Comparison of algorithms for the detection of marine vessels with machine vision

José Rodríguez-Gonzales, Junior Niquin-Jaimes, Ernesto Paiva-Peredo
Department of Electrical Engineering, School of Mechatronics Engineering, Universidad Tecnológica del Perú, Lima, Peru

Article Info

Article history:

Received Jan 21, 2024
Revised Jul 27, 2024
Accepted Aug 6, 2024

Keywords:

Algorithms
Detection
Machine learning
Neural networks
Ships
YOLO

ABSTRACT

The detection of marine vessels for revenue control has many tracking deficiencies, which has resulted in losses of logistical resources, time, and money. However, digital cameras are not fully exploited since they capture images to recognize the vessels and give immediate notice to the control center. The analyzed images go through an incredibly detailed process, which, thanks to neural training, allows us to recognize vessels without false positives. To do this, we must understand the behavior of object detection; we must know critical issues such as neural training, image digitization, types of filters, and machine learning, among others. We present results by comparing two development environments with their corresponding algorithms, making the recognition of ships immediately under neural training. In conclusion, it is analyzed based on 100 images to measure the boat detection capability between both algorithms, the response time, and the effectiveness of an image obtained by a digital camera. The result obtained by YOLOv7 was 100% effective under the application of processing techniques based on neural networks in convolutional neural network (CNN) regions compared to MATLAB, which applies processing metrics based on morphological images, obtaining low results.

Corresponding Author:

Ernesto Paiva-Peredo
Faculty of Engineering, Universidad Tecnológica del Perú
125 Natalio Sanchez Street, Santa Beatriz Urbanization, Cercado de Lima, Lima, Peru
Email: c22395@utp.edu.pe

1. INTRODUCTION

Currently, technology has allowed us to create systems designed to detect objects, which can be oriented to environmental monitoring of the oceans [1]–[3]. Likewise, object detection has developed, which allows us to implement it in different fields and areas where machine vision is beneficial [4]–[8]. First, one of the application areas is related to security systems through images obtained from digital surveillance cameras; these cameras capture information analogically stored and processed with a microprocessor, which transforms to digital signals to analyze them [9]. Secondly, it is oriented in medicine to obtain early diagnoses of diseases. Through the processing of X-ray images and computed tomography (CT) scans, the medical professional can make correct decisions for the final diagnosis [10]. Thirdly, we can find industrial automation processes that help better results in managing manufacturing control, object classification, verification, and detection of product defects [11], [12]. Finally, it is used in the guidance of autonomous vehicles, where cameras capture images at different angles and in real-time, as it allows obstacle detection, safe navigation, and traffic sign recognition, among others [13], [14].

The ship detection methods from images obtained from different spatial orientations have been aimed at analyzing the dimensions of ships for detection [5], [8]. For example, we found articles dealing with ship detection from satellite images [7]. In addition, this information on the detection and tracking of objects in the marine environment is based on real-time video recordings, where they work with a dynamic subtraction of backgrounds; the algorithm it uses allows it to draw differences from captured images to focus and track the target [8]. On the other hand, the analysis of the different ship detection methods starts from aerial platforms, where digital cameras capture the images to analyze them and validate the presence of ships in the area [15]. From these investigations, it is concluded that the methods used to detect vessels are based on orientation, which confirms the presence of ships at sea, in addition to monitoring fishing areas, marine displacement control, and search and rescue of vessels.

The capture of images from an unmanned surface vehicle presents an overly broad area of research with great projections for the future. Depending on the project's scope, two types of images can be used to detect objects in marine images: red, green, and blue (RGB) and infrared. In that sense, the RGB images are the most suitable for having the most detailed object detection area since it allows tracking with high accuracy [5], [16], [17]. On the other hand, infrared images present a better response when working in dark or cloudy environments, but the quality of details is deficient [18], [19]. Finally, it was detected that most research works employ complex methods for tracking marine vehicles with images, leaving aside the detection of objects used by morphological image processing commands.

From this premise, two object detection methods are compared: the method of morphological image processing and the other by image detection algorithm based on neural networks that can be trained. First, filtering, eroding, and binarizing algorithms are used in the MATLAB processing environment to highlight the presence of a ship that was taken by a digital camera [20]. Second, the image captured by the digital camera is loaded into the YOLOv7 environment, which has been previously trained to recognize objects [21]. Finally, the results will be compared based on 100 samples processed by both algorithms.

2. METHOD

2.1. Ship detection using MATLAB algorithms

2.1.1. Capturing an image and changing to grayscale

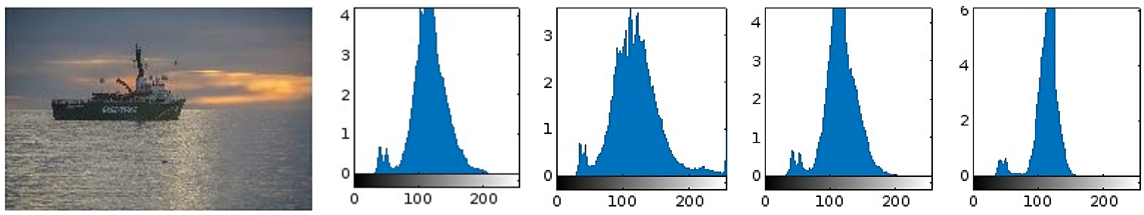
Initially, the program environment loads the image, depicted in Figure 1(a). To facilitate further analysis, we transition the image into a color scale, as demonstrated in Figure 1(b). Utilizing weighted grayscale enhances our ability to scrutinize color layers in greater detail; as depicted in Figures 1(c) to (e), we observe the grayscale representations of individual color layers. Figure 1 illustrates the image analysis based on layers, featuring Figure 1(a) showcasing the original RGB image devoid of any filters; Figure 1(b) exhibiting the image of the seabed transformed into weighted grayscale; Figure 1(c) portraying the darkened rendition of the seabed in the red layer; Figure 1(d) revealing the seabed image with light colors removed in the green layer; and Figure 1(e) displaying the seabed image with light colors eliminated in the blue layer.

Figure 2 analyzes the distortion behavior concerning the grayscale. Figure 2(a) shows the original image in RGB, which will be processed. On the other hand, Figure 2(b) shows the image's histogram converted to grayscale. The filter was applied from 0-200 in the grayscale application. Finally, the results of the histogram of the red layer (the applied filter ranges from 0-250), green layer (the used filter is reduced from 0-190), and blue layer (filter applied is reduced from 0-150) are presented in Figures 2(c) to (e), respectively.

(a) (b) (c)

(d) (e)

Figure 1. Image analysis based on layers: (a) original image without filters applied, (b) the image of the seabed applied to grayscale, (c) the image of the seabed in the process of darkening, (d) the image of the seabed without the light colors, and (e) the image of the seabed without light colors



(a) (b) (c) (d) (e)

Figure 2. Representation when applied to grayscale: (a) original image, (b) gray weighted, (c) red cape, (d) green cape, and (e) blue cape

2.1.3. Image thresholding

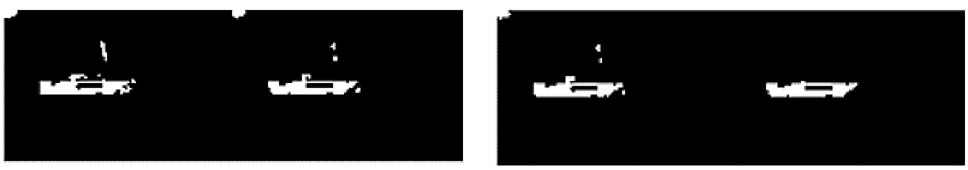
The original image undergoes grayscale filtering, applying a threshold where gray tones below 60 are eliminated, removing pixels exceeding this threshold. Figure 3 illustrates the outcomes of these procedures alongside the elimination of non-object pixels. In Figure 3(a), we observe the original image captured by the digital camera in RGB colors. Following this, in Figure 3(b), the applied layer showcases the image converted to grayscale. During this conversion process, a threshold of 60 is set for enhanced noise extraction, as depicted in Figure 3(c). The resulting images are presented to demonstrate the progression of image thresholding.

(a) (b) (c)

Figure 3. Image thresholding: (a) no filters applied, (b) image with the filter applied to grayscale, and (c) image used with the threshold of 60

2.1.4. Object filter

The image harbors external noises requiring removal through a noise extraction process based on a predetermined threshold. In Figure 4, it is crucial to ensure a clean process to minimize the risk of false positives during detection. Figure 4(a) illustrates the initial stage of noise extraction, known as opening, where noise analysis and extraction occur. Subsequently, in Figure 4(b), labeled extraction, the final elimination of noise or impurities is executed to facilitate accurate vessel detection. Figure 4 depicts the sequential steps of the noise cleaning process, emphasizing the importance of meticulous execution to prepare for vessel survey efficiently.



(a) (b)

Figure 4. Noise cleaning process: (a) noise analysis and extraction and (b) final preparation for vessel survey

2.1.5. Recognition of the ship

The image treated by filters recognizes the vessel. The identification process was noise extractions. During the process, a blue rectangle was established, as shown in Figure 5, where the target to be identified, in this case, the ship, is highlighted.

Figure 5. Vessel survey result

2.2. Vessel detection using the YOLOv7 algorithm
2.2.1. Algorithm training

To train an algorithm in YOLOv7, a database of at least 500 images for training and their respective labels of the object to be detected is necessary. Subsequently, the learning framework and the number of training epochs are defined, and the result of executing the training code will not deliver a neural network capable of recognizing objects like those of the labels of the training images [21], [22]. Subsequently, in the test phase, 100 images are set for processing and ship detection, obtained out of the 500 training images.

2.2.2. Image input

To effectively utilize the YOLOv7 algorithm, it is imperative to understand its functionality as a neural network specializing in bounding box predictions. Object detection through YOLO involves leveraging neural network architectures that employ learned models for detection tasks [23], [24]. These models, including Faster R-CNN and Mask R-CNN, excel in locating multiple objects within images, often generating pixel-level masks [25]. However, challenges such as noise and class imbalances persist in this process.

The chosen environment for implementing the algorithm is CONDA, a programming platform facilitating interaction with Python for executing object recognition tasks. Preceding its deployment, the algorithm undergoes training based on ship labels assigned within the image dataset [26]. Figure 6 delineates the sequential steps in training the neural network and detecting objects. It outlines key components such as the dataset for training images, the CONDA programming environment for algorithm development, the YOLOv7 neural network version applied, the direct image input captured by a digital camera, the training phase for the neural network, and the eventual detection results.

2.2.3. Execution of the detection algorithm

Before initializing the stopping process, you must enter the address of the image where you want to detect the object. Within a local environment, we must have a pre-trained model since the data set must be related to vessels for detection. Finally, after running the algorithm, the resulting image will look like Figure 7, where a box demarcates the vessel.

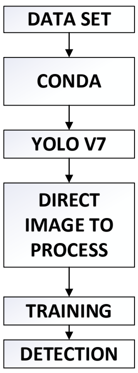


Figure 6. Process diagram for using YOLO

Figure 7. Result of vessel survey

3. RESULTS AND DISCUSSION

A comparison of the two methods for object detection was carried out, and the objective was to determine the accuracy and recognition time. The evaluation has led to different results; both times are contradictory concerning the detected, with a longer waiting time in recognition but with a successful outcome. The parameters to be considered for the evaluation are object detection time and detected area [27].

3.1. Object detection time

Response time is a crucial parameter in detection systems, reflecting the algorithm's efficiency. However, this parameter's significance is magnified when accompanied by precise detections, rendering it a reliable reference point. To assess response times accurately, an evaluation was conducted on 100 images distinct from the 500 used for training the network.

Figure 8 compares the image processing times between the two algorithms. The depicted values pertain to the processing times of these 100 distinct images. Specifically, "Time M(s)" denotes the response time within the MATLAB environment, while "Time Y(s)" represents the response time within the CONDA environment employing the YOLO neural network. Notably, while one algorithm may exhibit longer waiting times, the detection accuracy tends to be higher, aligning more closely with the desired targets [28]. These results underscore the trade-offs between processing speed and detection accuracy, highlighting the complexities inherent in algorithmic performance evaluation [29].

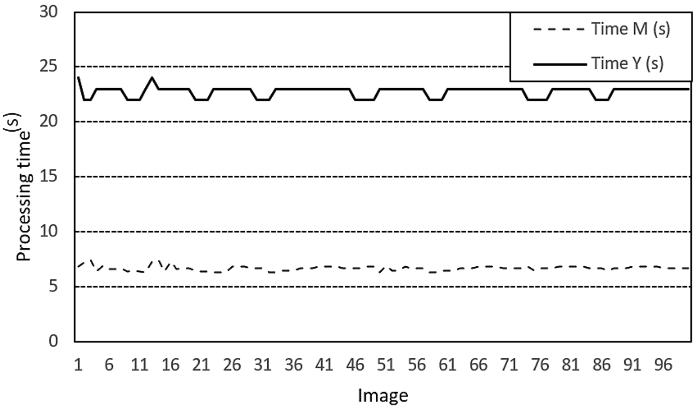


Figure 8. Comparison of image processing time

3.2. Detected area

To determine the correctness of image detection, a criterion is established where highlighting over half of the objects is deemed successful detection (assigned a value of 1) while highlighting less than half is considered unsuccessful (assigned a value of 0). Figure 9 compares the detection effectiveness between two algorithms based on evaluations conducted on 100 images distinct from the training dataset.

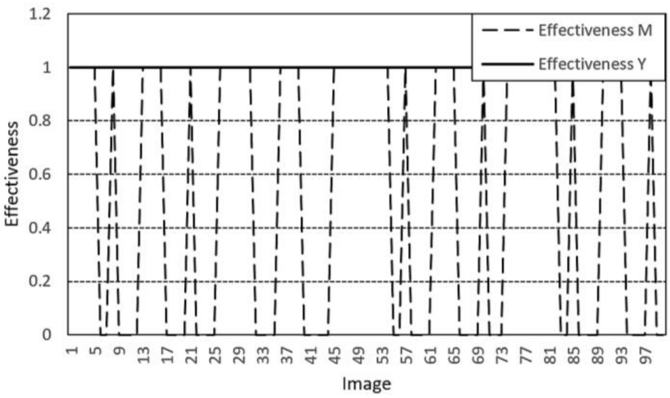


Figure 9. Comparison of detection effectiveness of the two algorithms

In Figure 9, "Effectiveness M" represents the response in the MATLAB environment, while "Effectiveness Y" signifies the response in the CONDA environment utilizing the YOLO neural network. The displayed values indicate the effectiveness of ship detection for each algorithm. Remarkably, the results demonstrate that the YOLO algorithm outperforms the MATLAB algorithm regarding detection effectiveness. This observation underscores the superiority of the YOLO algorithm in accurately detecting ships within images.

4. CONCLUSION

This study addressed the problem of detecting marine vessels for income control, highlighting the deficiencies in existing monitoring methods. A solution was proposed based on the underutilized use of digital cameras and the use of neural networks. Comparison between the two development environments and their respective algorithms revealed that the YOLOv7 algorithm, which incorporates neural network processing, outperforms MATLAB based on morphological image processing metrics. The evaluation focused on object detection time and detected area as critical parameters.

Object detection time analysis showed longer wait times in the YOLO algorithm but with more accurate and successful detection results. Despite the extended processing times, the YOLO algorithm outperforms MATLAB in terms of efficiency and accuracy in object detection. The evaluation of the detected area highlighted the superiority of the YOLO algorithm compared to MATLAB. The effectiveness of vessel detection was more prominent in the YOLO algorithm, providing more satisfactory and accurate results in more than 50% of the cases evaluated.

In summary, this study demonstrates the effectiveness of the YOLOv7 algorithm implementation in marine vessel detection compared to MATLAB. Despite the longer processing times, the YOLO algorithm offers more accurate and efficient results in terms of object detection. This work lays the foundation for improving revenue control through advanced vessel tracking technologies.

REFERENCES

[1] J. Woo, C. Yu, and N. Kim, "Deep reinforcement learning-based controller for path following of an unmanned surface vehicle," *Ocean Engineering*, vol. 183, pp. 155–166, Jul. 2019, doi: 10.1016/j.oceaneng.2019.04.099.

[2] Q. Cai, Q. Wang, Y. Zhang, Z. He, and Y. Zhang, "LWDNet-a lightweight water-obstacles detection network for unmanned surface vehicles," *Robotics and Autonomous Systems*, vol. 166, Aug. 2023, doi: 10.1016/j.robot.2023.104453.

[3] K. F. Hansen, L. Yao, K. Ren, S. Wang, W. Liu, and Y. Liu, "Image segmentation in marine environments using convolutional LSTM for temporal context," *Applied Ocean Research*, vol. 139, Oct. 2023, doi: 10.1016/j.apor.2023.103709.

[4] Q. Yu and Y. Su, "Local defogging algorithm for the first frame image of unmanned surface vehicles based on a radar-photoelectric system," *Journal of Marine Science and Engineering*, vol. 10, no. 7, Jul. 2022, doi: 10.3390/jmse10070969.

[5] Z. Liu, Y. Zhang, X. Yu, and C. Yuan, "Unmanned surface vehicles: an overview of developments and challenges," *Annual Reviews in Control*, vol. 41, pp. 71–93, 2016, doi: 10.1016/j.arcontrol.2016.04.018.

[6] H. Kim, D. Kim, and S.-M. Lee, "Marine object segmentation and tracking by learning marine radar images for autonomous surface vehicles," *IEEE Sensors Journal*, vol. 23, no. 9, pp. 10062–10070, May 2023, doi: 10.1109/JSEN.2023.3259471.

[7] D. K. Prasad, D. Rajan, L. Rachmawati, E. Rajabally, and C. Quek, "Video processing from electro-optical sensors for object detection and tracking in a maritime environment: a survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 8, pp. 1993–2016, Aug. 2017, doi: 10.1109/TITS.2016.2634580.

[8] U. Kanjir, H. Greidanus, and K. Oštir, "Vessel detection and classification from spaceborne optical images: a literature survey," *Remote Sensing of Environment*, vol. 207, pp. 1–26, Mar. 2018, doi: 10.1016/j.rse.2017.12.033.

[9] Z. Hu, J. Zhou, and E. Zhang, "Improving traffic safety through traffic accident risk assessment," *Sustainability*, vol. 15, no. 4, Feb. 2023, doi: 10.3390/su15043748.

[10] R. H. Moss *et al.*, "Skin cancer recognition by computer vision," *Computerized Medical Imaging and Graphics*, vol. 13, no. 1, pp. 31–36, Jan. 1989, doi: 10.1016/0895-6111(89)90076-1.

[11] Y. Tang *et al.*, "Optimization strategies of fruit detection to overcome the challenge of unstructured background in field orchard environment: a review," *Precision Agriculture*, vol. 24, no. 4, pp. 1183–1219, Aug. 2023, doi: 10.1007/s11119-023-10009-9.

[12] T. Luo *et al.*, "Classification of weed seeds based on visual images and deep learning," *Information Processing in Agriculture*, vol. 10, no. 1, pp. 40–51, Mar. 2023, doi: 10.1016/j.inpa.2021.10.002.

[13] L. Wang *et al.*, "A review of methods for ship detection with electro-optical images in marine environments," *Journal of Marine Science and Engineering*, vol. 9, no. 12, Dec. 2021, doi: 10.3390/jmse9121408.

[14] Z. Shao *et al.*, "Multi-scale object detection model for autonomous ship navigation in maritime environment," *Journal of Marine Science and Engineering*, vol. 10, no. 11, Nov. 2022, doi: 10.3390/jmse10111783.

[15] M. F. Fingas and C. E. Brown, "Review of ship detection from airborne platforms," *Canadian Journal of Remote Sensing*, vol. 27, no. 4, pp. 379–385, Aug. 2001, doi: 10.1080/07038992.2001.10854880.

[16] W. Han, A. Kuerban, Y. Yang, Z. Huang, B. Liu, and J. Gao, "Multi-vision network for accurate and real-time small object detection in optical remote sensing images," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022, doi: 10.1109/LGRS.2020.3044422.

[17] A. Zhou, W. Xie, and J. Pei, "Infrared maritime target detection using the high order statistic filtering in fractional Fourier domain," *Infrared Physics & Technology*, vol. 91, pp. 123–136, Jun. 2018, doi: 10.1016/j.infrared.2018.04.006.

[18] D. Zhou, S. Qiu, Y. Song, and K. Xia, "A pedestrian extraction algorithm based on single infrared image," *Infrared Physics & Technology*, vol. 105, Mar. 2020, doi: 10.1016/j.infrared.2020.103236.

[19] Z. Liu *et al.*, "Infrared salient object detection based on global guided lightweight non-local deep features," *Infrared Physics &*

Technology, vol. 115, Jun. 2021, doi: 10.1016/j.infrared.2021.103672.
[20] J. An *et al.*, "Enhancement of marine lantern's visibility under high haze using AI camera and sensor-based control system," *Micromachines*, vol. 14, no. 2, Jan. 2023, doi: 10.3390/mi14020342.
[21] X. Han, L. Zhao, Y. Ning, and J. Hu, "ShipYOLO: an enhanced model for ship detection," *Journal of Advanced Transportation*, vol. 2021, pp. 1–11, Jun. 2021, doi: 10.1155/2021/1060182.
[22] J.-Y. Shen, C.-K. Lu, and L. G. Lim, "A novel deep convolutional neural network pooling algorithm for small floating objects detection," in *2023 International Conference on Consumer Electronics - Taiwan (ICCE-Taiwan)*, Jul. 2023, pp. 175–176, doi: 10.1109/ICCE-Taiwan58799.2023.10226754.
[23] H. Fu, R. Zhang, X. Ning, and Y. Wang, "Ship detection based on improved YOLO algorithm," in *2021 40th Chinese Control Conference (CCC)*, Jul. 2021, pp. 8181–8186, doi: 10.23919/CCC52363.2021.9549967.
[24] H. D. Trung, "A comparison of YOLO networks for ship detection and classification from optical remote-sensing images," in *International Conference on Hybrid Intelligent Systems*, 2023, pp. 43–52, doi: 10.1007/978-3-031-27409-1_5.
[25] Y. Zhou, "A YOLO-NL object detector for real-time detection," *Expert Systems with Applications*, vol. 238, Mar. 2024, doi: 10.1016/j.eswa.2023.122256.
[26] S. Fefilatyev, D. Goldgof, M. Shreve, and C. Lembke, "Detection and tracking of ships in open sea with rapidly moving buoy-mounted camera system," *Ocean Engineering*, vol. 54, pp. 1–12, Nov. 2012, doi: 10.1016/j.oceaneng.2012.06.028.
[27] J. Yu, D. Huang, X. Shi, W. Li, and X. Wang, "Real-time moving ship detection from low-resolution large-scale remote sensing image sequence," *Applied Sciences*, vol. 13, no. 4, Feb. 2023, doi: 10.3390/app13042584.
[28] A. P. W. Aldryani *et al.*, "Development of ship detection using OPENCV YOLO method on unmanned prototype boat for monitoring national sea," in *2022 7th International Workshop on Big Data and Information Security (IWBIS)*, Oct. 2022, pp. 7–12, doi: 10.1109/IWBIS56557.2022.9924891.
[29] M. Kumar and R. Bhatt, "An approach towards real-time object detector using open CV," in *2022 11th International Conference on System Modeling & Advancement in Research Trends (SMART)*, Dec. 2022, pp. 981–985, doi: 10.1109/SMART55829.2022.10046751.

BIOGRAPHIES OF AUTHORS

José Rodríguez-Gonzales engineering student at the Universidad Tecnológica del Perú. Currently, he is an electronic technician for the engineering area of an electronic security company. He can be contacted by e-mail: 1630348@utp.edu.pe.

Junior Niquin-Jaimes Engineering student at the Universidad Tecnológica del Perú. He currently works as a technical specialist at telefónica. His research interests include energy systems, artificial intelligence, and mechatronic systems. He can be contacted by e-mail: 1612636@utp.edu.pe.

Ernesto Paiva-Peredo graduated as an electrical mechanical engineer from the Universidad de Piura in Peru in 2013. He pursued further studies and obtained a master's degree in electrical and mechanical engineering, specializing in automation and optimization. He was supported by CONCYTEC and awarded by the Universidad de Piura in 2016. During his academic journey, he was a research assistant at the Department of Technology and Innovation (DTI) at SUPSI. Presently, he holds the position of professor-researcher at the Universidad Tecnológica del Perú. He can be contacted by e-mail: c22395@utp.edu.pe.