International Conference on Recent Trends in Data Science and its Applications
DOI: rp-9788770040723.015

Vision Based Object Detection Method for Water Surface Garbage Capture Robot

Renit C
Department of ECE,
St.Xaviers Catholic College Engineering
Chunkankadai, Nagercoil India
renitsweety@gmail.com

Harrish Kumar G M
Department of Information Technology,
National Engineering College
Kovilpatti, India
1915036@nec.edu.in

Kumar S
Department of Information Technology,
National Engineering College
Kovilpatti, India
1915032@nec.edu.in

Vignesh A
Department of Information Technology,
National Engineering College
Kovilpatti, India
1915032@nec.edu.in

JerartJulusL
Department of Information Technology
National Engineering College
Kovilpatti, India
jerartjulus@nec.edu.in

***Abstract*—The environment is covered with many water resources. The living organism on earth relay more on water. Due to the growth of innovation and Technologies there is a need for protecting our environment. There are many plastics-based materials used in our day today life for packing, storing etc. These plastics pollute our environment and also reduces the quality of water. These pollutant affects the human life and also the aquatic organisms. The biggest challenge is to identify the object in the turbulence of water. In order to remove these various methods are used to detect the pollutants through object detection algorithms. This paper deals with overcoming the challenges of realizing the accurate pollutant detection techniques and provides a best method for real time garbage detection methods and its performance.**

Keywords—Deep learning, Floating Object, Intelligent Robot, Object detection, Water surface Cleaner.

I. Introduction

Water resources are potentially useful in multipurpose for human beings and other organisms. Water covers 71 percent of the earth's surface, but just 3 percent of that water is fresh. The water is contaminated by many unwanted substances which harms the organisms that relay on water. People relay on water for the purpose of drinking, cleaning, swimming and other essential activities. Although the crucial role automation plays in all industrial applications, properly disposing of sewage from industries and cleaning sewage remains a difficult task. The polluted water may affect the health of the organisms too. The major pollutants of water are industrial waste, oil spills, chemical fertilizers, wastewater and sewage. Humans pollute water in such a way that they throw plastics and other floating waste in water. Over decades, there has been an increase in the amount of water contamination brought on by human error. [1].

For every pollution there will always be several factor affecting it, likewise in case of water pollution the major factors affecting pollution are bacteria, viruses, parasites, weeds, chemicals from industries, food products and containers, and huge amount of plastics. Notably, plastics are the major cause of pollution which affects the lives of aquaticorganisms living in under water. major plastics do not degradeeven for years which affects all the living organisms in wateralso cause major health issues in humans. So this need for plastic removal is mandatory to lead a healthy life nowadays. Removing plastics from water bodies is not an easy task because of several factors affecting it such as some light weight plastics move along with the direction of airflow which is difficult to detect and collect the garbage from water [2]. Some heavyweight plastics which are submerged under water can cause toxicity for the animals and plants living in the water. Also, air flow plays the major role in plastic removalbecause the light weight plastics are carried out by the flow of air which cause hardship for the device to detect and collect the garbage as it should not leave any garbage whether it is static or moving in the water bodies.

The slippery and round shaped plastic bottles is a challenge to collect as the drag force usually makes the plastics flow away from the robot. Second, the robot is tele operated in a pond where Practical tests are used to gauge garbage collection capacity. Thirdly, the impact of matching the robot driving speed to the conveyor belt speed is investigated. To resist the disturbances that involves how to realize the real- time and accurate garbage detection while conduction of vision-based steering and despite the choppy water's surface, to grip the floating trash.

Detection: The moving robot makes a challenge to detect the plastics that flow away in the water surface. The object detection algorithm must be accurate and real time.

Navigation: The unpredictable dynamic factors such as wind and waves in the aquatic environment, that makes it necessary.

Collection: The objects will move on the surface of water due to the dynamic conditions and the robot itself. The plastics will move awayin according to the wind flow.

II. Need for Object Detection

DOI: rp-9788770040723.015

The various pollutants relay on water bodies which destroy the water and other organisms which depends on that water. So, the unwanted waste that float on water should be removed. In water it is difficult to remove waste as easy as possible due to the shaky movement of water. By human intervention activity it is not so easy to remove all those waste in case of large water bodies. Floating object Machine cannot work without any algorithms. We need to feed all the information in that machine in order to make it effective. The Fig 1. Shows the need for object detection. The first step tocollect waste is to detect waste. Therefore, there is a need forobject detection. There are many algorithms which give solution to this object detection with different accuracies. The primary goal of object detection is to identify occurrences ofa preset set of items and to use a bounding box to represent the locations of each object in the image.

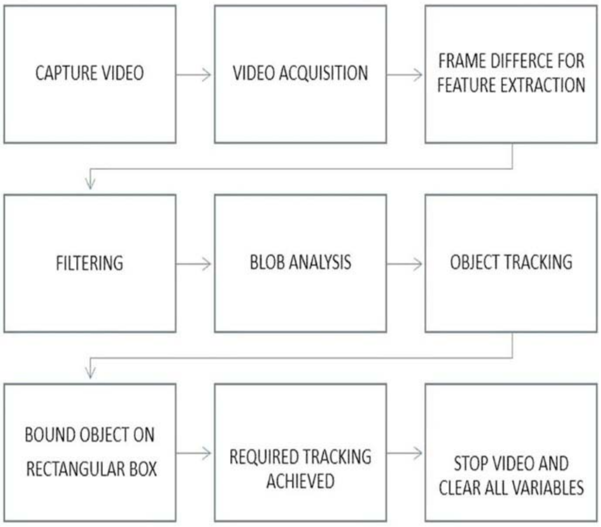

Fig. 1. Flowchart for need of Object Detection

IV. Object Detection

Detecting things in photos, movies, and other media is made possible by computer vision and image processing technology. The figure 2 shows the real time captured frame of image with floating garbage. AI is the fundamental concept behind object detection in this. Data is gathered by computer vision software and input into machine learning or deep. The camera gathers data from photos taken at various angles and transmits it to the image processing unit [4]. The findings are then presented based on the model's training data.

- Unfortunately, due to the saturation of hand-crafted features, object detection has reached a high after 2010. In 2012 the rebirth of convolutional neural and deep convolutional networks where successful at learning robust and representation of the image. The Convolutional Neural Network (CNN) has proposed and broken the deadlocks of object detection in 2014. Object detection has 2 genres: two stage detection and onstage detection.
- A set of object proposal starts with the extraction. The fixed sized image is fed into a pre trained CNN model

may not be in same position for a while, so by using waste boxes or with the hands of human is impossible [3]. So, to overcome these, a machine is to be invented. In case of plastic waste, the water may deposit inside the plastic bags which increases the weight of the waste so the machine needs to be designed which handle the weight of the waste.

to extract features. SVM classifiers are employed to identify object categories and forecast the existence of an item inside that region.

- R-CNN, SPP Net the image is processed at conv space layers for one time while the image is processed with R-CNN for as many times. However, it includes drawbacks like multistage and SPP Net ignores all preceding layers and solely fine-tunes its completely linked layers.
- In R-CNN, it is mostly unable to capture objects which are small. Pyramids can be used to solve this simple image to scale into different sizes. All the predictions can be combined when the detections are detected on each scale.

Fig. 2. Object Detection

IV. Proposed Methodologies

There are many methodologies that defines object detection in in a finer manner with different types of algorithms that detects objects and images. There exist many algorithms for object detection. some are very successful and some resulted in failures and some are still under testing state. Every algorithm has its own pros and cons. In spite of all drawbackssome notable object detection algorithms are also used widely which uses object detection and collection. some of the notable object detection algorithms are listed below:

A. CNN

CNN is a deep learning technique that stands for Convolutional Neural Network. Which is widely used in the few past years. It has high accuracy which is because used for image classification and recognition [5]. The computer scientist YannLecun proposed CNN algorithm in late 90’s.he also shared that he was inspired from human visual perception of recognizing things. The convolutional neural network also known as artificial neural network; it is specially designed to process data in pixel. It has built in layers through which the high dimensionality of images is

DOI: rp-9788770040723.015

reduced without losing the information. In simple CNN is a supervised deep learning method and widely used in image recognition and detection [6]. It has 4 layers of network (1) convolutional layer, (2) pooling layer, (3) ReLU correction layer (4) fully connected layer.

Steps Involved:

- The foremost step is to **choose a dataset** which is choosing the image for classification is the initial step.
- Next is to **prepare dataset for training** which mostly deals within creating paths, and categories or labelling and also in resizing the image.
- Third is to **create training** data which contains pixel values from the image list.
- Next is to **shuffle the dataset** and the fifth step is toassign labels and features.
- The following step is to normalize and transform labels to categorical data.
- The CNN model is then defined, compiled, and trained.

The main advantage of CNN is its automatic detection of important features which can be possible without any human supervision is shown in Fig.3. It is also appreciated by learning about the key features by itself. But the encoding the position and orientation of the object is not possible in CNN [7]. It requires lots and lots of training which can take more timeto complete. It lacks in the ability of shifting the input signalto equally shifted output signal.

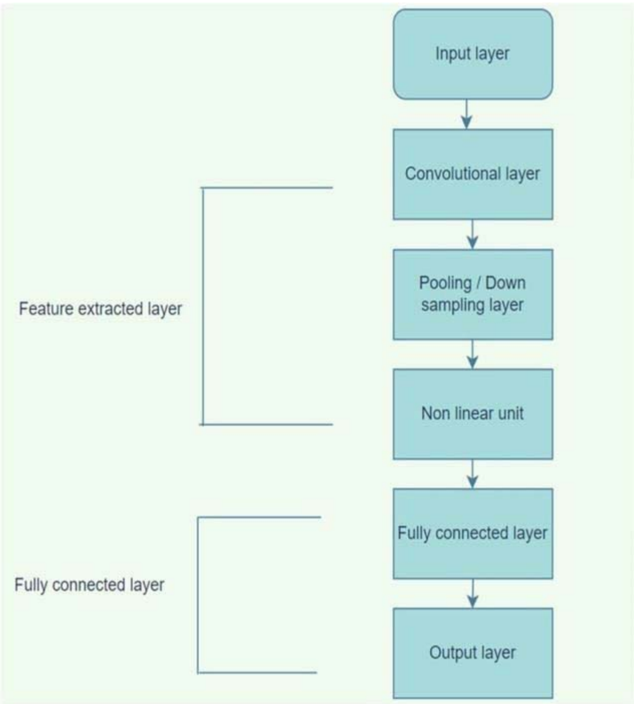

Fig. 3: Flowchart for CNN

B. Histogram of Oriented gradients (HOG)

The Fig 4. Shows the histogram frame works of image. Histogram of Oriented recognition gradients deals with face recognition that uses principal component analysis (PCA) to solve problems based on low accuracy. It is mostly used when the accuracy of the problem is low and it is used under nonrestrictive conditions. Most importantly the gradient orientation deals with the direction of greatest intensity change in the neighborhood of pixel. HOG uses features combined with linear classifier named SVM which is classification algorithm developed by V. Vapnik [8] along with his team. It is based on ideas of optimization and statistical learning. Both linear and nonlinear SVM training models are utilized for face recognition, and it is the primary approach for tackling pattern recognition problems.

Working steps in HOG

- Preprocessing: if a large image of size 720X475 is selected for a patch of size 100X200. The patch is cropped out of an image and resized to 64X128.
- Calculate gradient images: Sober operator is used to calculate the horizontal and vertical gradients
- Compute gradient histogram in 8X8 cells: the picture is divided into 8X8 cells, and the gradient is calculated for each 8X8 cell.
- 16X16 block normalization: A 16X16 matrix has four histograms that are combined to generate a 36X1 element vector.
- Calculate the histogram of oriented gradients feature vector: the concatenated vector is converted into one giant vector for calculation.

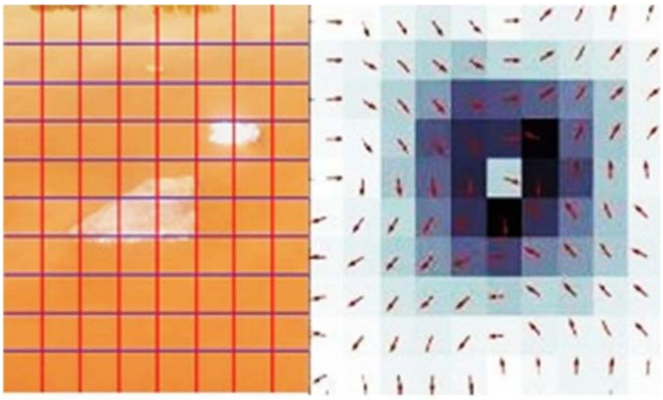

Fig. 4. Histogram Oriented Gradient

C. SSD

SSD stand for Single Shot Multi-Box Detector Single stage object detection method. Most of the object detection method detects the object by dividing it onto sliding windows, but SSD algorithm divides the object into grid cells which is responsible for detecting the object in that region of image. Detection of particular object is done by predicting thelocation of an object within that region. On different tiers of the output region, it employs default boxes of varied sizes, shapes, and aspect ratios. For better coverage of location, it uses 8732 boxes, scale and aspect ratios [9].

The majority of the predictions will be devoid of any item. It discards forecasts with a confidence value of less than 0.01. The main components of SSD are backbone model and SSD head it is shown Fig.5. The backbone model is a pretrained classification of image [10]. It is also like YOLO algorithm but it took only one shot to detect more

DOI: rp-9788770040723.015

than one object present in an image using multi-Box. To make this algorithm effective, training is mandatory. So, the input to SSD is givenas an input image with ground truth bounding boxes. During training, the default boxes are matched to the ground truth boxes in terms of aspect ratio, position, and scale. The intersection between predicted boxes and ground truth should be bigger than 0.5. Finally, we select the projected box with the greatest overlap with the ground truth. Thus, the object is detected using SSD algorithm.

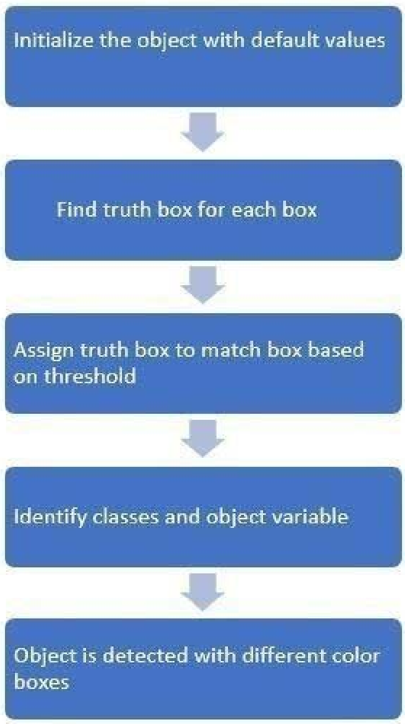

Fig. 5. Flowchart for SSD

D. YOLO (You Only Look Once)

YOLO is the most popular algorithms for the detection of objects and images the processing speed of YOLO is very fast that it is an algorithm to detect real time objects very fast. It produces high accuracy and overall processing speed which makes it popular [11],[12]. YOLO algorithm predicts the object and the location as bounding box of the input image. Using the four numbers it recognizes each bounding box; center, width and height of the bounding box.

Working of YOLO

At first the algorithm divides image into grids like 3x3 or 4x4, with the help of the grid it detects by objects per grid than objects per image, it encodes a vector for each grid to describe that cell. If a grid does not have any object, it is encoded as,

V1,1= (Pc, Bx, By, Bw, Bh, V1, V2) = (0, _, _, _, _, _) (1)

There is a possible that one issue might happen, if the algorithms predicts large number of bounding boxes for one class, to resolve this there is non-max suppression algorithmthat at first the box with maximum probabilityis considered[13] and it is compared with all the other boxes in that class with the use of Intersection Over Union (IoU), shown in equation 2.

$$\text{IOU} = \frac{\text{Area of intention between B1 and B2}}{\text{Area of union between B1 and B2}} \quad (2)$$

If the resultant IoU is exceeds the threshold, the box that has small probability is ignored. This process is continued and repeated until all the boxes are either predicted or excluded. The Fig. 6 shows the grid and cell vectors.

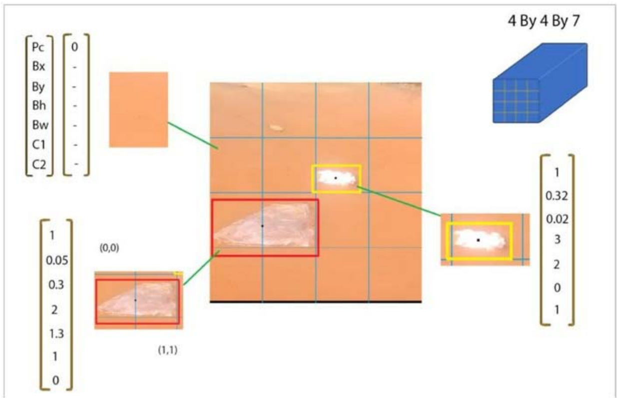

Fig. 6 Object detection using YOLO

V. Comparitive Analysis

Comparative analysis is the basic factor which defines the best among all. It is necessary to conduct comparative analysis to get a clear understanding of the problem or resultant relevant queries [14]. Notably when it comes to object detection, the chosen algorithm must be very effective to detect all the floating garbage's in spite of turbulence like airflow, heavy wind etc., because the algorithm should be trained to detect static as well as moving garbage's in water. In this paper, comparative analysis deals with finding the best object detection algorithm among the selected algorithms like CNN, SSD, YOLO, HOG [15]-[17]. Here the comparison is done based on the performance i.e., speed of the detecting algorithm, accuracy rate which defines how much accurate the algorithm works, and also with the speed of detecting the objects in the water surface. In addition to this analysis, the time taken for training is also a factor which makes this comparison effective.

TABLE 1: Cooperation of Object Detection Algorithms

ALGORITHMS	SPEED	ACCURACY RATE	TIME FOR TRAINING (hrs)	PERFOMANCE (Mb)
CNN	Slow	93.8	18	115
HOG	Medium	90.18	16	200
SSD	Medium	86.60	14	225
YOLO	High	95.02	10	25

Table 1 shows the comparative analysis of the above-mentioned algorithms

DOI: rp-9788770040723.015

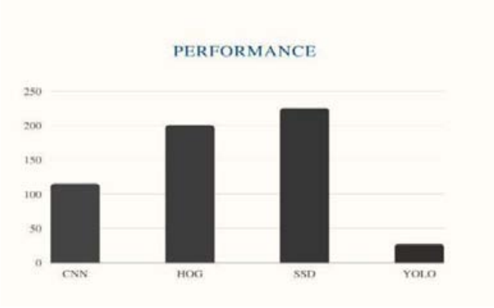

Fig. 7 Performance of Algorithms

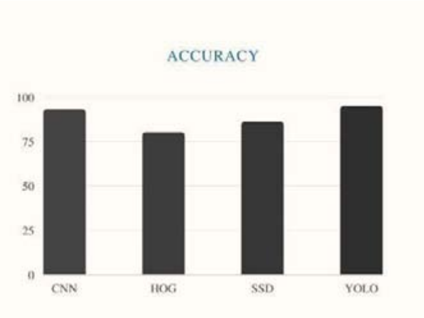

Fig. 8 Accuracy of Algorithms

Fig. 8 Time taken for training

From Figure 7, Figure 8, Figure 9 it is inferred that YOLO and SSD are efficient in comparison with CNN, HOG.

VI. Conclusion

This paper deals with introduction to the algorithms used for detecting objects and clarifies the need for object detection and a comparative analysis is done by comparing different parameters which results in the removal of floating garbage in surface of water. With these measured factors, the final result shows that YOLO algorithm is very effective than the other algorithms mentioned above. Yolo is also widely used nowadays for object detection with updated version releases.

Reference

[1] M. Barenboim, M. Masso, I. I. Vaisman and D. C. Jamison, "Statistical Geometry Based Prediction of Nonsynonymous SNP Functional Effects Using Random Forest and Neuro-fuzzy Classifiers", Proteins: Structure Function and Bioinformatics, vol. 71, no. 4, pp. 1930-1939, 2008.

[2] M. Barenboim, M. Masso, I. I. Vaisman and D. C. Jamison, "Statistical Geometry Based Prediction of Nonsynonymous SNP Functional Effects Using Random Forest and Neuro-fuzzy Classifiers", Proteins: Structure Function and Bioinformatics, vol. 71, no. 4, pp. 1930-1939, 2008.

[3] B. Goertzel, "Are there deep reasons underlying the pathologies of today's deep learning algorithms?", Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 9205, pp. 70-79, 2015.

[4] A.Shashua, Y.Gdalyahu, and G. Hayon, "Pedestrian detection for driving assistance systems: Single-frame classification and system level performance," In Proceedings of IEEE Intelligent Vehicles Symposium, 2004.

[5] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," In C. Schmid, S. Soatto, and C. Tomasi, editors, International Conference on Computer Vision and Pattern Recognition, vol. 2, PP. 886-893, June 2005.

[6] N.Bellas, S. M.Chai, M.Dwyer, and D.Linzmeier, "FPGA implementation of a license plate recognition SoC using automatically generated streaming accelerators," Parallel and Distributed Processing Symposium, 2006. IPDPS 2006. 20th International, pp. 8, 25-29 April 2006.

[7] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Fu, and A.C. Berg, "SSD: Single Shot Multibox Detector", Springer, vol.9905, pp.21-37, 2016.

[8] Z. Shen, Z. Liu, J. Li, Y. Jiang, Y. Chen, and X. Xue, "DSOD: Learning Deeply Supervised Object Detectors from Scratch", IEEE International Conference on Computer Vision (ICCV), Venice, Italy, pp.1937-1945, 2017.

[9] S. Zhai, D. Shang, S. Wang, and S. Dong, "DF-SSD: An Improved SSD Object Detection Algorithm Based on DenseNet and Feature Fusion", IEEE Access, vol.8, pp.24344-24357, 2020.

[10] Dhanabalan, S. S., Sitharthan, R., Madurakavi, K., Thirumurugan, A., Rajesh, M., Avaninathan, S. R., & Carrasco, M. F. (2022). Flexible compact system for wearable health monitoring applications. Computers and Electrical Engineering, 102, 108130.

[11] J.Redmon, and A. Farhadi, "YOLOv3: An Incremental Improvement", ResearchGate, 2018.

[12] J. Sang, Z. Wu, P. Guo, H. Hu, H. Xiang, Q. Zhang, ndB. Cai, "AnImproved Yolov2 for Vehicle Detection", Sensors, vol.18, issue.12, pp.4272, 2018.

[13] Gomathy, V., Janarthanan, K., Al-Turjman, F., Sitharthan, R., Rajesh, M., Vengatesan, K., &Reshma, T. P. (2021). Investigating the spread of coronavirus disease via edge-AI and air pollution correlation. ACM Transactions on Internet Technology, 21(4), 1-10.

[14] Julus, L. Jerart, and Krishnan Ayyappan. "Security-Enhanced Cloud for Serverless Computing and Its Applications," Privacy and Security Challenges in Cloud Computing. CRC Press, pp. 1-16.

[15] Julus, Jerart L., et al. "Advanced nonlinear equalizer for Filter Bank Multicarrier-based Long Reach-Passive Optical Network system," International Journal of Communication Systems 34.14, 2021.

[16] Julus, L. Jerart, D. Manimegalai, and S. SibiChakkaravarthy. "FBMC-based dispersion compensation using artificial neural network equalization for long reach-passive optical network," International Journal of Wavelets, Multiresolution and Information Processing 18.01,p. 1941011, 2020.