

Maritime Surveillance by Multiple Data Fusion: An Application Based on Deep Learning Object Detection, AIS Data and Geofencing

Sergio Ballines-Barrera¹^a, Leopoldo López¹^b, Daniel Santana-Cedrés²^c and Nelson Monzón^{1,2}^d

¹Qualitas Artificial Intelligence and Science, Spain

²CTIM, Instituto Universitario de Cibernética, Empresas y Sociedad, University of Las Palmas de Gran Canaria, Spain

www.qaisc.com

Keywords: Object Detection, AIS, PTZ cameras, Deep Learning, Geofencing, Maritime Environment.

Abstract: Marine traffic represents one of the critical points in coastal monitoring. This task has been eased by the development of Automatic Identification Systems (AIS), which allow ship recognition. However, AIS technology is not mandatory for all vessels, so there is a need for using alternative techniques to identify and track them. In this paper, we present the integration of several technologies. First, we perform ship detection by using different camera-based approaches, depending on the moment of the day (daytime or nighttime). From this detection, we estimate the vessel's georeferenced position. Secondly, this estimation is combined with the information provided by AIS devices. We obtain a correspondence between the scene and the AIS data and we also detect ships without VHF transmitters. Together with a geofencing technique, we introduce a solution that fuses data from different sources, providing useful information for decision-making regarding the presence of vessels in near-shore locations.

1 INTRODUCTION


Seaborne trade encompasses about 90 % of world trading (Kaluza et al., 2010). Despite the outbreak of the pandemic, it was affected just by a 3.8 % during 2020, with a recovery of 4.3 % in 2021 (Sirimanne et al., 2021). Hence, coastal regions, in particular those with port facilities, are clearly affected by a large influx of boats throughout the year. As an example, in 2021 the Port of Shanghai alone handled a volume of about 47 million TEU (*Twenty-foot Equivalent Unit*) in containers¹. Furthermore, if we consider the transit of other sorts of vessels (refueling, repairing, sports, or leisure), their number rises rapidly. In this regard, the importance of an accurate harbor traffic monitoring is quite relevant.


Automatic Identification Systems (AIS) provides the vessel position surrounding harbors and port authorities, although their use is not mandatory for all maritime vehicles. Consequently, the monitoring


tasks become arduous, complicating the surveillance of nearby waters. In this sense, we highlight the importance of using additional technologies, intending to complement and fill in the weaknesses of the current vessel monitoring systems.


The location of objects captured by cameras has been tackled through different Computer Vision techniques. From background subtraction techniques (Arshad et al., 2010; Hu et al., 2011) to optical flow algorithms (Li and Man, 2016; Larson et al., 2022), these approaches have been widely used in different environments, including coastal scenes. Object recognition in seaside scenes has been improved by several proposals (Zhang et al., 2021; Wang et al., 2022), albeit it remains considered an unsolved problem. Even with advances in Neural Networks, the most promising computational trend of Computer Vision in recent years, there are still limits to challenges that have not been overcome yet.

In this work, we integrate the accuracy of AIS together with the versatility of Neural Networks, supplied with images from a PTZ (*Pan-Tilt-Zoom*) camera. The main objective is to attain a broader awareness of the port environment, improving security in nearby waters. Some approaches (Bloisi et al., 2012; Simonsen et al., 2020) have proposed similar

^a <https://orcid.org/0000-0002-3327-5137>

^b <https://orcid.org/0000-0002-7066-4393>

^c <https://orcid.org/0000-0003-2032-5649>

^d <https://orcid.org/0000-0003-0571-9068>

¹Top 20 container ports

schemes but we use a more recent CNN's architecture (Bochkovskiy et al., 2020). Besides, our night approach is simple but also cheaper due to the fact that we only use an optical camera and no thermal information which cost is not negligible. In (Fu et al., 2021a; Fu et al., 2021b) were proposed an architecture based on YOLOv4 fused with a convolutional attention module. As said, our approach also detects vessels in optical nocturnal scenes. Also, our solution provides the vessels' location in georeferenced perimeters nearby the ports.

Additionally, in coastal regions that may suffer from poor connectivity, it is necessary that our AI strategies work efficiently on remote systems (Satyanarayanan, 2017; Zhao et al., 2019). Hence, our developments are carried out under the premise that they must be embedded in edge computing systems.

This research work is the first step in the strategy to achieve the future QAISC AI Coastal Surveillance Solution that is currently developed and validated in two complementary work environments:

- The Port of Las Palmas de Gran Canaria² where the ICTS PLOCAN³ has set up a sea observation station using HF radar technology and an optical camera. This environment is helping us to develop and validate approaches against very low visibility due to calima and haze situations representative of the South Atlantic and the coast of Africa and the Middle East.
- The Port of Gijón⁴, where a sea observation station with an optical camera has been recently installed and where the environment provides situations of very low visibility due to rain, fog and other complex weather conditions representative of the North Atlantic.

The rest of this paper is organized as follows: in section 2 we present the related works regarding the use of AIS systems and cameras to identify vessels. Afterward, section 3 includes the presented techniques. Then, section 4 outlines the experimental setup and results of ship detection, including the application of geo-fencing. Finally, the conclusions are summarized in section 5.

2 RELATED WORKS

Object detection in the marine environment has been classified by many experts as a difficult problem to overcome (Moosbauer et al., 2019; Prasad et al.,

²<http://www.palmasport.es/en/>

³<https://www.plocan.eu/>

⁴<https://www.puertogijon.es/en/>

2020). Factors such as weather conditions, waves or the magnitude of distances, appear as typical obstacles in this type of domain and interfere with the predictions made by the detectors. Furthermore, the absence of labeled datasets complicates the use of Neural Networks as an element to apply inference, which requires an additional effort to improve the quality of the results.

Depending on the perspective of the images used, this topic is classified into two broad categories: boat detection from an aerial or satellite view and ship detection from a height near the surface. As an example of some of the work done for the first of these groups, the work of (Chen et al., 2009) proposed the use of polarization cross-entropy in a variant of the CFAR (Constant False Alarm Rate) algorithm, while other researchers have experimented with different approaches using Deep Learning as a common factor (Kang et al., 2017; Yang et al., 2018). The latter aspect has also been one of the most popular methods to perform detection in scenes taken from a ground viewpoint, although other scientists have presented alternative techniques, such as the work (Frost and Tapamo, 2013) in which was proposed the use of predefined silhouettes with priority levels throughout the image.

In addition to these strategies, the maritime field has an element that provides highly useful data: AIS devices (Harre, 2000). This technology receives information about the geographical position of vessels, their course and speed, their dimensions, etc. However, this element presents several limitations, such as the periodicity of the reports issued by the vehicle itself, which can generate a disparity between the last situation reflected by the AIS and reality (Tu et al., 2018). Furthermore, the IMO (International Maritime Organization) indicates that the use of AIS is not mandatory for all vessels, introducing incongruities between the real position and the data provided by the AIS⁵.

For this reason, previous studies have already used Neural Networks (Wang et al., 2019; Chen et al., 2020) and AIS instruments together. Nonetheless, most of its use has been aimed at predicting the trajectory of ships using a succession of geographic locations emitted by these transmitters (Xiaorui and Changchuan, 2011; Last et al., 2014). Consequently, to try to contribute to this field, in this work we present a prototype that unifies Neural Networks Computer Vision Techniques, information from AIS devices, and geofencing aiming to improve maritime security and safety.

⁵AIS transponders regulation

3 METHODOLOGY

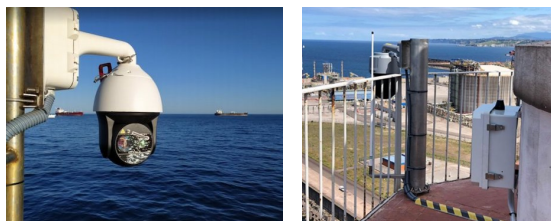
Outdoor object detection is affected by environmental conditions, as in the case of maritime areas. Thus, finding reliable tools to overcome a real-world scenario is challenging. In this sense, Neural Networks become useful to face a real context, although feature extraction is affected by, for instance, low-light conditions (i.e. night scenes). Due to this fact, in this work, we propose two different approaches in order to detect vessels, based on such lighting conditions, to then compute the correspondence between such vessels and AIS information.

First, to cope with images during the day, a YOLOv4 Neural Network architecture (Bochkovskiy et al., 2020) has been used. Among its advantages, this network has a fast loading and inference speed, low computational cost, and high quality of predictions, as presented in (Ballines Barrera, 2020b). On the other hand, detection under nighttime conditions relies on analyzing the distribution of the standard deviation inside an image, subdivided into a grid. Both approaches fit in our remote system.

In this section, we include the description of the hardware used, the dataset, and the details of both techniques, as well as the geofencing approach to introduce context information.

3.1 Hardware

Regarding the camera system used to capture maritime scenes, we use Hikvision PTZ cameras models DS-2DF8836I5X⁶ and DS-2DF9C435IHS-DLWT2⁷ installed in the Ports of Las Palmas de Gran Canaria and Gijón, respectively. These models are able to obtain images with a resolution of 4K and has a zoom of up to 36x so that distant objects can be recorded with high quality. We observe the cameras and their surrounding environment in the pictures 1(a) and 1(b).



(a) Port of Las Palmas de G.C. (b) Port of Gijón

Figure 1: Cameras installed in the context of this work.

⁶DS-2DF8836I5X specifications

⁷DS-2DF9C435IHS-DLWT2 specifications

Related to the processing hardware, we use an NVIDIA RTX 3090 to train the model. Our edge device is an NVIDIA Xavier NX 8GB with JetPack 4.5 installed (find more details at this link).

The vessels' detection also uses information from a dual-channel AIS receiver. This system uses an NMEA 0183 VDM protocol and includes USB 2.0 outputs delivering the data transmitted by AIS Class A and Class B transponder devices, AIS SARTs, and Aids to navigation.

3.2 Dataset

With the idea of improving the network's ability to detect ships, we have created a new dataset for training and validation. To this aim, the PTZ camera described above was used to obtain captures from multiple views and with different zoom levels. Nevertheless, due to the fact that many of the images did not contain vessels, a previous object detection was applied with a twofold purpose. On the one hand, this process allows for differentiating images with and without vessels. On the other hand, preliminary labeling is obtained. Afterward, this last set was revised to filter the false negatives that would happen during such first detection.

The obtained dataset consists of 3,000 samples, which were labeled in YOLO format. Moreover, to increase its quality, a data augmentation process was performed by using image transformation effects such as color or geometry alterations, or noise addition (among others). In this way, the model can be strengthened by including scenarios with fog, rain, and other adverse weather conditions. Finally, parts of the SeaShips (Shao et al., 2018) and Microsoft COCO (Lin et al., 2014) datasets were used to incorporate a wider diversity of perspectives and vessels, increasing the total amount of images up to 10,000. More details can be seen in (Ballines Barrera, 2020a).

3.3 Vessels Detection

As introduced above, vessel detection is a tough task, due to complex environment and lighting conditions. In this way, we propose two different approaches depending on the moment of the day, i.e. daytime and nighttime.

3.3.1 Daytime

Based on pre-trained weights from YOLOv4 network under COCO dataset, we perform transfer learning using the dataset described in section 3.2. For this purpose, this set was divided into training (80%) and

testing (20%). To evaluate the network performance, we consider the *mean Averaged Precision* (mAP). Initially, the model provided a score of 55.7% mAP, but after transfer learning, it reached a 85.3% mAP (20,000 epochs and using the standard network configuration). We rely on this last model to detect the vessels in the scene during the daytime.

3.3.2 Nighttime

Instead of using the inferences provided by the re-trained YOLOv4 network, we propose an alternative method to obtain the location of boats in nighttime images. This technique is based on the assumption that, on the presence of light sources on a maritime image during the night, such sources correspond to vessels. However, night scenes suffer from noise and artifacts, which should be minimized in order to provide better detection. In this sense, an automatic system was developed to adjust the camera's image parameters, aiming to reduce flashing lights at nightfall and reduce noise. Moreover, at dawn, it restores the daytime settings of the camera to analyze images using the method proposed in the above section.

The proposed technique works in the following way: firstly, the input image is converted to grayscale (Figure 2(a)). The points closer to a light source have higher levels of intensity, so thresholding is applied in order to obtain the whiter points. We have experimentally set a threshold value of 80. In practice, this value is enough to filter artifacts while preserving light sources. In Figure 2, we include the results for different threshold values (Figures 2(b)–2(d)). As observed, the higher the threshold, the lower the lights preserved, and vice versa. Afterward, using such a threshold, the image is converted to black and white.

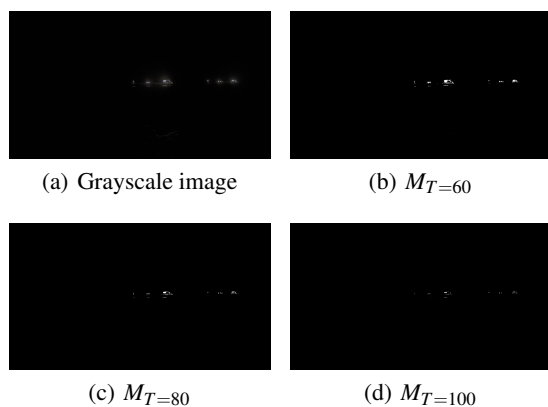


Figure 2: Effect of the threshold value on the BW mask: (a) input image converted to grayscale and the results for different thresholds to obtain the corresponding mask (M_T) (b) 60, (c) 80, and (d) 100).

Once the input image is filtered, it is divided into a 20×20 grid. The idea is to compute the standard deviation (σ) in each resulting cell, in order to detect those regions with white pixels. We have experimentally observed that using $\sigma > 20$ is enough to detect regions belonging to ship lights, avoiding false positives, and providing enough cells to build up a bounding box that contains the ship. Finally, from the selected cells, a new bounding box is generated, by annexing the contiguous regions. In this way, an approximation of detection boxes similar to the ones provided by the YOLOv4 network is obtained.

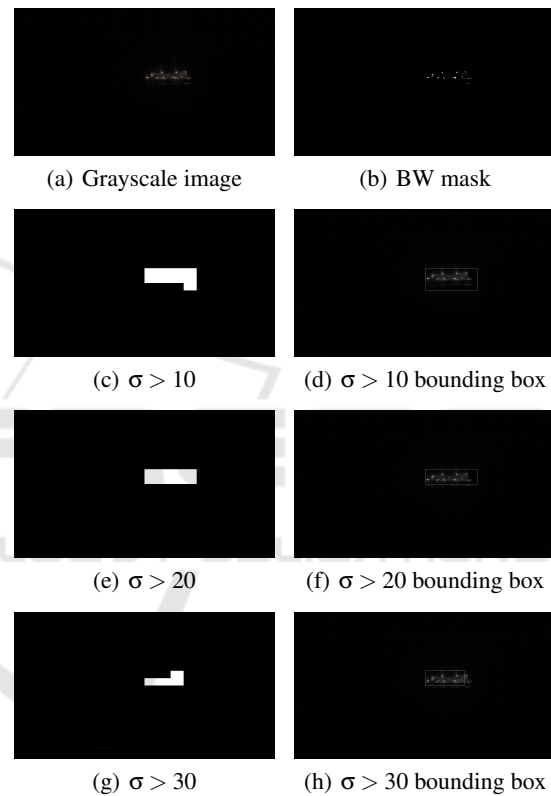


Figure 3: Effect of the threshold for σ on the bounding box result: (a) original image, (b) associated mask, and the results for different sigma thresholds and their associated bounding box ($\sigma > 10$ – (c)(d), $\sigma > 20$ – (e)(f), and $\sigma > 30$ – (g)(h)).

In order to illustrate the influence of different thresholds for σ we include Figure 3, which contains the original image in grayscale (Figure 3(a)), its associated mask (Figure 3(b)), and the effect of various threshold values. As observed, for lower values, the obtained region is bigger, and its associated bounding box includes water at the bottom. On the other hand, for greater threshold values, the resulting bounding box leaves out part of the ship stern. Hence, it seems suitable to use an intermediate value for the threshold.

3.4 Estimation of the Georeferenced Position

Independently of the approach used to detect the vessels in the scene, we can estimate their georeferenced position. This can be done by using the Python library *CameraTransform* (Gerum et al., 2019), which includes utilities to perform a projection of points from an image (2D) to a real-world scene (3D), given the intrinsic and extrinsic camera parameters. As we are not using panoramic sequences, a rectilinear projection (pin-hole camera) is used to estimate such a projection.

This way, it is possible to obtain the estimated latitude and longitude of any pixel in the image. In particular, we are interested in obtaining such information on the points belonging to bounding boxes from YOLOv4 inferences. Owing to the fact that bounding boxes are made up of more than a single pixel, if we compute the geolocation of a ship based on such pixels, we will obtain multiple distances. To avoid this, we use the center point of the lower side of the detection bounding box as a reference to estimate the location of each ship.

On the other hand, it is possible to know the precise location of near ships by using the information transmitted by AIS devices. As the source of such information is hosted in an external server, it is obtained with a one-minute frequency. The received file consist of pairs of NMEA-ZDA⁸ and AIVDM⁹ / AIVDO¹⁰ messages, thus managing to assign an instant of time to all transferred sequences.

To speed up the data file reading, we record the last processed position and use it as a reference for the next access. Furthermore, to obtain the most relevant data, we apply two types of filters to the sequence, based on the date of transmission and the position of the ship. For the latter, the Equations 1 and 2 were used to obtain a specific region of the environment, taking into account the direction and field of view of the camera. Finally, only the most recent message from each vessel is kept for the next steps in the process.

$$\varphi_2 = \arcsin\left(\sin(\varphi_1) \cos\left(\frac{d}{r}\right) + \cos(\varphi_1) \sin\left(\frac{d}{r}\right) \cos\theta\right), \quad (1)$$

$$\lambda_2 = \lambda_1 + \arctan\left(\frac{\sin(\theta) \sin\left(\frac{d}{r}\right) \cos(\varphi_1)}{\cos\left(\frac{d}{r}\right) - \sin(\varphi_1) \sin(\varphi_2)}\right), \quad (2)$$

where d denotes the haversine distance between both locations, r the sphere radius, φ and λ the latitude and

longitude of each perspective point, and θ the bearing (clockwise from North).

Finally, to determine the disparity between the AIS information and the estimated position, both are correlated. First, both data are ordered according to their latitude and longitude, and then Equation 3 is used to compute the haversine distance between all possible combinations.

$$d = 2r \arcsin\left(\sqrt{\sin^2\left(\frac{\varphi_1 - \varphi_2}{2}\right) + \cos(\varphi_1) \cos(\varphi_2) \sin^2\left(\frac{\lambda_1 - \lambda_2}{2}\right)}\right), \quad (3)$$

where r indicates the sphere radius, and φ and λ the latitude and longitude of each perspective point, respectively.

Thus, the real distance between the method estimation and the AIS information is computed, binding those pairs with the lowest distance. As a consequence, we obtain a correspondence between the scene captured by the camera and the information provided by AIS devices, by fusing both data.

3.5 Geofencing

Apart from estimating the georeferenced position of the vessels in the scene, it is also interesting to obtain information regarding context. Techniques like geofencing make it possible, which is very useful in decision-making tasks because it provides a relationship between the ship's position and its surrounding region.

With geofencing, we project synthetic lines, regions, or polygons delineated initially on another reference system. In particular, in our case, we are interested in performing a back-and-forth projection. On the one hand, the map includes regions we want to project on the captured scene. On the other hand, the vessels' detection provides as output the bounding boxes, which will be included in the map to represent the location. As in section 3.4, we use the library *CameraTransform*, that includes functionalities to obtain pixel coordinates from GPS data and vice versa.

4 EXPERIMENTAL RESULTS

In order to test the performance of the proposed techniques, we present different experiments. First, we include results regarding the estimation of vessels' geolocation at daytime and nighttime. Afterward, an analysis including geofencing is presented.

⁸Time and date

⁹Received data from other vessels

¹⁰Own vessel's information

4.1 Georeferenced Position Estimation

To evaluate the precision of the georeferenced position estimation, we measure the distance error by using equation 4:

$$E_d = \frac{|M_d - AIS_d|}{C_d} \cdot 100, \quad (4)$$

where E_d denotes distance error and M_d , AIS_d , and C_d are the distances given by our method, the AIS device, and the distance to the camera, respectively. In this way, we show the percentage error between the position provided by AIS and the distance obtained by our proposal. For this aim, we have averaged the results related to measures carried out within 2 days, by computing them at different zoom levels and considering diverse target distances.

4.1.1 Daytime

As described in section 3.3.1, during daytime we use the inferences given by the re-trained YOLOv4 network to estimate the vessels' location. In order to quantitatively characterize the results, in Table 1 we include figures regarding percentage error computed using equation 4, depending on the distance of the target boats and the zoom level used.

Table 1: Daytime geolocation error based on target ship distance and camera zoom level.

		Distance (m)			
		0–2,500	2,501–5,000	5,001–7,500	7,501–10,000
Zoom	0.1	3.38%	7.61%	12.51%	19.51%
	0.2	7.09%	6.18%	7.82%	8.48%
	0.3	9.17%	3.13%	8.39%	9.92%
	0.4	12.16%	4.93%	1.26%	6.11%
	0.5	12.73%	5.69%	2.90%	4.53%

As observed, in general, the zoom level is crucial to increase method precision. Nevertheless, this is not true for the closest distance, probably due to the lens distortion in near objects when using a long focal distance. Moreover, it is necessary to consider that several challenges are faced when we measure in a real scenario. Firstly, the low height camera position relative to sea level could lead to wrong distance estimation of distant objects due to the perspective. Secondly, wind introduces instability in the view plane. Finally, we rely on the precision of the ships' detection given by the network. Therefore, an imprecise detection results in a wrong estimation of the location.

Qualitative results are given in Figure 4. Images from 4(a) and 4(c) represent the estimation of the position given by our method, whereas 4(b) and 4(d) include the real location given by AIS devices. In the

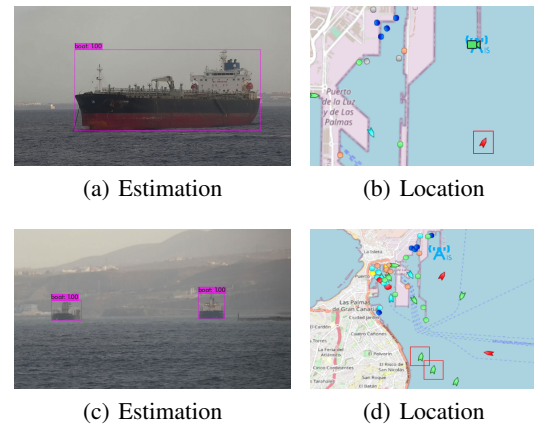


Figure 4: Daytime detection results: at a close distance ((a), (b)) and at a medium distance ((c), (d)). We compare the estimation ((a),(c)) with the real location ((b), (d) – red squares mark vessels' positions on the map).

case of a close distance (Figures 4(a) and 4(b)), the result given by our approach is 1,101 m whilst the AIS system provides a distance of 1,233 m. On the other hand, when a medium distance is considered (Figures 4(c) and 4(d)), the re-trained YOLOv4 model result on 5,639 m (left) and 5,301 m (right), and the AIS devices indicate 5,539 m and 5,180 m, respectively. Considering these results, we can estimate a mean error of about 118 m.

4.1.2 Nighttime

Aiming to estimate the position of vessels during nighttime, we apply the approach described in section 3.3.2. Table 2 includes numbers regarding the results of the detection, considering different zoom levels with several target distances, as in the case of daytime sequences.

Table 2: Nighttime geolocation error based on target ship distance and camera zoom level.

		Distance (m)		
		0–2,500	2,501–5,000	5,001–7,500
Zoom	0.1	26.96%	30.24%	—
	0.2	23.48%	21.15%	—
	0.3	23.25%	8.40%	—
	0.4	25.69%	6.89%	15.70%
	0.5	24.97%	4.26%	18.63%

As can be observed, nighttime detection is less precise than the daytime one. In this sense, it is necessary to consider that night has several limitations, mainly due to visibility constraints. Unlike the previous method, we have observed that the nighttime scheme is not able to estimate ships between 7,501 m

and 10,000 m, nor ships between 5,001 m and 7,500 m with a low zoom level. Such limitations are mainly due to the thresholds (black-and-white mask and σ) and to the grid configuration. Related to the first one, they can be modified but lead to detect false positives. Regarding the lattice, shorter zooms provide a more general view, which makes it tough to distinguish between ships at different distances.

Besides, since the lights of the ships are located at a certain height rather than at the waterline, this influences the quality of the bounding box obtained. This is observed mainly on nearby vessels, where the resulting bounding box is placed higher than expected. As a consequence, the estimated distance is farther than the real one.

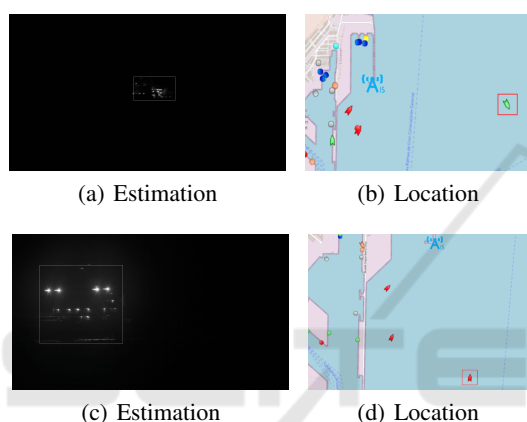


Figure 5: Nighttime detection results: at a close distance ((a), (b)) and at a medium distance ((c), (d)). We compare the estimation ((a), (c)) with the real location ((b), (d)) – red squares mark vessels' positions on the map).

Figure 5 depicts the results obtained during nighttime, comparing our method estimation (Figures 5(a) and 5(c)) with the position given by the AIS (Figures 5(b) and 5(d)). For the case of a close distance (Figures 5(a) and 5(b)), the proposed technique gives a distance of 1,483 m, whereas the AIS system provides 1,633 m. On the other hand, when a medium distance is considered, the nighttime approach results in 3,292 m, whilst AIS device supplies a 3,058 m distance. This yields to an averaged error of about 192 m.

4.2 Geofencing

Once we have presented different results related to vessels detection and location estimation, in this set of experiments we include an analysis regarding geofencing.

Figure 6 includes the geofencing applied on an image belonging to the Port of Las Palmas de Gran Canaria. On the map (6(a)), the vessels are represented

with points, which are given by GPS coordinates obtained from the projection resulting detection (represented with orange points and labeled with numbers from 1 to 5) in 6(b). Alike, in the scene can be seen the outcome of projecting the lines from the map to the scene. Thus, we obtain mutual information, regarding the vessels and their position inside a defined region.

A more complex example is presented in Figure 7, which includes a geofencing in the Port of Gijón. As represented in the map of Figure 7(a), an anchoring area has been defined. Again, in this map, the detected vessels are represented by orange points and labeled with numbers. The corresponding geofenced region in the scene can be observed in Figure 7(b), where the ship labeled with 1 is almost on the limit of the anchoring area, which also corresponds to the scene viewed in the map.

In order to illustrate the robustness of the re-trained YOLOv4 network, we include Figure 8. As observed in Figure 8(a), we have a scene with multiple vessels. In particular, it can be seen that a small boat is in the forefront, whose detail is illustrated on the right. One of the strengths of the presented technique is that is able to detect small boats, even those which are not required to use AIS devices. Similarly to previous figures, we also include the result of the geofencing in the case of this small vessel. Although in the scene in Figure 8(c) seems that the small boat and the crane ship behind are close, this is a matter of perspective. We can see in Figure 8(b), that the boat labeled with 2 is inside Navigation Channel I, whereas the crane ship marked with 4 is located in between Navigation Channels I and II. This can be also noticed in the geofencing depicted in Figure 8(c), where the green line is behind the boat and the orange one is on the starboard of the crane ship.

5 CONCLUSIONS

A solution that fuses data from different sources for decision-making tasks on coastal regions has been presented in this work. To this end, two approaches to detect vessels are used, depending on the moment of the day. From that outcome, the ships' distances are estimated and then paired with the information provided by AIS devices. Together with a geofencing technique, mutual information is correlated. This way, this application gives the vessels' GPS position on the map and projects virtual fences on the camera scene, and includes boats with no AIS systems.

When comparing the used technique to estimate vessels' geographic location with the AIS informa-

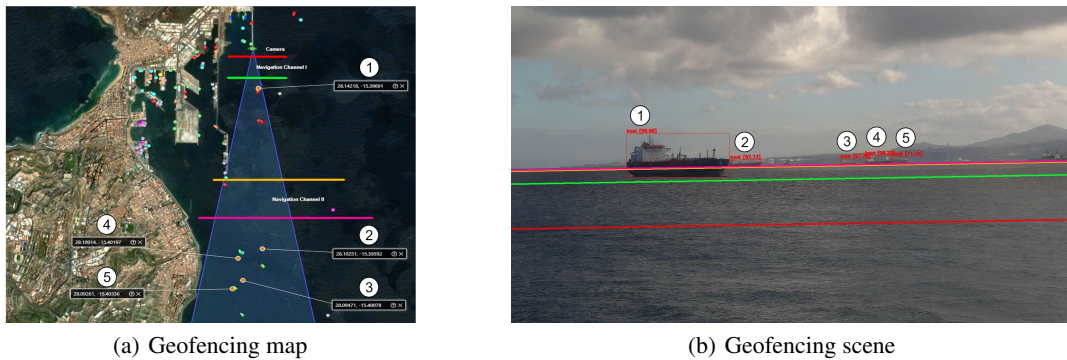


Figure 6: Geofencing in the Port of Las Palmas de Gran Canaria: (a) geofencing on the map and (b) geofencing on the scene.

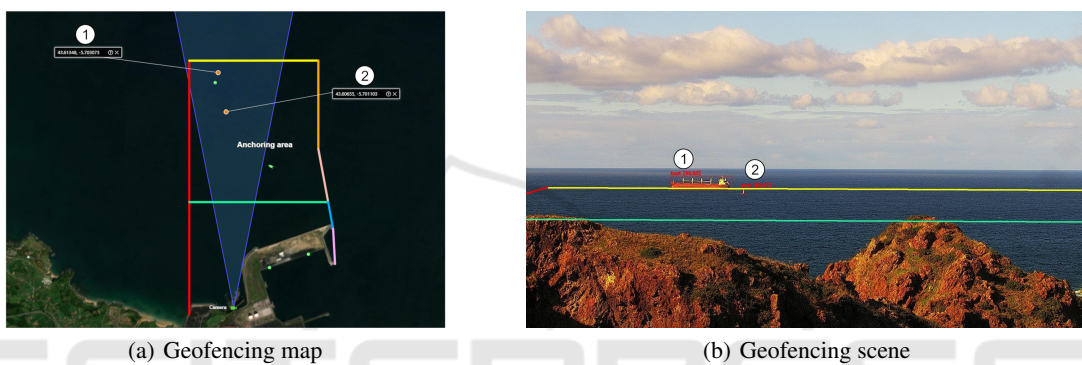


Figure 7: Geofencing in the Port of Gijón: (a) geofencing on the map and (b) geofencing on the scene.

tion during the daytime, we have observed that error ranges between 5 % and 10 %. In general, such an error is reduced when the zoom is applied, except for close targets. Elements like camera altitude, wind exposure, and the precision of the detection given by the network affect the results. As a consequence, the view and stabilization of the scene for further computations become a tough task.

Regarding nighttime results, nightly conditions strongly influence the estimation of far ships, providing imprecise estimations than the daytime approach. The use of different thresholds and a fixed grid configuration, affect the outcome as well. Thus, ships in different positions are assigned to very similar latitudes, and considering the magnitude of distances in the maritime environment, this fact induces errors in the remote evaluations.

Future work relies on hardware and software improvements. The daytime technique would be benefited by improving camera location, in order to provide a better perspective, paving the way for better distance estimation. Furthermore, the addition of complementary cameras could improve the detection of ships, mainly when they are hidden from the current single-camera perspective. This contributes to

a wider view, improving region coverage to a possible security system. To deal with discrepancies regarding moving vessels, using the heading and speed data from AIS devices to compute their new position would help to improve the results. The nighttime approach could be improved in two ways. Firstly, using a dynamic size and position for the grid in the regions of interest, including enhancing the lower margin of the bounding box, to contain the boat's waterline. Secondly, the use of a night camera is another possibility, although its cost is a disadvantage.

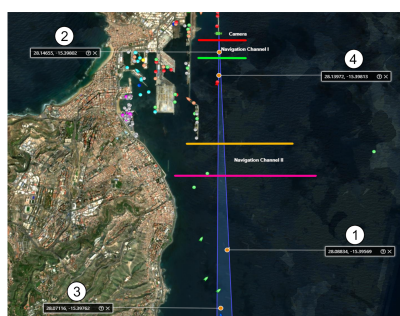
ACKNOWLEDGEMENTS

This research is the result of a collaboration between QAISC and researchers from CTIM, within the framework of the research contract C2021/193 signed between the company and The Canarian Science and Technology Park Foundation of the University of Las Palmas de Gran Canaria. We specially thank to Prof. Agustín Trujillo.

It also has been partially supported by Vicepresidencia Primera, Consejería de Vicepresidencia Primera y de Obras Públicas, Infraestructuras,



(a) Ships' detection and detail



(b) Geofencing map



(c) Geofencing scene

Figure 8: Detection and geofencing of a small boat in the Port of Las Palmas: (a) original scene and detail of a small boat without AIS system, (b) geofencing on the map, and (c) geofencing on the scene.

Transporte y Movilidad from Cabildo de Gran Canaria, through the project of reference Resolution No. 45/2021 and by The Canarian Science and Technology Park Foundation of the University of Las Palmas de Gran Canaria through the research project F2021/05 FEI Innovación y Transferencia empresarial.

We also thank the Oceanic Platform of the Canary Islands (PLOCAN) and The Port Authority of Las Palmas for their authorization to use their coastal observation facility and their support within the framework of the "Smart Coast AI SOLUTIONS BLUE ECONOMY LAB" initiative led by QAISC in collaboration with CTIM.

REFERENCES

- Arshad, N., Moon, K.-S., and Kim, J.-N. (2010). Multiple Ship Detection and Tracking Using Background Registration and Morphological Operations. In Kim, T.-h., Pal, S. K., Grosky, W. I., Pissinou, N., Shih, T. K., and Slkezak, D., editors, *Signal Processing and Multimedia*, pages 121–126, Berlin, Heidelberg. Springer Berlin Heidelberg.
- Ballines Barrera, S. (2020a). Detección de objetos en el entorno marítimo mediante dispositivos AIS y cámaras PTZ. Master's thesis, University of Las Palmas de Gran Canaria.
- Ballines Barrera, S. (2020b). Identificación y detección de objetos móviles mediante redes neuronales empleando el sistema NVIDIA Jetson TX2. Master's thesis, University of Las Palmas de Gran Canaria.
- Bloisi, D., Iocchi, L., Fiorini, M., and Graziano, G. (2012). Camera based target recognition for maritime awareness. In *2012 15th International Conference on Information Fusion*, pages 1982–1987. IEEE.
- Bochkovskiy, A., Wang, C., and Mark Liao, H. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. *arXiv preprint arXiv:2004.10934*, abs/2004.10934.
- Chen, J., Chen, Y., and Yang, J. (2009). Ship Detection Using Polarization Cross-Entropy. *IEEE Geoscience and Remote Sensing Letters*, 6(4):723–727.
- Chen, Z., Chen, D., Zhang, Y., Cheng, X., Zhang, M., and Wu, C. (2020). Deep learning for autonomous ship-oriented small ship detection. *Safety Science*, 130:104812.
- Frost, D. and Tapamo, J.-R. (2013). Detection and tracking of moving objects in a maritime environment using level set with shape priors. *EURASIP Journal on Image and Video Processing*, 2013(1):1–16.
- Fu, H., Song, G., and Wang, Y. (2021a). Improved yolov4 marine target detection combined with cbam. *Symmetry*, 13(4).
- Fu, H., Zhang, R., Ning, X., and Wang, Y. (2021b). Ship detection based on improved yolo algorithm. In *2021*

- 40th Chinese Control Conference (CCC), pages 8181–8186.
- Gerum, R. C., Richter, S., Winterl, A., Mark, C., Fabry, B., Le Bohec, C., and Zitterbart, D. P. (2019). Camera-Transform: A Python package for perspective corrections and image mapping. *SoftwareX*, 10:1–6.
- Harre, I. (2000). Ais adding new quality to vts systems. *Journal of Navigation*, 53(3):527–539.
- Hu, W.-C., Yang, C.-Y., and Huang, D.-Y. (2011). Robust real-time ship detection and tracking for visual surveillance of cage aquaculture. *Journal of Visual Communication and Image Representation*, 22(6):543–556.
- Kaluza, P., Kölzsch, A., Gastner, M. T., and Blasius, B. (2010). The complex network of global cargo ship movements. *Journal of the Royal Society Interface*, 7(48):1093–1103.
- Kang, M., Ji, K., Leng, X., and Lin, Z. (2017). Contextual Region-Based Convolutional Neural Network with Multilayer Fusion for SAR Ship Detection. *Remote Sensing*, 9(8):1–14.
- Larson, K. M., Shand, L., Staid, A., Gray, S., Roesler, E. L., and Lyons, D. (2022). An optical flow approach to tracking ship track behavior using goes-r satellite imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15:6272–6282.
- Last, P., Bahlke, C., Hering-Bertram, M., and Linsen, L. (2014). Comprehensive Analysis of Automatic Identification System (AIS) Data in Regard to Vessel Movement Prediction. *Journal of Navigation*, 67(5):791–809.
- Li, H. and Man, Y. (2016). Moving ship detection based on visual saliency for video satellite. In *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 1248–1250.
- Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., and Zitnick, C. L. (2014). Microsoft COCO: Common Objects in Context. In *European Conference on Computer Vision – ECCV 2014*, pages 740–755. Springer, Springer International Publishing.
- Moosbauer, S., König, D., Jäkel, J., and Teutsch, M. (2019). A Benchmark for Deep Learning Based Object Detection in Maritime Environments. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 916–925.
- Prasad, D. K., Dong, H., Rajan, D., and Quek, C. (2020). Are object detection assessment criteria ready for maritime computer vision? *IEEE Transactions on Intelligent Transportation Systems*, 21(12):5295–5304.
- Satyanarayanan, M. (2017). The Emergence of Edge Computing. *Computer*, 50(1):30–39.
- Shao, Z., Wu, W., Wang, Z., Du, W., and Li, C. (2018). SeaShips: A Large-Scale Precisely Annotated Dataset for Ship Detection. *IEEE Transactions on Multimedia*, 20(10):2593–2604.
- Simonsen, C., Theisson, F., Holtskog, Ø., and Gade, R. (2020). Detecting and locating boats using a ptz camera with both optical and thermal sensors. In Farinella, G., Radeva, P., and Braz, J., editors, *Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, volume 5, pages 395–403. SCITEPRESS Digital Library. 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, VISIGRAPP 2020 ; Conference date: 27-02-2020 Through 29-02-2020.
- Sirimanne, S. N., Hoffmann, J., Asariotis, R., Ayala, G., Assaf, M., Bacrot, C., Benamara, H., Chantrel, D., Cournoyer, A., Fugazza, M., Hansen, P., Kulaga, T., Prenti, A., Rodríguez, L., Salo, B., Tahiri, K., Tokuda, H., Ugaz, P., and Youssef, F. (2021). *Review of Maritime Transport 2021*. United Nations Conference on Trade and Development (UNCTAD), New York: United Nations.
- Tu, E., Zhang, G., Rachmawati, L., Rajabally, E., and Huang, G.-B. (2018). Exploiting AIS Data for Intelligent Maritime Navigation: A Comprehensive Survey From Data to Methodology. *IEEE Transactions on Intelligent Transportation Systems*, 19(5):1559–1582.
- Wang, N., Wang, Y., and Er, M. J. (2022). Review on deep learning techniques for marine object recognition: Architectures and algorithms. *Control Engineering Practice*, 118:104458.
- Wang, Y., Wang, C., Zhang, H., Dong, Y., and Wei, S. (2019). Automatic ship detection based on retinanet using multi-resolution gaofen-3 imagery. *Remote Sensing*, 11(5).
- Xiaorui, H. and Changchuan, L. (2011). A Preliminary Study on Targets Association Algorithm of Radar and AIS Using BP Neural Network. *Procedia Engineering*, 15:1441–1445.
- Yang, X., Sun, H., Fu, K., Yang, J., Sun, X., Yan, M., and Guo, Z. (2018). Automatic Ship Detection in Remote Sensing Images from Google Earth of Complex Scenes Based on Multiscale Rotation Dense Feature Pyramid Networks. *Remote Sensing*, 10(1):1–14.
- Zhang, R., Li, S., Ji, G., Zhao, X., Li, J., and Pan, M. (2021). Survey on deep learning-based marine object detection. *Journal of Advanced Transportation*.
- Zhao, H., Zhang, W., Sun, H., and Xue, B. (2019). Embedded deep learning for ship detection and recognition. *Future Internet*, 11(2).