

Research Article

Faster R-CNN Algorithm for Detection of Plastic Garbage in the Ocean: A Case for Turtle Preservation

Muhammad Faisal ¹, Sushovan Chaudhury ², K. Sakthidasan Sankaran ³,
S. Raghavendra ⁴, R. Jothi Chitra ⁵, Malathi Eswaran ⁶, and Rajasekhar Boddu ⁷

¹Department of Computer Science, Sekolah Tinggi Manajemen Informatika Dan Komputer Profesional, A.P Petarani No. 27 Road, Makassar 90231, Indonesia

²Department of Computer Science and Engineering, University of Engineering and Management, Kolkata, India

³Department of ECE, Hindustan Institute of Technology and Science, Chennai, India

⁴Department of Information and Communication Technology, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal 576104, India

⁵Department of ECE, Velammal Institute of Technology, Chennai, Tamilnadu, India

⁶Department of Computer Technology-PG, Kongu Engineering College, Erode, Tamilnadu, India

⁷Department of Software Engineering, College of Computing and Informatics, Haramaya University, Dire Dawa, Ethiopia

Correspondence should be addressed to S. Raghavendra; raghavendra.s@manipal.edu and Rajasekhar Boddu; rajsekhar.boddu@haramaya.edu.et

Received 16 March 2022; Accepted 25 April 2022; Published 26 May 2022

Academic Editor: Parikshit Narendra Mahalle

Copyright © 2022 Muhammad Faisal et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Turtles are one of the ancient marine animals that live today. However, the population is threatened with extinction, so its existence needs to be protected and preserved because turtles often eat plastic waste in the ocean whose shape, texture, and color are similar to jellyfish. The technology in the computer vision area can be used to find the solution related to the case of reducing plastics and bottles trash in the ocean by implementing robotics. The region-based Convolutional Neural Network (CNN) is the latest image segmentation and has good detection accuracy based on the Faster R-CNN algorithm. In this study, the training image was built based on two different objects, namely plastic bottles and plastic bags. The target is that the two objects can be recognized even if there are other objects in the vicinity, or the image quality will be affected by the color of the seawater. The results obtained are that plastic objects and bottles can be recognized correctly in the picture. Of the five-color hues tested, the results show that the object detection process is valid on the average color hue, sepia, bandicoot, and grayscale. In contrast, the object detection process is invalid in black-and-white tones. The test results shown in the table explain that the object detection that gets the highest results is an image with normal coloring, while the lowest value is on bandicoot. The average accuracy of all types of images tested is 96.50. However, the accuracy value still needs to be improved to apply feasibility permanently to hardware such as diving robots.

1. Introduction

Turtles are one of the ancient marine animals that live to this day. However, the population is threatened with extinction, so it needs to be protected and preserved were. Of the seven sea turtles globally, six are found in Indonesia [1]. Several species of turtles have become endangered, which is why there is a need for search and rescue [2]. Efforts are needed to protect it through turtle conservation [3]; so that it is expected

to prevent the extinction of turtle habitat. One of the turtle foods is jellyfish; however, turtles often eat plastic waste in the oceans because of their shape, texture, and color, similar to jellyfish, so that they will think of them as jellyfish. Over the last decade, breakthroughs in the domains of machine learning, statistics, and computer vision have piqued researchers' interest in advanced deep learning techniques [4].

The Faster R-CNN methodology can be used to detect plastic waste objects at sea that are applied to diving tools or

robots to reduce the distribution of plastic waste in the sea. In datasets with occlusion and overlap, the Faster R-CNN beats other masks in terms of computation precision and mean precision [5]. The dataset used in this study is made up of photos with .jpg and .png extensions that were gathered from Google Image and Shutterstock.com in the form of images of plastic bags and plastic bottles in the sea, with a total of 400 image datasets, 140 of which are .jpg files and 260 of which are .png files. Each step in this research is shown by a flowchart in Figure 1.

2. Background

2.1. Computer Vision. Computer vision is a branch of technology that identifies, tracks, and measures targets for further image processing using a camera and a computer as an image to the human eye [6]. Deep learning approaches have made significant contributions to computer vision applications such as picture classification, object detection, and image segmentation [7]. Computer vision and machine-learning algorithms have mainly been studied in a centralized setting, where all processing is done in one central location. Object detection, object classification, and extraction of useful information from photos, graphic documents, and videos are among the most recent machine-learning applications in computer vision [8].

The machine-learning paradigm for computer vision supports vector machines, neural networks, and probabilistic graphical models. Machine learning in computer vision plays an essential role in object recognition, and image classification uses a tensor-flow library that can improve accuracy when recognizing objects [9]. Figure 2 shows the object detection process in a machine learning and computer vision environment.

Based on the illustration in Figure 2. It is explained that after detecting objects in the image, the next feature will be extracted from the given image, where every single image is broken down into small pieces containing a collection of information. The extraction process is seen in Figure 3.

2.2. Region Convolutional Neural Network R-CNN. Region Convolutional Neural Network (R-CNN) is based on deep learning object detection, commonly used for object detection. R-CNN uses a selective search algorithm to propose the image, where the input image will be grouped into 2000 regions that are selected based on texture, intensity, and color. This is done to cover the weakness of CNN, which divides the image region with a large regional scale which makes the identification process slower as shown in Figure 4.

Mask R-CNN is a Region-based Convolutional Neural Network that is state-of-the-art in image segmentation and has a good Faster R-CNN method [12]. This Deep Neural Network variant detects objects in the image and generates a high-quality segmentation mask for each instance [13]. Image segmentation is becoming a significant task in computer vision and image processing with essential

applications such as scene understanding, medical image analysis, robotic perception, video surveillance, augmented reality, and image compression [14]. Weaknesses in R-CNN include the relatively slow data training process because it uses 2000 proposal regions for each image. Besides that, it cannot be implemented for real-time classification because it takes about 47–50 seconds to process per image. The last one is that R-CNN can only do selective search algorithms in the introduction process and cannot use other algorithms for a selective search [11].

The R-CNN Mask is simple to implement and adds only a slight overhead to the Faster R-CNN, which runs at five frames per second. Furthermore, the R-CNN Mask is simple to apply to other tasks.

2.3. Faster Region Convolutional Neural Network (Faster R-CNN). Faster R-CNN is a method based on deep learning object detection, which is commonly used for object detection developed from the R-CNN algorithm to cover the weaknesses that exist in R-CNN. The advantage of Faster R-CNN is that Faster R-CNN uses RPN, where RPN is a neural network that replaces the role of selective search to propose regions. The role of selective search is replaced because the process is slow in processing images, which is about 2 seconds per image [15]. RPN serves to generate several bounding boxes where each box has 2 probability scores whether there are objects at that location or not, with the RPN processing is not repeated as is done in R-CNN and makes the whole model one that can be trained by end to end. Figure 5 shows the General architecture Faster R-CNN. One disadvantage of Faster R-CNN is that, unlike RPN, all anchors in the minibatch are extracted from a single image. The Faster R-CNN algorithm is very effective for cult problems in detecting some small still has some limitations in detecting camouflaged objects. As a result, the test is performed on five types of object coloring, including Normal, Sepia, Bandicoot, Grayscale, and black-white, to determine which types of coloring do not support a good recognition process. All samples from one image may be correlated. Because the network may take a long time to reach convergence, Mask R-CNN can return a mask for each detected object [16, 17].

Faster R-CNN is easy to deploy and train due to the Faster R-CNN framework, which facilitates a variety of flexible architectural designs and consists of mask branches that add only a small computational overhead, thus enabling fast systems and short experiments [18]. We chose Faster R-CNN due to its extremely high precision, which outperforms other algorithms, and its ability to detect small objects. Because plastic is a transparent object in water, our primary goal is to detect plastic waste with the most excellent precision possible [19]. As a result, we compromise the FPS rate because we are satisfied with the faster R-CNN. Seven frames per second is a faster rate.

Table 1 describes the object detection performance, with the Faster R-CNN algorithm appearing to achieve the highest precision (mAP).

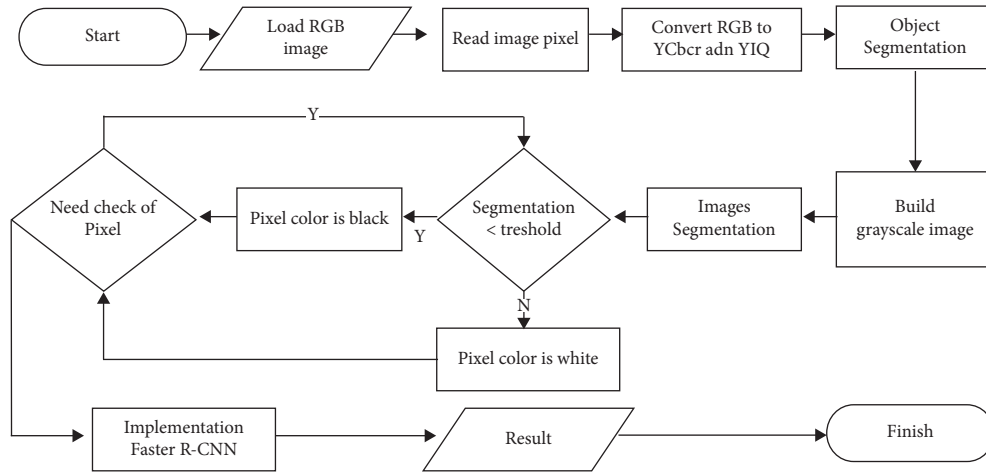


FIGURE 1: Research process flowchart.



FIGURE 2: Image processing and object detection process [8].



FIGURE 3: Turtle image before and after features extraction [10].

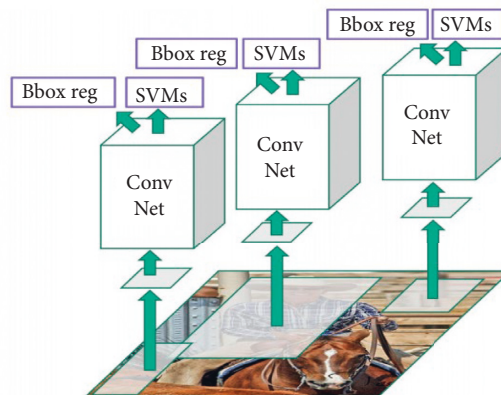


FIGURE 4: General architecture R-CNN [11].

3. Literary Review

The research that forms part of this paper has been based on the work studied through several publications. Table 1. Summarizes all of the publications. (Table 2)

4. Methodology

A research approach is an action plan that provides direction for conducting research systematically and efficiently. The explanation of the method used. One disadvantage of Faster

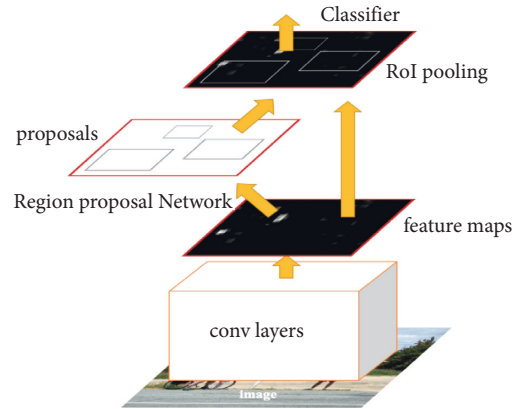


FIGURE 5: General architecture Faster R-CNN [11].

TABLE 1: Object detection results on aerial images collected by drone [20].

No.	Method	Car	Cat	Dog	Person	mAP
1	Yolo	87.2	100.0	81.0	88.7	79.4
2	SSD 300	100.0	100.0	66.7	92.9	21.6
3	SSD 500	93.2	100.0	69.6	81.7	82.6
4	Faster R-CNN	89.7	100.0	77.3	81.7	83.9

TABLE 2: The research publication.

Publication	Contribution	Comments
Bin Liu et al. (2017). Study of object detection based on faster R-CNN	Implementation of the faster R-CNN algorithm to detect objects such as cats, humans, cars, and horses [21]	Related research demonstrates the classification process using faster R-CNN + PVANET, but produces a not high average precision of 84.9%. Therefore, the author does not combine several methods based on the consideration that the faster R-CNN method can still produce higher precision values
Shih-Chung Hsu et al. (2018). Vehicle detection using simplified fast R-CNN	Implementation of fast R-CNN algorithm to detect vehicles [22]	Related research presents a fast and straightforward method by modifying fast R-CNN to be able to detect and localize vehicles in various displays effectively. Therefore, the author uses the faster R-CNN method, which develops the fast R-CNN method, where faster R-CNN can process faster recognition
Beibei Zhu et al. (2016). Automatic detection of books based on faster RCNN	Implementation of the faster R-CNN algorithm to detect books based on the object's shape [23]	In a related study, we adopted the faster R-CNN code framework, which was created to implement efficient and accurate book detection. Although in line with the author in using the faster R-CNN method, the negative object recognition test is more complex because it is carried out on various types of coloring
Xiaochun Mai et al. (2018). Faster R-CNN with classifier fusion for small fruit detection	Implementation of the faster R-CNN algorithm using five convolution screens to detect almonds still on the tree [24]	Related research explores the effectiveness of faster R-CNN to improve object classification, but some labeling and annotation errors result in uncertainty. Based on the constraints found in previous research, the authors consider using a larger sample training dataset to reduce the uncertainty during the object recognition process

TABLE 2: Continued.

Publication	Contribution	Comments
Wei Zhang et al. (2018). Deconv R-CNN for small object detection on remote sensing images	Implementation of the R-CNN algorithm succeeded in detecting microscopic aircraft objects taken from a height [25]	Related research can detect small objects, aircraft, and ships with time efficiency and high-detection precision. However, the author still needs to prove the results of the unique detection of ships if they are in an ocean area with sea color conditions that can change. This follows the plastic object detection tests carried out based on various hues categories The related research applies the faster R-CNN method for turtle detection through the onboard camera mounted on the drone, which effectively contributes to ecosystem solutions and environmental research in general and turtle conservation projects. Therefore, as a support, the authors conducted research on different cases but used the same method and had the same goal, namely focusing on ecosystem solutions for turtles
Mohamed badawy (2020). Sea turtle detection using faster R-CNN for conservation purpose	The method suggested an intelligent system for sea turtles detection where the faster R-CNN algorithm is employed impressively and gives promising results [26]	

R-CNN is that for RPN, all anchors in the minibatch are extracted from one image because all samples from one image may be correlated. The network may take a long time to reach convergence; therefore, Mask R-CNN can return a mask for each detected object. A research approach is an action plan that provides direction for conducting research systematically and efficiently. The explanation of the method used. Faster R-CNN integrates candidate region extraction, deep feature extraction, classification, and bounding box regression into a deep neural network faster.

4.1. Library Tensor-Flow. Tensor Flow is an open-source machine-learning library for research and software development. Tensor Flow offers beginners and specialists APIs for desktop, mobile, web, and cloud computing-based application development. In implementing the object detection process using the Faster R-CNN, one of the backend engines, tensor-flow, is first installed. This paper focuses on the primary use of the tensor-flow library working on the backend. The Object Detection API in TensorFlow is a powerful tool that allows anyone to quickly design and deploy practical picture recognition applications [27]. Object detection entails classifying and recognizing items in a picture and localizing and attracting bounding boxes for those objects. Tensor flow is also cross-platform, which means it can run on any platform, including GPUs, CPUs, and even mobile platforms. It also has specialized hardware for tensor math, known as a tensor-processing unit (TPU) [28].

4.2. Data Item Collection. Process Dataset aims to train and test neural networks and develop algorithms in computer vision [29]. The formation of the dataset begins by placing the folder containing the images in the.zip archive, then uploading it to Google Colab. A collection of images contained in the dataset is shown in Table 3.

4.3. Label Images Annotation. Image annotation is a branch of image retrieval used to label or tag images with a set of keywords based on the content of the idea that produces labels that can be used for grouping images based on the content of the image for easy management [30]. Figure 6 shows the image annotations process.











Object detection is consolidated into instance segmentation, with the purpose of classifying and localizing each object using bounding boxes. The purpose is to assign each pixel to a certain object class [31]. The point coordinates will be stored in a JSON file for each image in the annotation process. Although sometimes a minor error occurs during the image annotation process, it will not affect the overall model evaluation [32, 33].

4.4. The Method Faster R-CNN Recognition and Results. In this research, the proposed recognition and result are shown in Figure 7.

The illustration is shown in Figure 7 and explained that the process is part of two processes, namely:

4.4.1. Training Session. The system is given input data in images from plastic bag waste and plastic bottles where the file resizing is carried out no more than 200 KB so that the image size is reduced horizontally or vertically. Through the convolutional layer, the image features will be extracted and studied, and the essential parts that can be a characteristic of an object through the feature map created by the convolutional layer, which contains information about the vector representation of the captured image. RPN (Region Proposal Network) is a module that functions through two convolution layers where one layer is responsible for detecting the location of objects and one layer functions to predict bounding boxes. The output of RPN is the proposal region of the image. ROI is the layer responsible for equalizing the size of the feature map and proposal region that has been processed by the RPN and sending feature map

TABLE 3: Data items collection.

Plastic bags					
Plastic bottles					

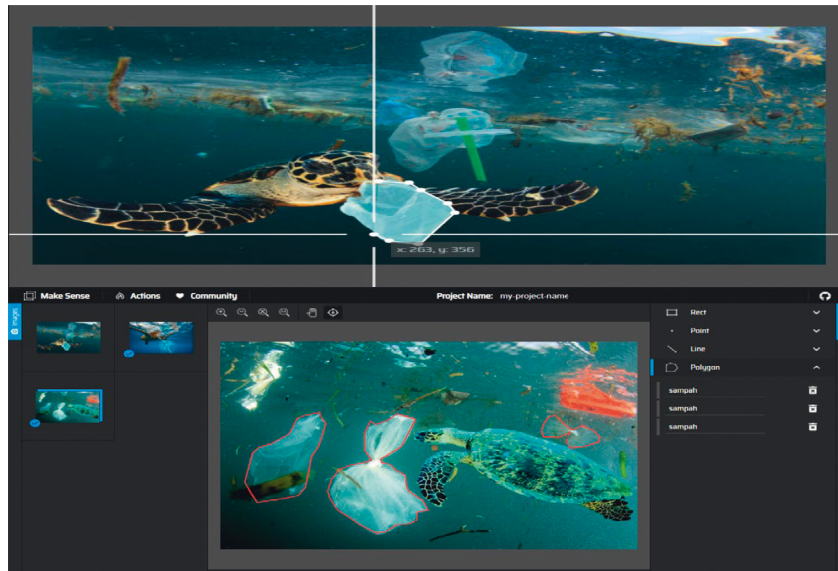


FIGURE 6: Image annotations process.

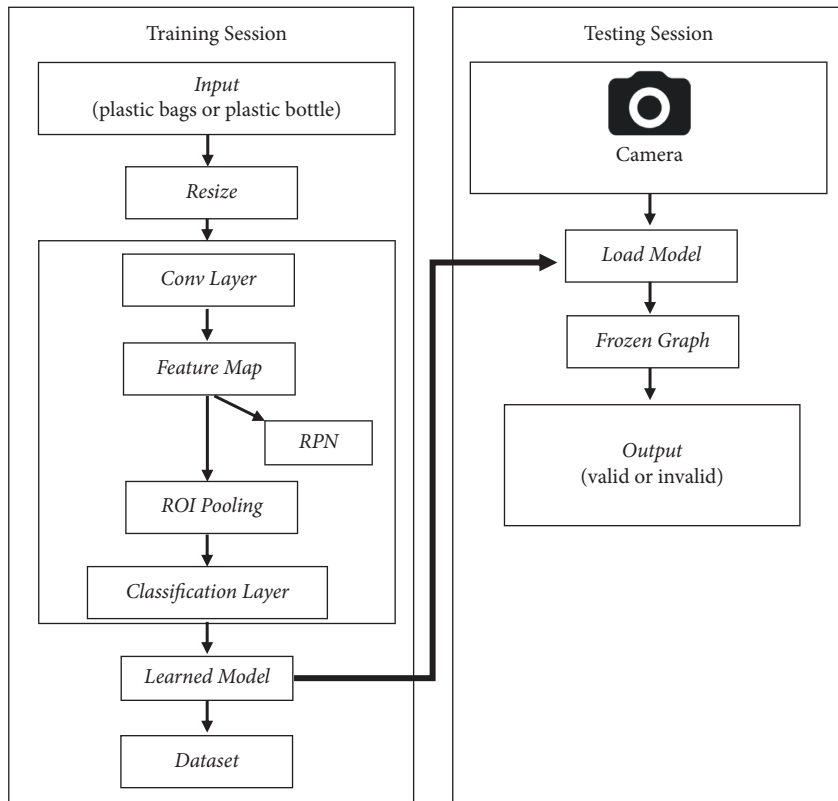


FIGURE 7: the proposed recognition and result.

information and proposals to be classified at the classification layer. The function of the classification layer is to group objects that have been detected in the RPN and perform labeling and assigning a bounding box to an object. Finally, after the system runs the learned model process, a dataset containing dataset weight information.

4.4.2. Testing Session. At the initial testing stage, the system will be given input data in the form of images from plastic bags and plastic bottles, then run the load model process so that the system reloads the model stored during the training session. Frozen Graph is a process where input data received through the camera will be processed on a graph stored in a frozen model to identify and assign a bounding box based on the weights that have been stored in the model that has been trained. The output is the result of the identification carried out and bounding boxes and labels for objects that have been classified.

5. Discussion

In this study, the plastic image object detection process stages are the following: first, the object image is obtained with a self-made image acquisition device. Second, the objects will be processed, labeled, and inserted into the Faster R-CNN for training. Finally, the trained model is used to segment the picture of the training object in order to acquire an indicator of the item's features.

5.1. Training Image Samples. Data were gathered from the Internet for this investigation, and the dataset was separated into training sets, validation sets, and test sets for the experiment. The training was carried out on plastic images and beverage bottles. Most methods for object instance segmentation require all training instances to be labeled with a segmentation mask [34]. Image training is frequently used to decide what heterogeneity should be included in a multi-point statistical reservoir model [35]. In this study, the training image is built based on two different objects, namely plastic bottles and plastic bags. The target is that both objects can be identified even though there are other objects around the object, or the image quality will be affected by the color of the seawater. The training results are shown in Figure 8.

Faster R-CNN parameters used at the training stage can be seen in Table 4.

In this study, to build a detection model using Faster R-CNN, a total of 92,998 images were used, consisting of 22,461 images of plastic bags and 69,996 images of plastic bottles. Before the training process is carried out, a text file contains the image name, bounding box size, and class (label) information. The training data is divided into training data with as many as 76,990 images and data validation as many as 15,467 images. The CNN architecture used is Resnet50, a model that has been trained using the ImageNet Dataset to produce good feature extraction. According to the default anchor Faster R-CNN, the number of anchors used is nine. Anchor is an important part that is used to determine an essential part of the image (proposal region) that will be

included in the RPN. The optimizer used is a plastic bag with a learning rate of 0.00001. In addition, Stochastic Gradient Distance is used to optimize the convolution layer, RPN weight, and fully connected layer. The epoch length used is 50,000, with a total epoch of 25.

Table 5 shows the results obtained at the training stage. Based on the table, the highest accuracy value is obtained in the 25th epoch with a value of 96.66%, the error rate is getting lower, and the execution time is 7 hours 21 minutes 12 seconds.

In this research, the Faster R-CNN recognition method and the result are shown in Figure 7.

The RPN loss is the sum of the classification loss and the bounding box regression loss, where the classification loss penalizes incorrectly classified boxes using cross-entropy loss, and the regression loss penalizes incorrectly predicted regression coefficients using a function of the distance between the accurate regression coefficients and the regression coefficients predicted by the network. The neural network is trained by specifying a multitask loss function:

$$L(\{P_x\}\{I_x\}) = \frac{1}{N_{cls}} \sum_x L_{cls}(P_x P_x^*) + \lambda \frac{1}{N_{reg}} \sum_x P_x^* L_{reg}(I_x P_x^*), \quad (1)$$

where N_{cls} , N_{reg} , and λ balance the normalized weights of classification loss and regression loss, and I is the index of the x candidate frame in small-batch processing. P_x is the probability that the x candidate box is the target. If the x candidate box is a candidate target, then $P_x^* = 1$; otherwise, $P_x^* = 0$.

The classification and regression loss functions are defined as formulae in equations (2) and (3):

$$\sum_x L_{cls}(P_x P_x^*) = -\log[P_x P_x^* + (1 - P_x^*)(1 - P_x)], \quad (2)$$

$$\sum_x L_{reg}(t_x t_x^*) = R(t_x - t_x^*), \quad (3)$$

where R is the smooth_{L1} function $t_x = \{t_z, t_y, t_w, t_h\}$ is a vector prediction parameterized candidate frame coordinates and $t_x^* = \{t_z^*, t_y^*, t_w^*, t_h^*\}$ is the coordinate vector of actual boundaries.

5.2. Testing. The Faster R-CNN approach recognizes objects using random images in the test. The results show that plastic objects and bottles in the image can be identified correctly. The results are seen in Table 6.

Based on the test results shown in Table 6, the image used in the testing process uses several types of color shades. This is done as an example of conditions in seawater, where the color of seawater can be affected by certain conditions, which can also affect the accuracy of the object detection process. The five-color hues tested show that the object detection process is valid in ordinary, sepia, bandicoot, and grayscale color tones. In contrast, the object detection process is invalid in black-and-white tones. The author assumes that the black-and-white process is the condition of the seawater at night or the seawater is polluted by waste oil,

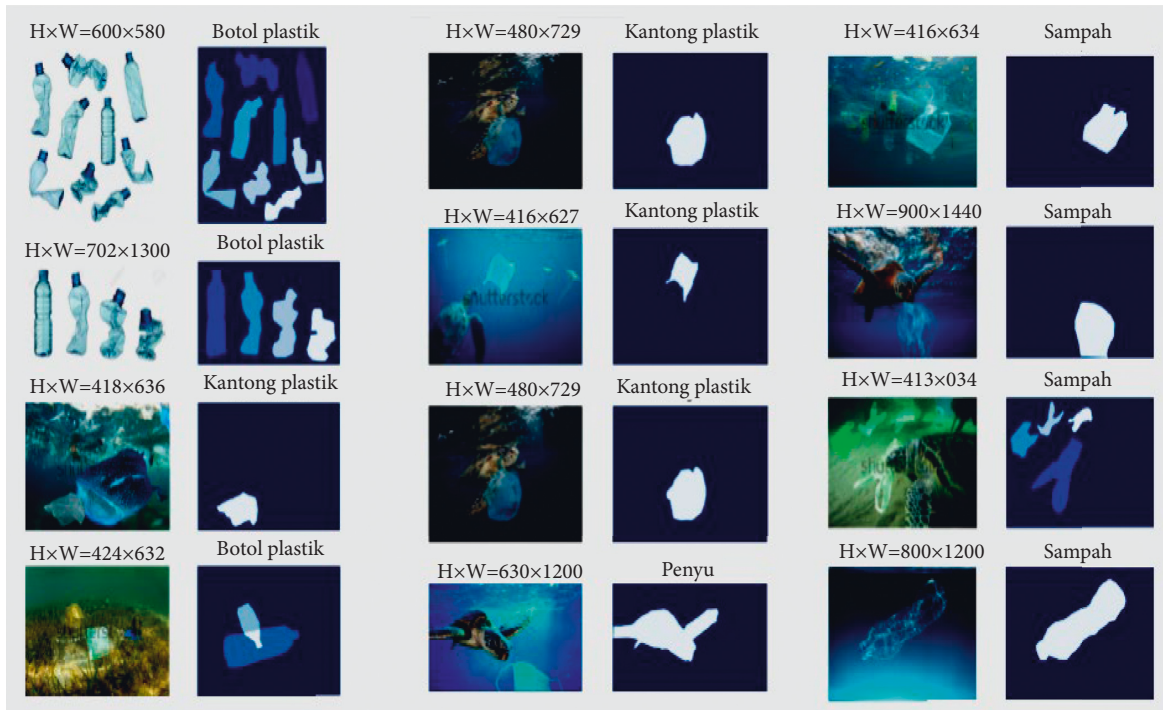


FIGURE 8: Training image samples.


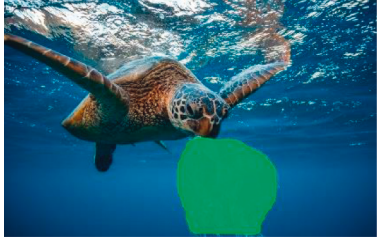

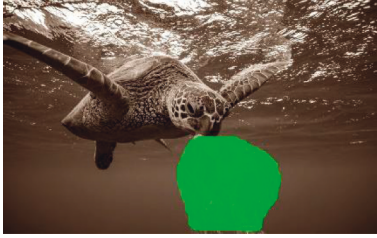

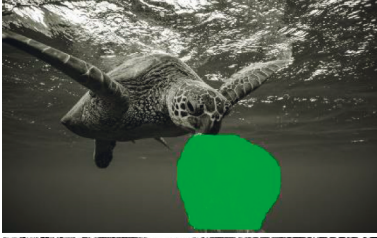

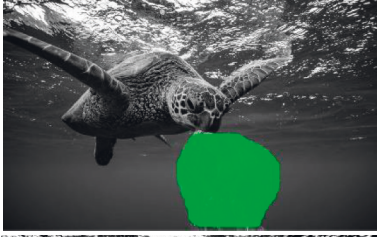

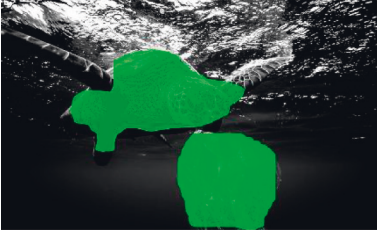
TABLE 4: Parameters of training.

Parameter	Weight
Input node	1
Output node	2
Iteration	50000
Learning rate	0.00001
Loss function	0.005
IOU	0.75

TABLE 5: Results of the training process.

Epoch	Loss RPN Classifier	Loss RPN Regression	Loss Detector Classifier	Loss Detector Regression	Accurate	Waktu (H : M : S)
1	0.308	0.21	0.13	0.06	93.92	07 : 25 : 00
2	0.213	0.009	0.085	0.026	95.55	07 : 14 : 57
3	0.197	0.007	0.084	0.022	95.62	07 : 13 : 25
4	0.202	0.005	0.080	0.020	95.84	07 : 18 : 11
5	0.113	0.005	0.082	0.019	95.76	07 : 15 : 10
6	0.011	0.004	0.082	0.018	95.80	07 : 15 : 21
7	0.009	0.004	0.078	0.017	95.97	07 : 22 : 05
8	0.007	0.004	0.075	0.016	96.04	07 : 17 : 07
9	0.008	0.003	0.074	0.015	96.09	07 : 16 : 29
10	0.006	0.003	0.071	0.014	96.25	07 : 21 : 27
11	0.007	0.003	0.072	0.014	96.17	07 : 19 : 45
13	0.007	0.003	0.068	0.013	96.38	07 : 20 : 22
14	0.006	0.003	0.069	0.013	96.33	07 : 19 : 13
15	0.006	0.002	0.067	0.013	96.39	07 : 17 : 37
16	0.005	0.002	0.066	0.013	96.44	07 : 19 : 16
17	0.004	0.002	0.067	0.012	96.40	07 : 18 : 50
18	0.004	0.002	0.064	0.012	96.54	07 : 17 : 48
19	0.003	0.002	0.062	0.012	96.58	07 : 21 : 57
20	0.002	0.002	0.062	0.012	96.58	07 : 22 : 08
21	0.003	0.002	0.061	0.011	96.62	07 : 21 : 12
22	0.002	0.002	0.061	0.011	96.64	07 : 21 : 18
23	0.003	0.003	0.060	0.010	96.65	07 : 22 : 02
24	0.002	0.002	0.060	0.010	96.65	07 : 21 : 08
25	0.002	0.002	0.059	0.010	96.66	07:21:12

TABLE 6: Testing images.

No	Original image	Result	Hues	Validation
1			Normal	Valid
2			Sepia	Valid
3			Bandicoot	Valid
4			Grayscale	Valid
5			Black white	Invalid

so it can be considered when applied to a diving machine or robot so that it does not work when the color of the seawater is black, or the level of clarity is very cloudy.

Various sorts of plastic bags and plastic bottles pictures will be used in the tests. The Confusion Matrix is used for testing, and the values of accuracy are used, where data was collected from 400 images of plastic bags and bottles. (Table 7).

The accuracy in the table is using the formula in (4):

$$\begin{aligned}
 \text{Accuracy} &= \frac{T}{N} \times 100\%, \\
 &= ((98 + 96 + 95 + 97)/400) \times 100\% \\
 &= (386/400) \times 100\% \\
 &= 96.50\%.
 \end{aligned}
 \tag{4}$$

TABLE 7: Testing images.

Plastic bag	Plastic bottle		Bandicoot	Grayscale	Total
	Normal	Sepia			
Normal	98	1	1	0	100
Sepia	1	96	2	1	100
Bandicoot	1	2	95	2	100
Grayscale	0	1	2	97	100
Total	100	100	100	100	400
	Accuracy			96.50%	

The test results shown in the table explain that the object detection that gets the highest results is an image with normal coloring, while the lowest value is on bandicoot. The average accuracy of all types of images tested is 96.50.

6. Conclusion

This study concludes that turtle population extinction can be prevented by helping reduce marine pollution by plastic waste. When applied to robotic technology or diving equipment, the Faster R-CNN approach can assist in segmentation and target item detection. Object identification algorithms must be able to run in near real-time on robotic platforms in order to be beneficial for the purpose of eliminating those plastics and other waste. The work given here is an algorithm for the protection of turtle species, which can become endangered if such measures are not implemented. When the item is black and white, the Faster R-CNN approach has limitations and is therefore recommended for use in clear seawater conditions. In the future, we want to expand on this work by evaluating analogous algorithms on a dataset collected from our own observations of marine trash in real-world settings. We'd also like to consider other approaches to accomplishing this project.

Data Availability

The data are available on request.

Conflicts of Interest

The author(s) declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors are thankful to the support by the STMIK Professional Makassar. The present research work is self-funded.

References

- [1] R. Ario, E. Wibowo, I. Pratikto, and S. Fajar, "Pelestarian habitat penyu Dari ancaman kepunahan di turtle conservation and education center (TCEC), bali," *Jurnal Kelautan Tropis*, vol. 19, no. 1, p. 60, 2016.
- [2] Y. Ma and W. Wei, "Posture data automatic extraction of ornamental turtle based on computer vision technology," *Journal of Physics: Conference Series*, vol. 2025, no. 1, Article ID 012063, 2021.
- [3] A. Nurhayati, T. Herawati, I. Nurruhwati, and I. Riyantini, "Tanggung jawab masyarakat lokal pada konservasi penyu hijau (*Chelonia mydas*) di Pesisir selatan jawa barat," *Jurnal Perikanan Universitas Gadjah Mada*, vol. 22, no. 2, p. 77, 2020.
- [4] P. Bharati and A. Pramanik, "Deep learning techniques-R-CNN to mask R-CNN: a survey," *Computational Intelligence in Pattern Recognition*, vol. 999, pp. 657–668, 2020.
- [5] B. Xu, W. Wang, G. Falzon et al., "Automated cattle counting using Mask R-CNN in quadcopter vision system," *Computers and Electronics in Agriculture*, vol. 171, Article ID 105300, 2020.
- [6] H. Tian, T. Wang, Y. Liu, X. Qiao, and Y. Li, "Computer vision technology in agricultural automation -A review," *Information Processing in Agriculture*, China Agricultural University, vol. 7, no. 1, pp. 1–19, 2020.
- [7] X. Feng, Y. Jiang, X. Yang, M. Du, and X. Li, "Computer vision algorithms and hardware implementations: a survey," *Integration*, vol. 69, pp. 309–320, 2019.
- [8] A. I. Khan and S. Al-Habsi, "Machine learning in computer vision," *Procedia Computer Science*, vol. 167, pp. 1444–1451, 2020.
- [9] R. Bandi and J. Amudhavel, "Object recognition using keras with backend tensor flow," *International Journal of Engineering & Technology*, vol. 7, no. 3.6, p. 229, 2018.
- [10] Z. Attal and C. Direkoglu, "sea turtle species classification for environmental research and conservation," in *Advances in Intelligent Systems and Computing* vol. 1095 AISC, , pp. 580–587, Springer, 2020.
- [11] R. Gandhi, "R-CNN , fast R-CNN , faster R-CNN, YOLO — object detection algorithms understanding object detection algorithms R-CNN," 2018, <https://towardsdatascience.com/r-cnn-fast-r-cnn-faster-r-cnn-yolo-object-detection-algorithms-36d53571365e>.
- [12] Y. Su, D. Li, and X. Chen, "Lung nodule detection based on faster R-CNN framework," *Computer Methods and Programs in Biomedicine*, vol. 200, Article ID 105866, 2021.
- [13] S. Saralch, V. Jagota, D. Pathak, and V. Singh, "Response surface methodology-based analysis of the impact of nanoclay addition on the wear resistance of polypropylene," *The European Physical Journal - Applied Physics*, vol. 86, pp. 1–13, Article ID 10401, 2019.
- [14] S. Minaee, Y. Y. Boykov, F. Porikli, A. J. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image segmentation using deep learning: a survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, p. 1, 2021.
- [15] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017.

- [16] J. Hong, M. Fulton, and J. Sattar, "Trashcan: a semantically-segmented dataset towards visual detection of marine debris," 2020, <https://arxiv.org/abs/2007.08097>.
- [17] M. Fulton, J. Hong, M. J. Islam, and J. Sattar, "Robotic detection of marine litter using deep visual detection models," in *Proceedings of the 2019 International Conference on Robotics and Automation (ICRA)*, pp. 5752–5758, IEEE, Montreal, QC, Canada, May 2019.
- [18] L. Wang, P. Kumar, M. E. Makhatha, and V. Jagota, "Numerical simulation of air distribution for monitoring the central air conditioning in large atrium," *International Journal of System Assurance Engineering and Management*, vol. 13, no. S1, pp. 340–352, 2021.
- [19] K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask R-CNN," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 2, pp. 386–397, 2020.
- [20] J. Lee, J. Wang, D. Crandall, S. Sabanovic, and G. Fox, "Real-time, cloud-based object detection for unmanned aerial vehicles," in *Proceedings of the 2017 First IEEE International Conference on Robotic Computing (IRC)*, pp. 36–43, Taichung, Taiwan, April 2017.
- [21] B. Liu, W. Zhao, and Q. Sun, "Study of object detection based on Faster R-CNN," in *Proceedings of the 2017 Chinese Autom. Congr. CAC 2017*, vol. 2017-Janua, pp. 6233–6236, Jinan, China, October 2017.
- [22] S.-C. Hsu, C.-L. Huang, and C.-H. Chuang, "Vehicle detection using simplified fast R-CNN," in *Proceedings of the 2018 International Workshop on Advanced Image Technology (IWAIT)*, pp. 1–3, Chiang Mai, Thailand, January 2018.
- [23] J. Wu, S. Zhuo, and Z. Wu, "National innovation system, social entrepreneurship, and rural economic growth in China," *Technological Forecasting and Social Change*, vol. 121, pp. 238–250, 2017.
- [24] X. Mai, H. Zhang, and M. Q.-H. Meng, "Faster R-CNN with classifier fusion for small fruit detection," in *Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 7166–7172, Brisbane, Australia, May 2018.
- [25] W. Zhang, S. Wang, S. Thachan, J. Chen, and Y. Qian, "Deconv R-CNN for small object detection on remote sensing images," in *Proceedings of the IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, vol. 2018, pp. 2483–2486, Valencia, Spain, July 2018.
- [26] M. Badawy and C. Direkoglu, "Sea turtle detection using faster r-cnn for conservation purpose," *Advances in Intelligent Systems and Computing*, AISC, vol. 1095, , pp. 535–541, 2020.
- [27] R. A. Canessane, R. Dhanalakshmi, and V. M. Anu, "Implementation of tensor flow for real-time object detection," *International Journal of Recent Technology and Engineering*, vol. 8, no. 2S11, pp. 2342–2345, 2019.
- [28] B. N. K. Sai and T. Sasikala, "Object detection and count of objects in image using tensor flow object detection API," in *Proceedings of the 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT)*, pp. 542–546, Tirunelveli, India, November. 2019.
- [29] R. Singhla, P. Singh, R. Madaan, and S. Panda, "Image classification using tensor flow," in *Proceedings of the 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS) 2021*, pp. 398–401, Tamilnadu, India, March 2021.
- [30] B. Berenguel-Baeta, J. Bermudez-Cameo, and J. J. Guerrero, "Omnidirectional image data-set for computer vision applications," *Jornada de Jóvenes Investigadores del I3A*, vol. 8, 2020.
- [31] V. Muhammed Anees, G. Santhosh Kumar, and M. Sreeraj, "Automatic image annotation using SURF descriptors," in *Proceedings of the 2012 Annual IEEE India Conference (INDICON)*, vol. 68, no. 4, pp. 920–924, Kochi, India, December 2012.
- [32] S. Ramesh and V. Vinod Kumar, "A review on instance segmentation using mask R-CNN," in *Proceedings of the International Conference on Systems, Energy & Environment (ICSEE) 2021*, Sydney, Australia, August 2021.
- [33] G. Zhu, Z. Piao, and S. C. Kim, "Tooth detection and segmentation with mask R-CNN," in *Proceedings of the 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)*, pp. 070–072, Fukuoka, Japan, February 2020.
- [34] R. Hu, P. Dollar, K. He, T. Darrell, and R. Girshick, "Learning to segment every thing," in *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4233–4241, Salt Lake, UT, USA, June 2018.
- [35] A. J. Mitten, J. Mullins, J. K. Pringle, J. Howell, and S. M. Clarke, "Depositional conditioning of three dimensional training images: improving the reproduction and representation of architectural elements in sand-dominated fluvial reservoir models," *Marine and Petroleum Geology*, vol. 113, Article ID 104156, 2020.