Conference (pp. 1-10).*

A.G. Bruzzone et al. (2019b). A digital twin approach to develop a new autonomous system able to operate in high temperature environment within industrial plants. *Summer Simulation Conference (pp. 22-24)*

A.G.Bruzzone et al. (2020). Modeling Human Physiology Coupled with Hyperbaric Plant Simulation for Oil and Gas. *International Journal of Privacy and Health Information Management (IJPHIM), 8(1), 1-12.*

L.-C Chen, Z., P., S. and A. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. *Proceedings of the European Conference on Computer Vision (ECCV)*

G.-R Jocher et al. (2020). yolov5. Code repository

Q. Lei et al. (2021). Garbage Classification System with YOLOV5 Based on Image Recognition. *Intelligent Robot & Equipment Center Guangzhou Institute of Advanced Technology, Chinese Academy of Sciences*

Q. Liu and C. Li. (2021). A Garbage Detection and Classification Method Based on Visual Scene Understanding in the Home Environment. *Wiley, Complexity Volume 2021*

S. Majchrowska et al. (2022). Deep learning-based waste detection in natural and urban environments. *Waste Management, Volume 128*

L. Parker. (2018). Here's how much plastic trash is littering the Earth. *National Geographic.*

U. S. Environmental Protecting Agency (EPA). (2020). EPA's Report on the Environment (ROE). Human Exposure & Health: Disease and Conditions

W. Zeng, R.L. Church. (2007). Finding shortest paths on real road networks: the case of A*. *Department of Geography, University of California*

Z. Wu et all. (2021). Using YOLOv5 for Garbage Classification. *2021 4th International Conference on Pattern Recognition and Artificial Intelligence (PRAI)*