

RESEARCH ARTICLE

BoatNet: automated small boat composition detection using deep learning on satellite imagery

Guo Jialeng^{1*}, Santiago Suárez de la Fuente¹ and Tristan Smith¹

How to cite

Jialeng G, Fuente S, Smith T. BoatNet: automated small boat composition detection using deep learning on satellite imagery. *UCL Open: Environment*. 2023;(5):05. Available from: <https://doi.org/10.14324/111.444/ucloe.000058>

Submission date: 20 July 2022; Acceptance date: 3 April 2023; Publication date: 24 May 2023

Peer review

UCL Open: Environment is an open scholarship publication, this article has been peer-reviewed through the journal's standard open peer review process. All previous versions of this article and open peer review reports can be found online in the *UCL Open: Environment* Preprint server at ucl.scienceopen.com

Copyright and open access

©2023 The Authors. Creative Commons Attribution Licence (CC BY) 4.0 International licence <https://creativecommons.org/licenses/by/4.0/>

Open access

This is an open access article distributed under the terms of the Creative Commons Attribution Licence (CC BY) 4.0 <https://creativecommons.org/licenses/by/4.0/>, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited.



*Corresponding author

E-mail: jialeng.guo@ucl.ac.uk

¹UCL Energy Institute, The Bartlett School of Environment, Energy and Resources, University College London, London, UK

Abstract

Tracking and measuring national carbon footprints is key to achieving the ambitious goals set by the Paris Agreement on carbon emissions. According to statistics, more than 10% of global transportation carbon emissions result from shipping. However, accurate tracking of the emissions of the small boat segment is not well established. Past research looked into the role played by small boat fleets in terms of greenhouse gases, but this has relied either on high-level technological and operational assumptions or the installation of global navigation satellite system sensors to understand how this vessel class behaves. This research is undertaken mainly in relation to fishing and recreational boats. With the advent of open-access satellite imagery and its ever-increasing resolution, it can support innovative methodologies that could eventually lead to the quantification of greenhouse gas emissions. Our work used deep learning algorithms to detect small boats in three cities in the Gulf of California in Mexico. The work produced a methodology named BoatNet that can detect, measure and classify small boats with leisure boats and fishing boats even under low-resolution and blurry satellite images, achieving an accuracy of 93.9% with a precision of 74.0%. Future work should focus on attributing a boat activity to fuel consumption and operational profile to estimate small boat greenhouse gas emissions in any given region.

Keywords: object detection, deep learning, machine learning, transfer learning, small boat activity, climate change

Introduction

Energy crisis, energy security and climate change

The Intergovernmental Panel on Climate Change (IPCC) explains, in its latest report, that humans and nature are being pushed beyond their abilities to adapt due to the anthropogenic emissions caused by economic, industrial and societal activities [1]. Carbon-intensive resources still comprise a large proportion of the energy system [1] – about 80% in 2017 [2]. However, the share of electricity production from renewables increased from 20.8% to 29.0% between 1985 and 2020 [3]. Still, carbon emissions have not been reduced in line with the ambitions of the Paris Agreement, and it is predicted that in the next few years, the gains in carbon reduction due to the Covid-19 pandemic will be erased, faster than expected [4]. However, even under all these pressures and projections, it is still possible for humanity to keep the global temperature below 1.5°C from pre-industrial levels by 2100 if substantial changes are made to the current energy systems.

However, energy security is an important part of the strategies proposed by countries to support economic growth and provide essential services to their populations. Currently, nations deposit most of their energy security into fossil fuels while expanding their renewable power capacity. Fossil fuels and their conversion systems (e.g., internal combustion engines) permit operators to react quickly to changes in the energy demand (i.e., more control over energy deployment) while offering acceptable volumetric energy densities. However, heavy reliance on fossil fuels, coupled with the fuel's geographical origin, is at the mercy of important price fluctuations due to geopolitical and logistical events, such as Russia's invasion of Ukraine. These can disrupt global energy systems and affect the stability of nations and human livelihoods [5,6]. On the other hand, renewable energy production and distribution tend to lie within a country's boundaries. Over the last few years, the price of renewable energy has been catching up with those of subsidised fossil fuels – with some specific examples already undercutting fossil fuel prices [7]. In fact, from 1987 to 2015, the cost of oil and coal rose by approximately 36% and 81%, respectively, and from 1989 to 2015, the cost of natural gas rose by approximately 53% [8]. More recently in March 2022, the UK experienced increases in natural gas to around £5.40/therm, a rise above 1,100% from the price levels seen in 2021 [9]. Nevertheless, it is important to note that renewable energy variability and investment requirements are significant challenges to grid stability and energy security.

Shipping sector, small boat fleet and emission inventory

Shipping, the backbone of market globalisation, plays an important role in the carbon reduction of human activities as it moves around 90% of all goods around the globe [10]. However, its reliance on fossil fuels, coupled with robust economic growth, saw total carbon dioxide (CO₂) emissions grow from 962 megatonne (Mt) in 2012 to 1056 Mt in 2018, representing more than 10% of the total global transportation emissions [11]. Furthermore, if nothing is done in the sector, it is projected that by 2050 shipping CO₂ emissions could grow to 1500 Mt. In this light, the International Maritime Organisation (IMO) produced its ambitions to decarbonise international shipping [12] in 2018. However, this vision only covers international navigation composed of large vessels and does not consider the small boat fleet – vessels below 100 gross tonnages that tend to measure less than 24 m in length [13].

There are good reasons for this decision. First, the IMO focuses mainly on ships that navigate international waters or large ships performing domestic voyages [14]. These vessels are required to have the automatic identification system (AIS) transponders for safe navigation. On the other hand, small boats tend not to have an AIS or a global positioning system (GPS) transponder [15,16], which makes the study of their movements more challenging. Second, small boats are typically registered and monitored by national and regional bodies, and the comprehensiveness of data depends on capital and human resources in addition to the infrastructure to maintain the registry [17]. Third, small boats are a diverse segment of shipping and usage depends on the geographical location, type of activity, construction and operating costs and accessibility to fuel or bunkering infrastructure [18]. Similarly, engine providers are extensive, giving a broad range of fuel consumption curves and emissions [19–21].

Furthermore, fuel selection is equally diverse: petrol, diesel, petrol mixed with engine oil – mainly for two-stroke engines, ethanol and bio-fuels – or a mix of bio-fuel with different fossil fuels. Finally, not

all small boats are powered by an internal combustion engine. They can instead be powered by sail, battery-electric or paddles [22–24].

Nevertheless, with all these challenges, the small boat fleet can significantly contribute to the shipping segment's emission footprint based on its activity [25,26]. Emissions inventories aid our understanding of what measures must be taken to enable governments and industry to start the road to full decarbonisation in a just and equitable way [27–29]. Furthermore, creating effective policies and regulations based on accurate emissions accounting can incentivise the use of energy-efficient technologies, electrification and scalable zero-emission fuels [30,31]. Additionally, if countries want to meet their ambitious decarbonisation emissions targets, they cannot afford to ignore the role played in greenhouse gas (GHG) emissions by the small boat fleet [32–35].

Although it is possible to estimate emissions from large vessels using AIS data sent from a ship's transponder to be coupled with technical models [11], small vessels depend on the national registration system. Their operation is typically assumed or captured by national fuel sales, which tend to be highly aggregated (e.g., [36]). Developed economies, such as the UK, tend to have a national registry of smaller vessels [37] that provides a sense of their activity level and hence can infer CO₂ emissions.

However, in developing countries, it tends to be a mixed bag in terms of the level of precision and availability. For instance, in Mexico, only fishing vessels are counted in the national registry [38]. Still, it is not easy to know where they are located and infer their activities. Overall, Mexico does not have a regional CO₂ inventory specialised in the small boat fleet; instead, they are aggregated as part of the *maritime and fluvial navigation [1A3d]* class in the national annual emission inventory developed by the Instituto Nacional de Ecología y Cambio Climático (INECC) [39] in a top-down approach based on the IPCC Guidelines [40]. Therefore, quantifying and categorising the small boat fleet will allow a better precision of where and how the emissions are being emitted and will enhance the maritime emission inventories.

Observing shipping activity in the Gulf of California is essential due to its unique geographical location, conformation and biophysical environment [41–43]. Furthermore, the Gulf of California, includes the largest fishing state (Sonora) in Mexico [44] and the most prominent sports fishing destination (Los Cabos, Baja) [45]. Additionally, the region is one of the most protected areas in Mexico due to its diversity of flora and fauna; the area includes the upper part of the Gulf of California, Bahía Loreto and Bahía de los Angeles [46,47].

Bringing deep learning to small ship detection in satellite imagery

Bringing deep learning, especially convolutional neural networks (CNNs), to the field of satellite image recognition is essential. Satellite image recognition is an important technology for various fields, such as environmental monitoring, natural resource management and disaster response [48–50]. It involves analysing satellite imagery to extract useful information, such as identifying objects, patterns and changes in the earth's surface. Traditional methods for satellite image recognition rely on hand-crafted features and rules, which can be time-consuming and error-prone [51–53].

Deep learning is a type of artificial intelligence (AI) that has shown great promise in solving complex problems in fields such as computer vision and natural language processing. It involves training large neural networks on vast amounts of data, which allows them to automatically learn complex patterns and relationships [54]. CNNs are a type of deep learning model that is particularly well-suited for image recognition tasks. They can learn hierarchical representations of visual data and can handle large amounts of data, making them efficient and effective for satellite image recognition [55,56].

Recent advances in satellite image recognition using deep learning have shown promising results. For example, researchers have used CNNs to detect objects or patterns in satellite imagery with high accuracy, such as roads, buildings and vegetation [57,58]. They have also applied deep learning to tasks such as land use classification, land cover mapping and disaster damage assessment [59–61].

In conclusion, bringing deep learning, especially CNNs, to the field of satellite image recognition is a large area of opportunity. It allows leveraging the power of AI to automatically learn complex patterns and relationships in satellite imagery. This can lead to improved accuracy, efficiency, automation and scalability compared to traditional methods, and has the potential to benefit a range of fields that rely on satellite imagery data.

Contributions

The contributions of this study are summarised as follows:

- A purpose-built methodology for this work, BoatNet, was developed. This work shows that BoatNet detects many small boats in low-resolution, blurry satellite images with considerable noise levels. As a result, the precision of training can be up to 93.9%, and detecting small boats in the Gulf of California can be up to 74.0%.
- This work demonstrated that BoatNet could detect the length of small boats with a precision up to 99.0%.
- BoatNet has allowed for a better understanding of the small boat activity and physical characteristics. Based on this, it has been possible to answer questions about the composition of small boats in the Gulf of California. Regarding the authors' knowledge of the literature, this is a first but essential step in constructing a way, based in object recognition, to estimate the maritime carbon footprint of the small boat fleet.

Related work

Small boat fleet and carbon emissions

Previous work related to estimating small-scale vessels without machine learning methods includes using top-down and bottom-up approaches and the use of statistical assumptions.

Parker et al. [62] used a top-down approach to estimate fishing sector emissions in 2011, which reached about 179 Mt carbon dioxide equivalent (CO₂e), representing 17.1% of the total large fishing ship emissions in that year [63]. However, their work only distinguished between motorised and non-motorised fishing vessels. Greer et al. [64] took a bottom-up approach to classify the fishing fleet in six different sizes, three below 24 m long. The findings show that the small fishing boat fleet in 2016 emitted 47 Mt CO₂, about 22.7% of the total fishing fleet. Ferrer et al. [65] used an activity-based method using GPS, landing and fuel-used data to estimate the fishing activity around the Baja California Peninsula in Mexico. They found that just the small-scale fishing fleet produced 3.4 Mt of CO₂e in 2014. To put this into context, Mexico's national inventory for the domestic shipping sector, but not accounting for fishing activity, in 2014 was recorded at just 2.2 Mt CO₂e, clearly placing into perspective the role of this fleet segment on national inventories [39].

Several authors have proposed using AIS to monitor the carbon emissions of the fleet [66–70]. Johansson et al. [71] proposed a new model Finnish Meteorological Institute - boat emissions and activities simulator (FMI-BEAM) to describe leisure boat fleet emissions in the Baltic Sea region with over 3000 dock locations, the national small boat registry, AIS data and vessel survey results. However, the method cannot cover countries with no national registry for small boats. Besides, small boats are not just leisure boats. Ugé et al. [72] estimated global ship emissions with the help of data from AIS. They used more than three billion daily AIS data records to create an activity database that captured ship size, speed, and meteorological and marine environmental conditions. This method is highly dependent on AIS data; however, these transponders are not normally installed on board small boats to capture their activity.

Zhang et al. [73] included unidentified vessels in the AIS-based vessel emission inventory. In doing so they developed an AIS-instrumented emissions inventory, including both identified and unidentified vessels. In particular, missing vessel parameters for unidentified vessels were estimated from a classification regression of similar vessel types and sizes in the AIS database. However, the authors did not discuss whether the regression model applies to vessels in most coastal areas. Nor did they explore regional vessel diversity in the database, so statistical inferences and levels of uncertainty about the applicability of their method to other unidentified vessels in a defined single region (e.g., small boats in the Gulf of California, Mexico) cannot be made.

Convolutional neural network architecture

Neural networks originate from the human perception of the brain. In 1943, American neuroscientists McCulloch and Pitts proposed a theory that every neuron is a multiple-input single-output structure [74]. Furthermore, there are only two possibilities for this output signal: zero or one, which is very similar to a computer.

In image recognition, a 7×7 image, for example, has 49 elements or cells. If 'X' is inputted to the grid, as shown in Fig. 1a, the computer will interpret it as a series of numbers (e.g., zeros and ones) as seen in Fig. 1b. If each cell is either black or white, for example, black can be assigned as one while white would be zero, resulting in a 7×7 matrix filled with zeros and ones. After feeding the algorithm as much data as is available, it will be trained to find parameters to determine if the object is an 'X' or not. For example, if it is a grey-scale picture, each number is neither zero nor one, but rather a grey-scale value from 0 to 255. If it is a colour image, it will use the red–green–blue (RGB) colour range. Essentially, no matter what the image is, it can be interpreted as a combination number inside a matrix, this eventually working as the input of the neural network. The goal of training a neural network is to find the parameters that make the loss function – it measures how far an estimated value is from its true value – smallest. However, the method described above is time-consuming and computationally expensive to train real-world images. Besides, the algorithm will be hard to recognise once the image is dilated, rotated or changed.

Based on the Neocognitron Model of Fukushima and Miyake [75], LeCun and Bengio [76] invented a practical method for image recognition, called the convolutional neural network. The role of convolution is to use a mathematical method to extract critical features from the image. This is achieved by extracting the features to use a convolution kernel to carry out the convolution operation. The convolution kernel is a matrix, usually 3×3 or 5×5 . For instance, if the convolution kernel is 3×3 , see Fig. 1c, then a convolution operation will be undertaken with the 7×7 'X' matrix (Fig. 1b) and the kernel (Fig. 1c). The operation result is also known as a feature map (Fig. 1d) [77].

The feature map reinforces the features of the convolution kernel. The 3×3 convolution kernel portrayed in Fig. 1c has only three oblique blocks of pixels that are ones. So if the original 7×7 matrix (Fig. 1b) also has diagonal pixel blocks of ones, the number would be extensive when the convolution operation is complete, which means the desired feature has been extracted. The smaller the value of the pixel block in the other positions of the feature map (Fig. 1d), the less it satisfies the feature. In general, different convolution kernels make it possible to achieve different feature maps.

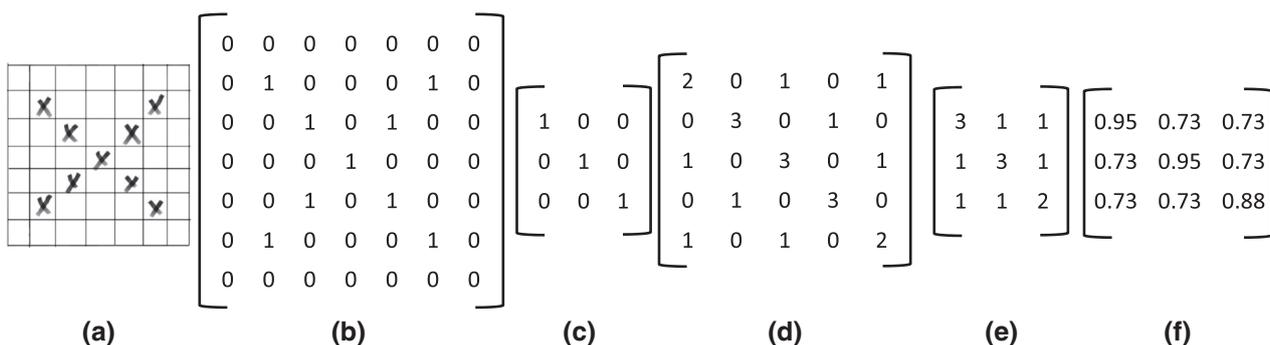
The next step after convolution is pooling. The pooling method can reduce the feature map size and maintain similar features to the feature map before the pooling process. Figure 1e shows the relatively small feature map after pooling the 5×5 matrix (Fig. 1d).

The step after pooling is activation. The activation function decides whether the neuron should be activated by computing the weighted sum and further adding the bias. The essence of the activation function is to introduce nonlinear factors to solve problems that a linear model cannot solve [78]. For example, after activating the sigmoid function, each element in the feature map would be between zero and one, as shown in Fig. 1f.

It is worth noting that the initial convolution kernel may be artificially set. Nevertheless, machine learning will go backwards to adjust and find the most suitable convolution kernel based on its data. As an image generally has many features, there will be many corresponding convolution kernels. After many convolutions and poolings, features can be found, including the diagonal lines of the image, the contours and the colour features. This information is taken and fed into the fully connected network for training, and it is finally possible to determine what the image is.

Figure 1

From left to right: (a) Letter X in a 7×7 image; (b) letter X in a 7×7 matrix; (c) a 3×3 convolution kernel; (d) a 5×5 feature map; (e) a 3×3 feature map after pooling; (f) a 3×3 feature map after activating with sigmoid function.



Convolutional neural networks in image recognition

The above literature review has demonstrated that the past literature on shipping carbon inventories has not focused on small boats. Thus, the topic of activity-based emission inventories for this segment is an important gap in the literature. There is still considerable work to be done to understand how the small boat fleet operates, what fuels are used, and the level of activity. However, with the development and maturation of a range of computer vision techniques such as CNNs, it may be possible to accurately identify small vessels from open satellite imagery and support understanding of this segment of shipping.

One of the computer vision's most fundamental and challenging problems is target detection. The main goal of target detection is to determine the location of an object in an image based on a large number of predefined classes. Deep learning techniques, which have emerged in recent years, are a powerful method for learning features directly from data and have led to significant breakthroughs in the field of target detection. Furthermore, with the rise of self-driving cars and face detection, the need for fast and accurate object detection is growing.

In 2012, AlexNet, a deep CNN (DCNN) proposed by Krizhevsky et al. [79], achieved record accuracy in image classification at the ImageNet Large-Scale Visual Recognition Challenge (ILSRVC), making CNNs the dominant paradigm for image recognition. Next, Girshick et al. [80] introduced Region-based Convolutional Neural Networks (R-CNN), the first CNN-based object detection method. The R-CNN algorithm represents a two-step approach in which a region proposal is generated first, and then a CNN is used for recognition and classification. Compared to the traditional sliding convolutional window to determine the possible regions of objects, R-CNN uses a selective search to pre-extract some candidate regions that are more likely to object in order to avoid computationally costly classification and object searches, which makes it faster and significantly less computationally expensive [80,81]. Overall, the R-CNN approach is divided into four steps:

1. Generate candidate regions.
2. Extract features using CNN on the candidate regions.
3. Feed the extracted features into a support vector machine (SVM) classifier.
4. Correct the object positions by using a regressor.

However, R-CNN also has drawbacks: the selective search method is slow in generating positive and negative sample candidate regions for the training network, which affects the overall speed of the algorithm; R-CNN needs to perform feature extraction once for each generated candidate region separately; there are a large number of repeated operations which limits the algorithm performance [82].

Since its inception, R-CNN has undergone several developments and iterations: Fast R-CNN, Faster R-CNN and Mask R-CNN [83–85]. The improvement of Fast R-CNN is the design of a pooling layer structure for the region of interest (ROI). The pooling stage effectively solves the R-CNN operation that crops and scales image regions to the same size, speeding up the algorithm. Faster R-CNN replaces the selective search method with the region proposal network (RPN) [84]. The selection and judgment of candidate frames are handed over to the RPN for processing, and candidate regions are subjected to multi-task loss-based classification and localisation processes.

Several CNN-based object detection frameworks have recently emerged that can run faster, have a higher detection accuracy, produce cleaner results and are easier to develop. Compared to the Faster R-CNN model, the You Only Look Once (YOLO) model can better detect smaller objects, that is, traffic lights at a distance [86], which is important when detecting objects in satellite images. Also, the YOLO model has a faster end-to-end run time and detection accuracy than the Faster R-CNN [86]. Mask R-CNN upgrades the ROI pooling layer of the Fast R-CNN to an ROI align layer and adds a branching FCN layer, the mask layer, to the bounding box recognition for semantic mask recognition [85]. Thus, the Mask R-CNN is essentially an instance segmentation algorithm, compared to semantic segmentation.¹ Instance segmentation is a more fine-grained segmentation of similar objects than semantic segmentation.

However, even traditional CNNs can be very useful for large-scale image recognition. For example, Simonyan and Zisserman [87] researched the effect of convolutional network depth on its accuracy in large-scale image recognition settings. Their research found that even with small (3 × 3) convolution filters, significant accuracy is achieved by pushing the depth from 16 to 19 weight layers.

In this research, the YOLO framework was selected. It uses a multi-scale detection method, which enables it to detect objects at different scales and to adapt to changes in the size and shape of the objects being observed [88]. Besides, YOLO is highly effective in detecting small objects with high accuracy and precision [89]. This makes it an ideal choice for detecting small objects in satellite imagery contexts, such as small boats in coastal waters. Additionally, YOLO is highly scalable, making it suitable for use in large-scale applications [90].

Finally, this study intends to develop the first stages of BoatNet. This image recognition model aims at detecting small boats, especially leisure and fishing boats in any sea area which, in turn and with further development, could significantly reduce uncertainty in the estimation of small boat fleet emission inventories in countries where access to tracking infrastructure, costly satellite databases and labour-intensive methodologies are important barriers.

Convolutional neural network configurations

Target areas in the Gulf of California and dataset statistical analysis

The Gulf of California in Mexico was chosen as an area of study. Ideally, to analyse a sufficient amount of satellite image data, the ports of each of the major harbour cities in the Gulf of California would need to be included in the scope of our study. Thus, the first step in this work was to determine if there was enough satellite data for the area. In this study, the Gulf of California was split into a few zones based on the Mexican state limits: (1) Baja California, (2) Sinaloa, (3) Sonora and (4) Baja California Sur. The satellite dataset used in this analysis included 690 high-resolution (4800 pixels × 2908 pixels) images of ships collected from Google Earth, where the imagery sources are Maxar Technologies and CNES/Airbus. From the imagery dataset, a statistical analysis was performed on how many times, temporally speaking, the satellite database captured the region of interest. As a result of this analysis, it was found that:

- most cities in the Gulf of California do not have enough open-access satellite data in 2018 and 2021, while many cities have relatively rich satellite data between 2019 and 2020;
- there has been a steady increase in the collection of satellite data in the Gulf of California from 2018 to 2020;
- the open-access and high-quality satellite data from Google Earth Pro is not immediately available to the public;
- differences in data accessibility are still evident among different cities. For example, Guaymas in the state of Sonora has rich satellite images in 2019 and 2020. However, other cities, such as La Ventana in the state of Baja California Sur, did not appear on Google Earth Pro between 2019 and 2020.

For this reason, continuing with the previous strategy of analysing the satellite data for each city in the Gulf of California would lead to a relatively large information bias and thus would not achieve an effective object detection model. Therefore, the following three cities with the richest data-accessibility in Google Earth Pro were chosen as the target areas for this study: Santa Rosalia, Loreto and Guaymas (see Fig. 2 for their geographical locations). The number of times captured by Google Earth Pro [91] is shown in Fig. 3 with a database of 583 images with timestamps between 2019 and 2020 for the three Mexican coastal cities.

Preprocessing

Each satellite image used for training was manually pre-labelled with a highly precise label box [92]. The original dataset contained images larger than 9 MB, which is an efficiency burden for neural network training, especially when few objects are detected. For this reason, all images were resized from 4800 pixels × 2908 pixels to 416 pixels × 416 pixels, with the file sizes reduced to between 10 KB and 40 KB [93].

Each satellite image for targeting or testing can be directly extracted from Google Earth Pro. Before downloading these images, a few things were done initially. First, all the layers from Google Earth Pro needed to be removed. Then, it was necessary to open the 'Navigation' tab of the 'Preferences' menu; 'Do not automatically tilt while zooming' needs to be clicked. This allowed the images

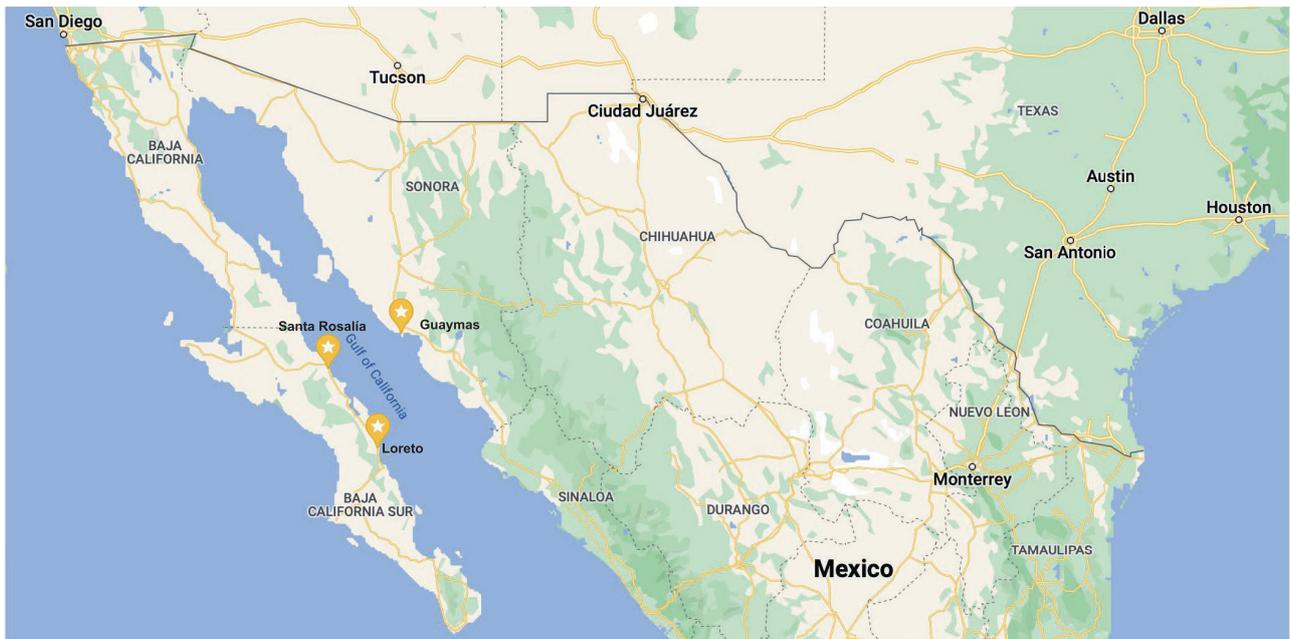


Figure 2

The geographic locations of the three target cities – Santa Rosalia, Loreto, and Guaymas. (Source: Google Maps 2021.)

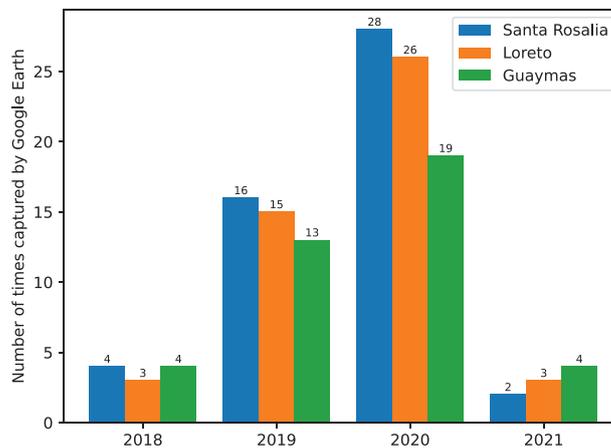


Figure 3

Number of times the three cities (Santa Rosalia, Loreto, Guaymas) were captured by Google Earth Pro from 2018 to 2021.

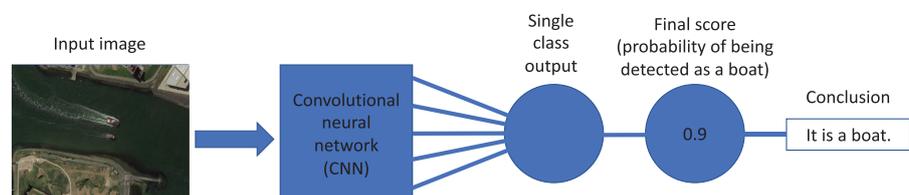
available to be acquired which were directly above sea level. Finally, the eye altitude was set to 200 m and the images were saved in 4800 pixels × 2908 pixels.

Single object detection architecture

Figure 4 shows a schematic of the models being used for detecting boats, where satellite images in the Gulf of California are the input of a pre-trained CNN. The detection accuracy was determined by computing the mean probability score from the Gulf’s satellite images. In the section Convolutional neural networks in image recognition, recent literature and the development of CNNs, including the YOLO model, were discussed. YOLO version 5 (YOLOv5) has four different categories of models, YOLOv5s, YOLOv5m, YOLOv5l and YOLOv5x [94]. They have 7.3 million, 21.4 million, 47.0 million and 87.7 million parameters, respectively. The performance charts can be seen in Fig. 5, which shows that the YOLOv5l model can achieve higher average precision with the same faster computing speed. Thus, in this study, Google Colab’s Tesla P100 GPU² and the YOLOv5 framework were used.

Figure 4

Model architecture. Detection model architecture for obtaining a conclusion from an input satellite image of boats. Images are preprocessed and passed through a CNN. The model’s output is a score, $y \in (0,1)$, representing the probability of being detected as a boat.



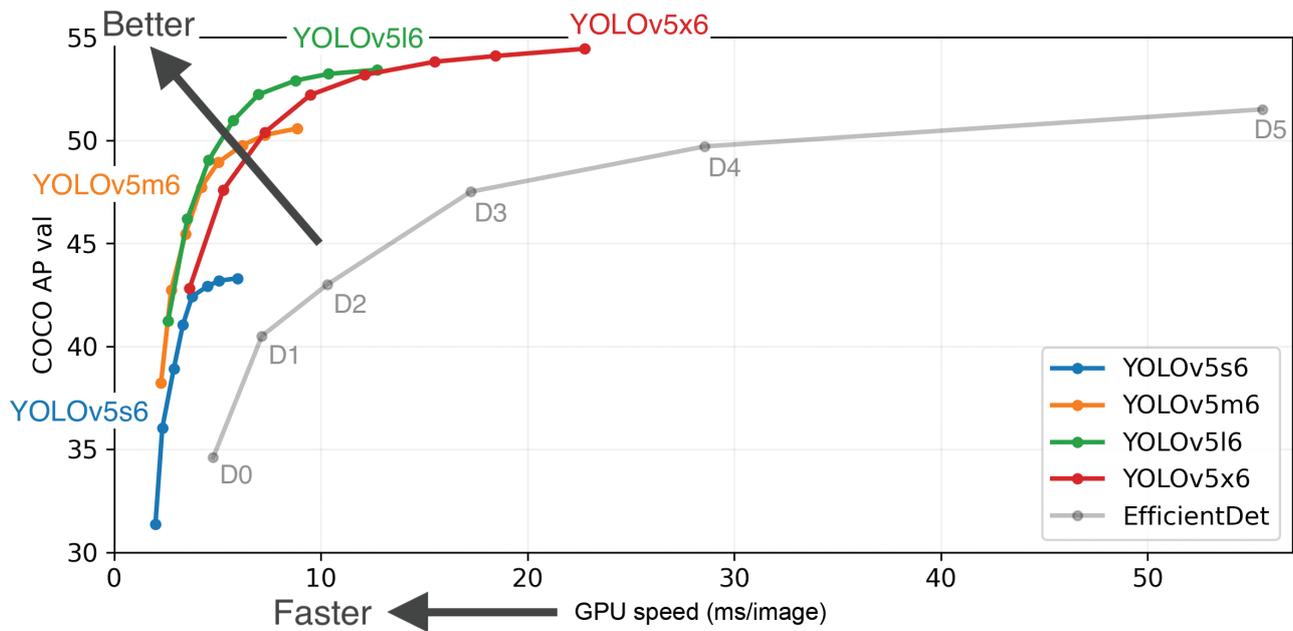


Figure 5

Average Precision (AP) versus GPU Speed in the 6th generation of YOLOv5 model under COCO data set [86,94].

Satellite images often contain noise such as shadows cast by water on the sea surface or haze clouds in the atmosphere, which make the training data inaccurate and often cause problems ensuring the model's correctness. He et al. [95] proposed a simple but effective image prior-dark channel before removing haze from a single input image. The prior-dark channel can be used as a statistic of outdoor haze-free images. Based on critical observation, most local patches in outdoor haze-free images contain some pixels whose intensity is very low in at least one colour channel. Using this prior-dark channel before the haze imaging model, the thickness of the haze can be estimated, and a high-quality haze-free image can be recovered. Moreover, a high-quality depth map can also be obtained as a byproduct of haze removal. In the same way, shadows can be removed using the prior-dark channel.

Similar to the principle of using convolution kernels, specific image kernels can sharpen the image. While the sharpening kernel does not produce a higher-resolution image, it emphasises the differences in adjacent pixel values, making the image appear more vivid. Overall, sharpening an image can significantly improve its recognition accuracy with a 5×5 image kernel.

Object measurement and classification

Measuring the length of a ship was one of the most challenging topics in this study. As Google Earth Pro does not provide an application programming interface (API) for accurate scales, manually measuring the size of a particular scale became the core process to calculate the size of any given ship. To achieve that it is important that all of the captured satellite images have the same eye altitude. By measuring only one real length of the object through the Google Earth Pro measurement tool and knowing the pixel length of this object, the length of one pixel in the satellite image of the fixed eye altitude can be calculated.

As the dataset for the training model was created with each edge tangent to the edge of the detected object, it can roughly treat the boat's length as the length of the diagonal within the detection box. Second, as the scale is central to the detection of the small boat fleet, the imagery scale should adhere to the following rules:

- Cannot be too large. The image should contain the full area in which boats may be found.
- Cannot be too small. If this is not followed it is highly probable that the group of boats are detected as a single but larger boat.
- Be sufficiently clear. This characteristic allows the algorithm to quantify the boat's length and accurately classify the measurements.

The eye altitude was set to 200 m based on the above rules. This study used a satellite image of Zurich Lake, Switzerland, on 16 August 2018 as the standard image for defining the scale (Fig. 6).

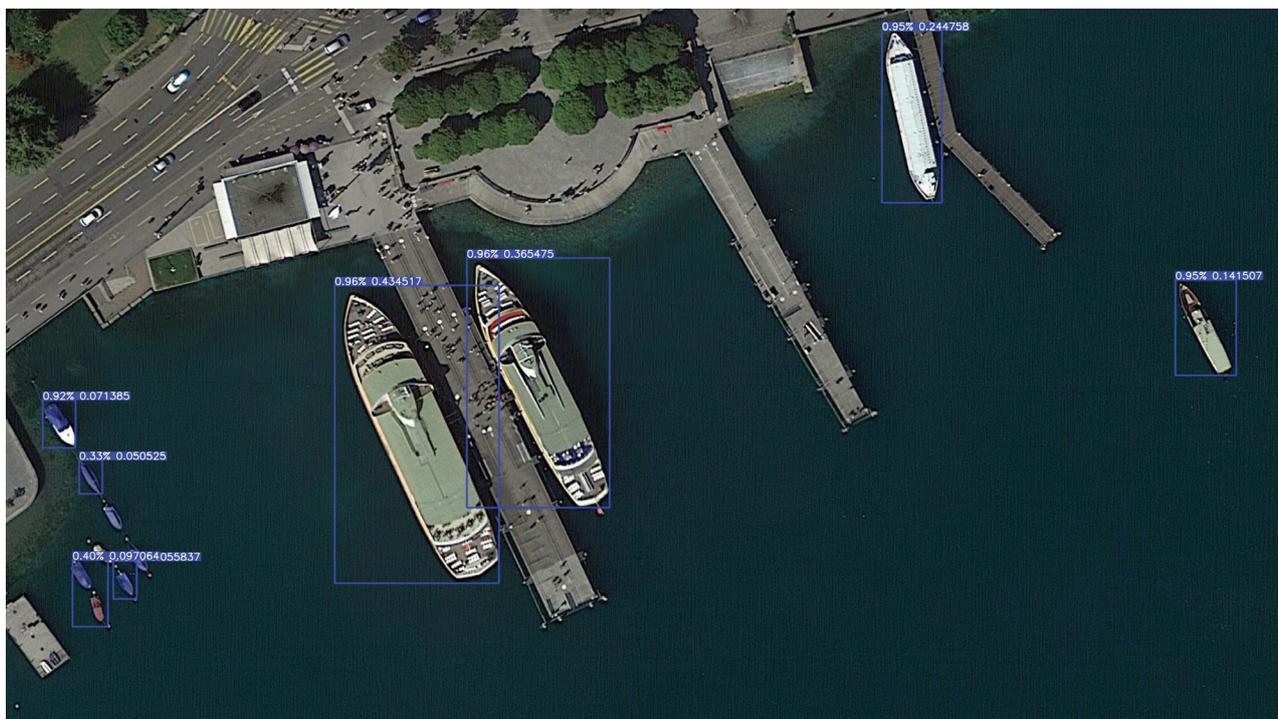


Figure 6

An image from Google Earth Pro for Zurich Lake on 16 August 2018 when eye alt is 200 m. (Source: Google Earth Pro, 2021.)

Compared with other regions, the satellite image of Zurich Lake complies with the rules, and it is a suitable candidate as the standard for measuring the length of small boats. This standard was then used for the rest of the imagery database.

A length measurement comparison was made between the BoatNet framework and Google Earth Pro to validate the length method. The vessel in Fig. 6 had a YOLO length of 0.43 that, after the scale conversion, represented 55.17 m. With an eye altitude of 200 m, the ratio of the absolute length to the YOLO length was approximately 127. Finally, after several verification tests, this ratio returned a small margin of error (around 1–3%) and hence was deemed a suitable scaling ratio for the remaining images. Moreover, having the same ratio and eye altitude was not enough. The resolution of each image must be the same, so measurements are standardised. For this purpose, all datasets that detect small boats will maintain a resolution of 3840×2160 pixels.

In a certain sense, large vessels (e.g., cargo ships) and small vessels (e.g., small boats for domestic use) are distinguished when creating the dataset for the area of interest. However, due to the scaling, some large vessels such as general cargo ships do not appear fully in an image. Hence they are not considered in the statistical results of this work. On the other hand, some of the larger vessels, slightly shorter in length than the 200 m eagle eye scale, are identified correctly by the algorithm and counted as part of the number of large vessels in the area. A Python script was then designed to count the number of small and large boats between regions.

After distinguishing between large and small boats, it is necessary to distinguish between small recreational boats for domestic use and fishing boats. The model used the detected deck colour of the small boat to distinguish between them. If the deck was predominantly white, it was assigned as a recreational boat, while any other mix of colour would be designated as a fishing boat. It is recognised that this is a broad and simplistic classification method, but it is an effective one to test the categorisation power of the model. Each of the detected boats (i.e., the objects within the four coordinate anchor box) were analysed whether the colour was white or mainly white to assign it to each category. As a final step, the model performed the category counting for each image, where a Python script was designed to count the small white boats.

Results

Train custom data: weights, biases logging, local logging

As shown in Fig. 7, the average accuracy, precision and recall of the model all show a significant increase with the model training number when the intersection over union (IOU)³ is between 50%

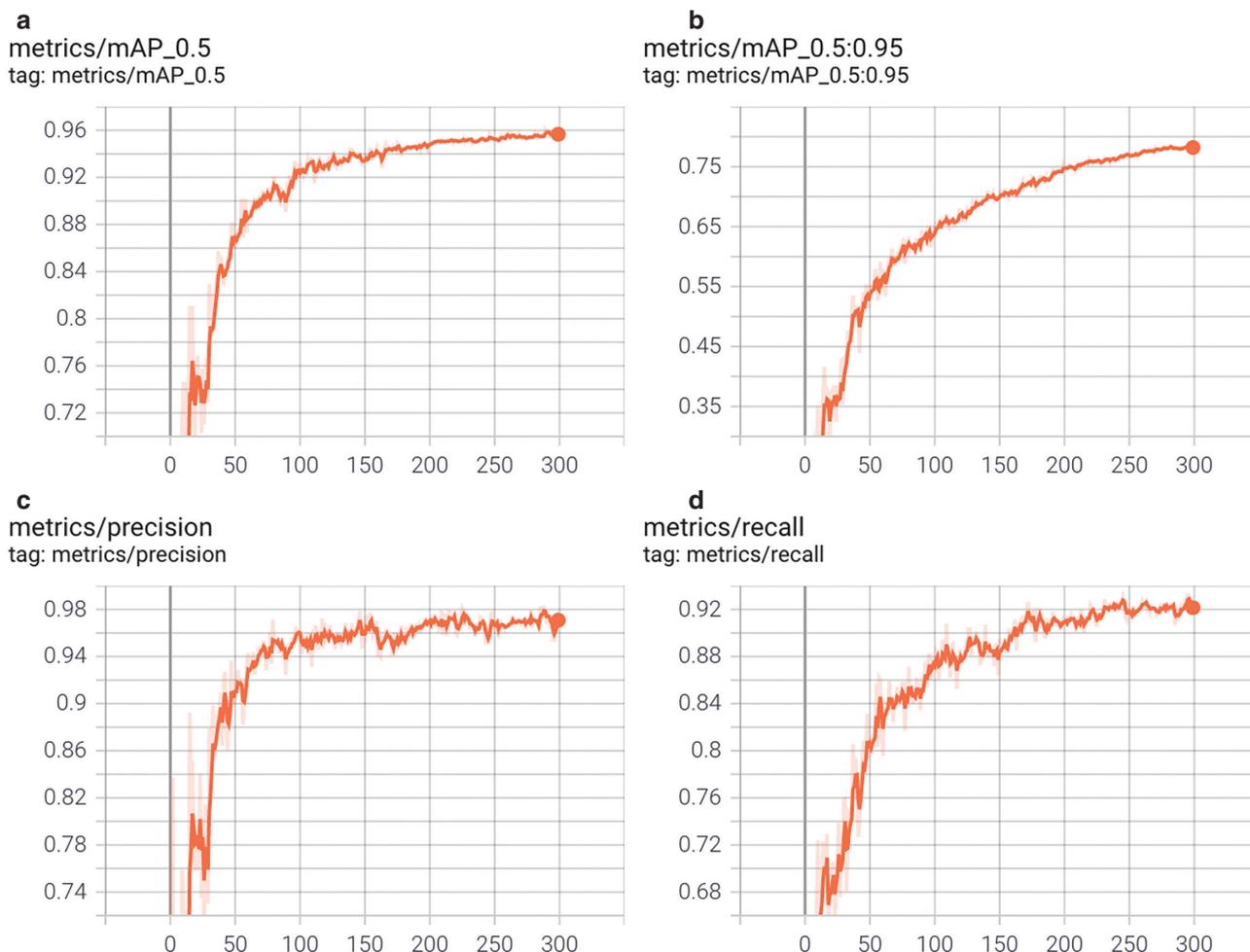


Figure 7

Average model precision when IOU is larger than 0.50; average model precision when IOU is between 0.50 and 0.95; The model precision; The model recall rate.

and 95%. In particular, the precision of the model can eventually reach a level close to 96%. However, this does not necessarily mean that the model will also fit the satellite imagery of the Gulf of California. First, such high-accuracy results only tell us that the model can achieve a relatively high recognition accuracy, which gradually increases and reaches 96% after 300 training repetitions. In the case that the algorithm needs to be trained for this area, consideration must be given to purposefully selecting many small boats in or near the area as a data source for training the model.

To train models faster, the images' resolution was reduced by about 70 times, resulting in images of 416 pixels × 416 pixels. The training could otherwise take two weeks if the images used had a resolution of 4800 pixels × 2908 pixels.

Similarly, as shown in Fig. 8, the loss rate of the box can eventually reach 1% as the number of training sessions increases. As this study has defined only one class of object (i.e., boat), the probability that the detection box does not detect that it is a boat at all is 1%. Similarly, because there is only one class, the class loss rate is zero. Figure 9 presents the prediction results during the training of the model, and shows that the model can detect the presence of vessels in 100% of the tested ranges and gives the corresponding range boxes. Most detected boats have a 90% probability of being boats, an acceptable value for object detection. As only one class was set, some were also considered a 100% probability of being boats.

Detection results and small boat composition

Starting with the length measurement comparison of a detected small boat and a large ship by BoatNet against Google Earth Pro measuring tools, Fig. 10 shows that the small boat detected

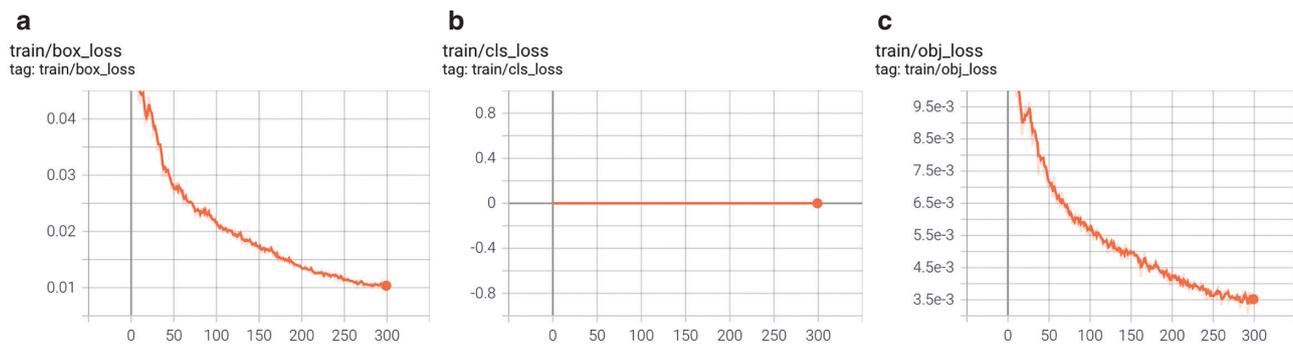


Figure 8

The box loss rate of the model; the class loss rate of the model; the object loss rate of the model.

measured 6.98 m using Google Earth Pro, while BoatNet estimated 6.74 m. The error between them is 3.4%. Google Earth Pro measured the larger ship at 41.38 m and BoatNet at 40.98 m. The error between them is less than 1.0%.

As explained in the section Single object detection architecture, it was unproductive for the model to select the entire region for the study due to the varying amount of publicly available regional images from Google Earth Pro over the past three years. Ultimately, the satellite image database was built from 690 images. However, as stated in that section, some of the slightly earlier satellite images offered inferior detail representation capabilities, which resulted in the model not accurately detecting the target's features. To improve model accuracy, an image enhancement process using a 5×5 sharpening kernel allowed for a higher recognition rate. However, the following situations still occur:

1. **Figure 11:** When the detailed representation of the image is indigent, that is, the images are blurred, and two or three small boats are moored together, the model is very likely to recognise the boats as a whole. There are two reasons for this problem. First, the training data is primarily a 'fuzzy' data source. Thus, when two or three small boats are moored together, the model cannot easily detect the features of each small boat individually. In contrast, it may seem more reasonable to the model that the boats as a whole have the same features. The second reason is that most data sources are individual boats on the surface or boats docked close to each other. As the data sources do not fully consider the fuzzy nature of the detail needed to detect the object and the fact that they are too close together, the model naturally does not recognise such cases.
2. **Figure 12:** When a large cargo ship is moored, the ship appears as a 'rectangle' from the air, much like a long pier, and is sometimes undetectable because small vessels with a rectangular shape were not common at the time the data feed was compiled. This also applies to uncommon vessels such as battleships. This could be corrected if the model considered larger ships, but this was outside the scope of this work.
3. **Figure 13:** The recognition rate was also significantly lower when the boats sometimes lay on the beach rather than floating on the water. This is because most of the training data are based on images in the water rather than boats on the beach.

Nevertheless, as **Figs 11–13** demonstrate, the model still detects most small boats in poorly detailed satellite images, even those that the human eye cannot easily detect. The number of small and large ships between regions can be seen in **Figs 14** and **15**. Two different types of port cities are exemplified by Guaymas, Loreto and Santa Rosalia:

1. Santa Rosalia and Loreto have a much smaller number of small boats and almost no large ships.
2. The port of Guaymas presented a larger number of small boats when contrasted to the other two coastal cities. There were between 1.37 and 8.00 times more small boats detected in Guaymas than in Santa Rosalia (dependent on the month and year of the image) and 3.00 times more than in Loreto.
3. Guaymas also had a larger number of large ships. There were more than 10 times more large ships detected in Guaymas than in Santa Rosalia and Loreto.
4. In the relatively large seaports of Guaymas and Loreto, there is a tendency for both large and small vessels to decrease with time.

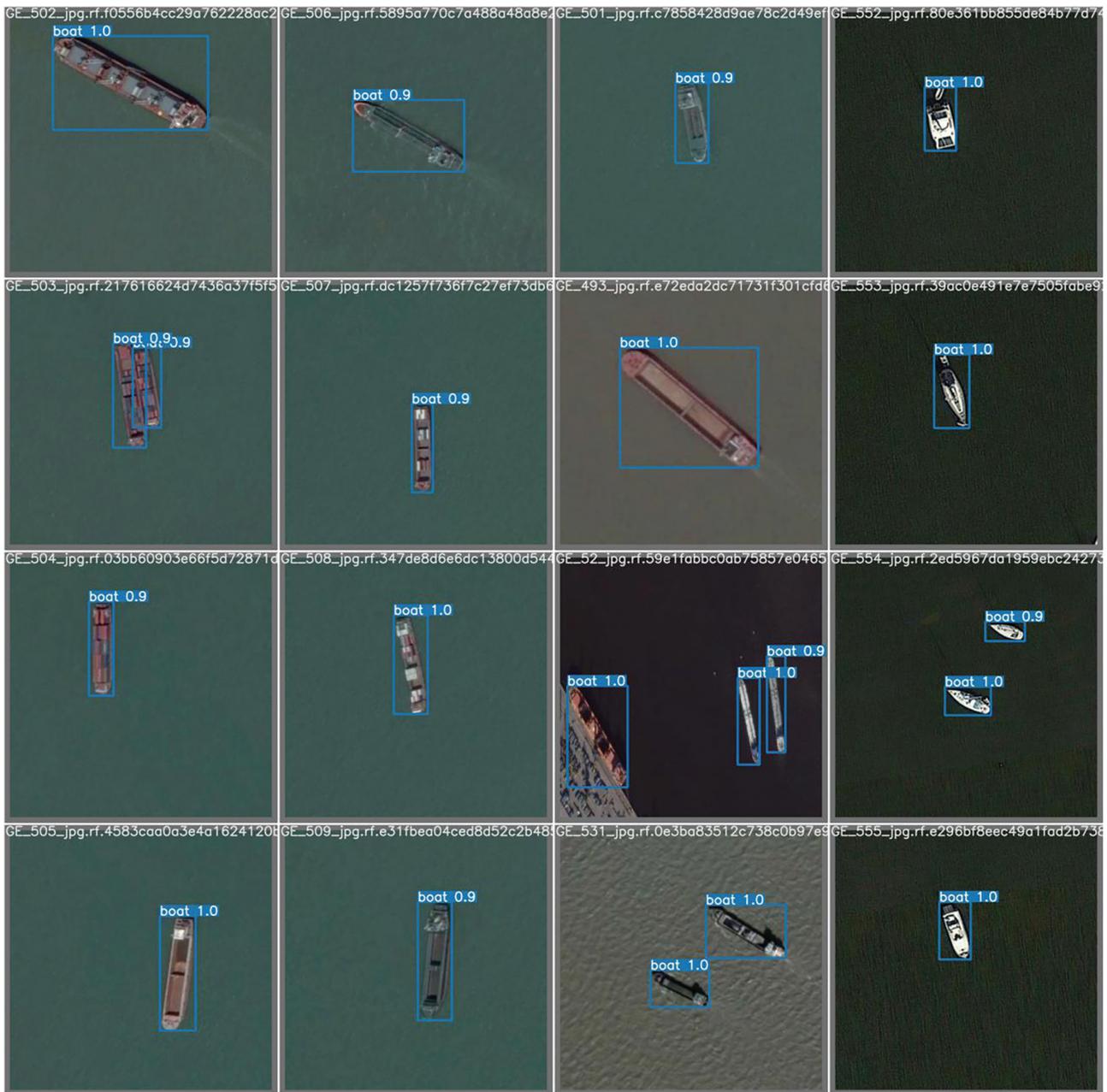


Figure 9

Test result of a trained model for detecting ships. (Source: Author-originated, based on Google Earth Pro imagery, 2021).

In fact, according to statistics [96] from the Mexican government, in 2020, the populations of Guaymas, Loreto and Santa Rosalia were 156,863, 18,052 and 14,357, respectively. Therefore, the results infer that the number of detected boats is correlated to the number of habitats, which makes sense as, by probability, there would be more economical and leisure activities around larger coastal cities.

Leisure and fishing boats in the Gulf of California

According to the statement in the section Single object detection architecture, determining whether a boat is white can be used as a criterion to determine whether a boat is used for recreation or fishing. As seen in Table 1, most of the small boats captured in the photo of Guaymas in 2020 are white (i.e., all can be classified as recreational boats). However, as previously discussed, this conclusion is limited as the algorithm does not consider, for example, other colours as part of the characteristics of leisure boats. Furthermore, although there are many uncertainties in detecting the colour of the boats, the algorithm also considers situations where the colour is not fully white due to atmospheric refraction, weather conditions or cloud interference. The algorithm also considers cases where the boat's colour is light. Therefore, this approach is acceptable from the point of view of algorithm complexity, results and detecting data quality.

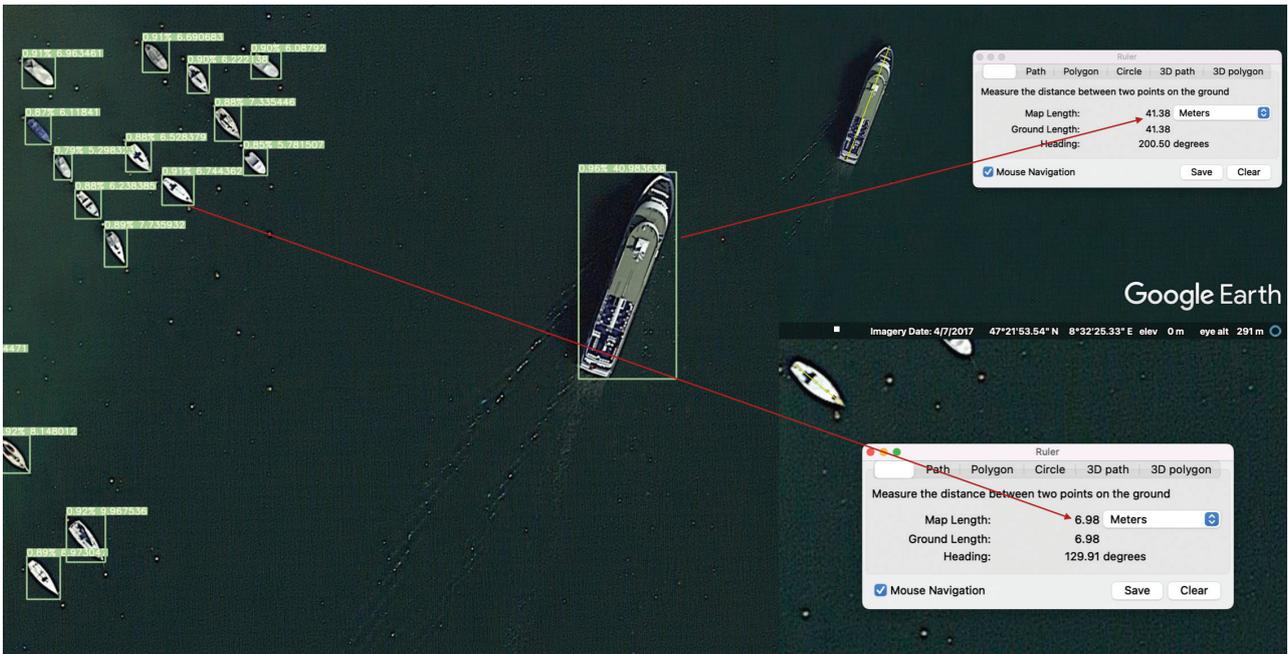


Figure 10

Comparing the length of boats using Google Earth ruler and computer vision algorithm. This example shows the image from Google Earth Pro for Zurich Lake on 16 August 2018 when the eye alt was 200 m. (Source: Author-originated, based on Google Earth Pro imagery, 2021).

Discussion

This study demonstrated the capabilities of a deep learning approach for the automatic detection and identification of small boats in the waters surrounding three cities in the Gulf of California with a precision of up to 74.0%. This work used CNNs to identify types of small vessels. Specifically, this study presented an image detection model, BoatNet, capable of distinguishing small boats in the Gulf of California with an accuracy of up to 93.9%, which is an encouraging result considering the high variability of the input images.

Even with the model's level of performance using large and highly ambiguous training images, it was found that image sharpening improved model accuracy. This implies that access to better quality imagery, such as that available through paid-for services, should considerably improve model precision and training times.

The results of this research have several important implications. First, the study used satellite data to predict the number and types of ships in three important cities in the Gulf of California. The resulting analysis can contribute to the region's shipping fleet composition, level of activity

Figure 11

When small boats are moored closely together in the harbour, the model may recognise two small boats as one. The image is from Guaymas, January 2020. (Source: Author-originated, based on Google Earth Pro imagery, 2021).

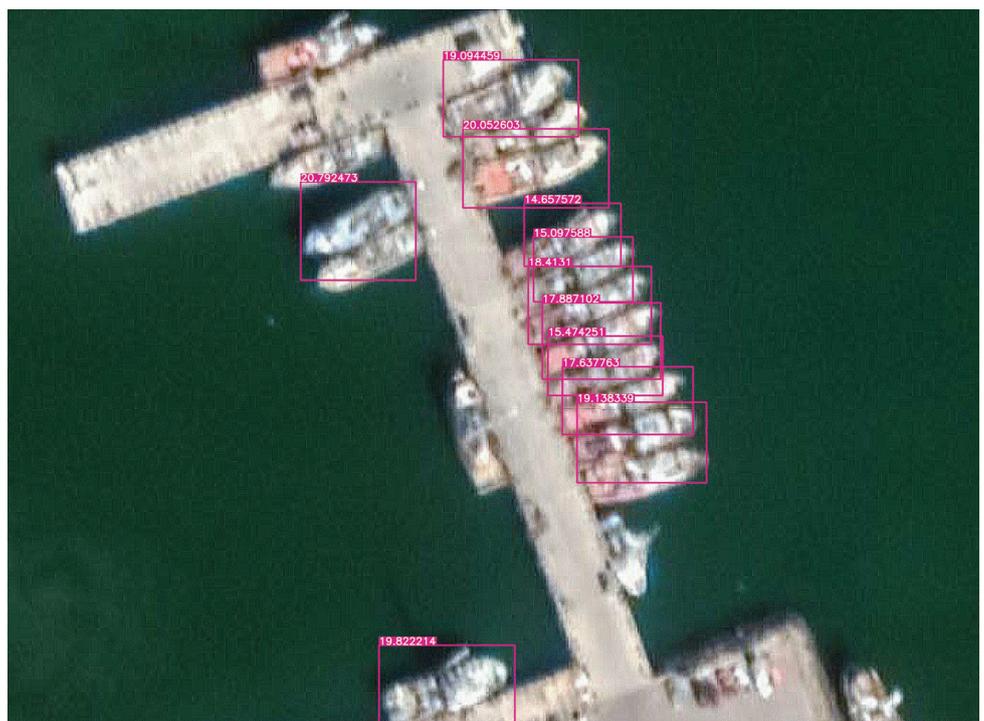


Figure 12

When the cargo ship is full of cargo, the ship looks like a rectangular jetty from above and loses the normal shape of a ship. The image is from Guaymas, January 2020. (Source: Author-originated, based on Google Earth Pro imagery, 2021).



and ultimately their carbon inventory by adding the emissions produced by the small boat fleet. Furthermore, through this approach, it is also possible to assign emissions into regions supporting the development of policies that can mitigate local GHG and air pollution. In addition, the transfer learning algorithm can be pre-trained in advance and immediately applied to any sea area worldwide. This will provide a potential method to increase efficiency for scientists and engineers worldwide who need to estimate local maritime emissions. In addition, the model can quickly and accurately identify the boat's length and classify them, allowing researchers to allocate more time to the vessels they need concentrate on, not just small boats. Finally, all of the above benefits can be exploited in under served areas with a shortage of infrastructure and resources.

This work is the first step to building emission inventories through image recognition, and it has some limitations. The study considered the ship as a single detection object. It did not evaluate whether the model can improve the accuracy of identifying ships in the case of multiple

Figure 13

Models may have difficulty detecting small boats moored on the beach. The image is from Santa Rosalia, February 2021. (Source: Author-originated, based on Google Earth Pro imagery, 2021).



Figure 14

The seasonal daily average count of small boats in Santa Rosalia, Loreto and Guaymas between 2019 and 2020.

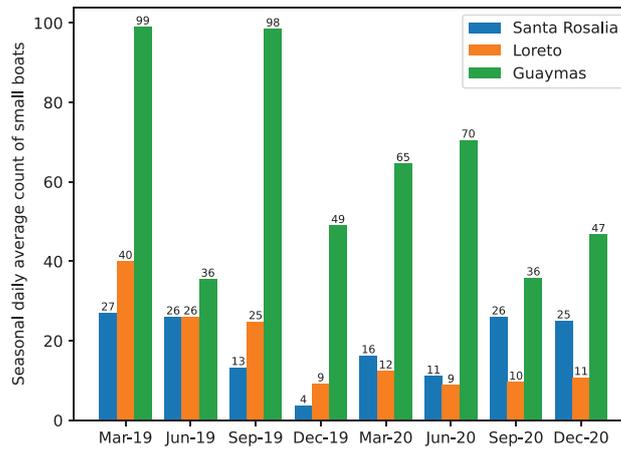
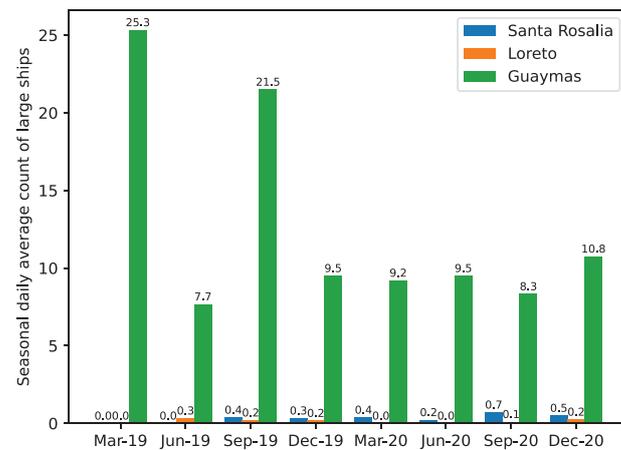


Figure 15

The seasonal daily average count of large ships in Santa Rosalia, Loreto and Guaymas between 2019 and 2020.



detection objects. For instance, BoatNet was not trained to detect docks to improve the metrics of detecting boats. By down-sampling the image to 416 pixels × 416 pixels, it is possible to mask some of the boats at the edges of the photograph. Furthermore, deep learning models train faster on small images [97]. A larger input image requires the neural network to learn from four times as many pixels, increasing the architecture’s training time. In this work, a considerable proportion of the images in the dataset were large images of 4800 pixels × 2908 pixels. Thus, BoatNet was set to learn from resized small images measuring 416 pixels × 416 pixels. Due to the low data quality of the selected regions, the images are less suitable as training datasets. However, using datasets from other regions or higher-quality open-source imagery may result in inaccurate coverage of all types of ships in the region. When focusing on the small boat categorisation and the data used, understanding the implications of different environments (e.g., water or land) on object classification accuracy through the AI fairness principle deserves further study. From this point of view, large-scale collection of data sources in the real physical world would be costly and time-consuming. That said, it is possible that reinforcement learning, or building simulations in the virtual world, could reduce the negative impact of the environment on object recognition and thus improve its categorisation precision. Of all these limitations, model detection still achieves excellent performance in detecting and classifying small boats. To enrich the analysis, one of

Table 1. Detection example. Small boats, large ships and small white boats in Guaymas in 2020

	Mar 20	Jun 20	Sep 20	Dec 20
Small boats	323	283	215	187
Large ships	46	38	50	43
Small white boats	302	273	210	174
Shipping boats	21	10	5	13

the future works planned is comparing research results with different algorithms for the same problem.

It is important to remember that BoatNet currently only detects and classifies certain types of small boats. Therefore, to estimate fuel consumption and emissions, it is necessary to couple it with small boat behaviour datasets [65], typical machinery, fuel characteristics and emission factors unique to this maritime segment [39].

Finally, this work has demonstrated that deep learning models have the potential to identify small boats in extreme environments at performance levels that provide practical value. With further analysis and small boat data sources, these methods may eventually allow for the rapid assessment of shipping carbon inventories.

Notes

- 1 In computer vision, image segmentation is the process of partitioning an image into multiple image segments, also known as image regions or image objects.
- 2 Colab Pro limits RAM to 32 GB while Pro+ limits RAM to 52 GB. Colab Pro and Pro+ limit sessions to 24 hours.
- 3 IOU is an evaluation metric used to measure the accuracy of an object detector on a particular dataset.

Acknowledgements

We thank UCL Energy Institute, University College London for supporting the research, and academic colleagues and fellows inside and outside University College London: Andrea Grech La Rosa, UCL Research Computing, Edward Gryspeerdt, Tom Lutherborrow.

Open data and materials availability statement

The datasets generated during and/or analysed during the current study are available in the repository: <https://github.com/theiresearch/BoatNet>.

Declarations and conflicts of interest statement

Research ethics statement

Not applicable to this article.

Consent for publication statement

The authors declare that research participants' informed consent to publication of findings – including photos, videos and any personal or identifiable information – was secured prior to publication.

Conflicts of interest statement

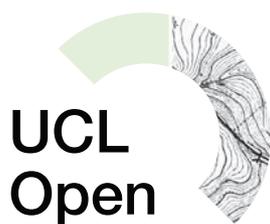
The authors declare no conflicts of interest with this work.

References

- [1] Pörtner H-O, Roberts DC, Tignor M, Poloczanska ES, Mintenbeck K, Alegra A, et al. Climate change 2022: impacts, adaptation and vulnerability. Cambridge: Cambridge University Press; 2022.
- [2] Raturi AK. Renewables 2019 global status report. Renewable energy policy network for the 21st century. 2019.
- [3] BP. Statistical review of world energy 2021. BP Statistical Review, London, UK; 2021.
- [4] IEA. World energy outlook 2021 [online]. 2021. Available from: <https://iea.blob.core.windows.net/assets/4ed140c1-c3f3-4fd9-acae-789a4e14a23c/WorldEnergyOutlook2021.pdf>.
- [5] Bhattacharyya SC. Fossil-fuel dependence and vulnerability of electricity generation: case of selected European countries. *Energy Policy*. 2009;37(6):2411–20.
- [6] DBIS. Russia–Ukraine and UK energy: factsheet. Department for Business, Energy & Industrial Strategy, UK Government. 2022. Available from: <https://www.gov.uk/government/news/russia-ukraine-and-uk-energy-factsheet>.
- [7] IRENA. Renewable power generation costs in 2020. International Renewable Energy Agency. 2021. Available from: <https://www.irena.org/publications/2021/Jun/Renewable-Power-Costs-in-2020>.

- [8] BP. Statistical review of world energy 2016. BP Statistical Review, London, UK; 2016.
- [9] Trading Economics. UK natural gas – 2022 data – 2020-2021 historical – 2023 forecast – price – quote. Trading Economics. 2022. Available from: <https://tradingeconomics.com/commodity/uk-natural-gas>.
- [10] Walker TR, Adebambo O, Feijoo MCDA, Elhaimer E, Hossain T, Edwards SJ, et al. Environmental effects of marine transportation. In: World seas: an environmental evaluation. Elsevier; 2019. p. 505–30.
- [11] Healy S. Greenhouse gas emissions from shipping: Waiting for concrete progress at IMO level. Policy department for economic, scientific and quality of life policies directorate-general for internal policies. 2020; 19.
- [12] IMO. Adoption of the initial IMO strategy on reduction of GHG emissions from ships and existing IMO activity related to reducing GHG Emissions in the shipping sector. 2018. Available from: https://unfccc.int/sites/default/files/resource/250_IMO%20submission_Talanoa%20Dialogue_April%202018.pdf.
- [13] UK Government. Operational standards for small vessels. 2021. [Accessed 30 August 2021]. Available from: <https://www.gov.uk/operational-standards-for-small-vessels>.
- [14] An K. E-navigation services for non-solas ships. *Int J e-Navigat Maritime Econ*. 2016;4:13–22.
- [15] Stateczny A. AIS and radar data fusion in maritime navigation. *Zeszyty Naukowe/Akademia Morska w Szczecinie*. 2004. p. 329–36.
- [16] Vachon P, Thomas S, Cranton J, Edel H, Henschel M. Validation of ship detection by the RADARSAT synthetic aperture radar and the ocean monitoring workstation. *Can J Remote Sens*. 2000;26(3):200–12.
- [17] Gillett R, Tauati MI. Fisheries of the Pacific Islands: regional and national information. FAO Fisheries and Aquaculture Technical Paper. 2018;625:1–400.
- [18] Calderón M, Illing D, Veiga J. Facilities for bunkering of liquefied natural gas in ports. *Transp Res Procedia*. 2016;14:2431–40.
- [19] Devanney J. The impact of the energy efficiency design index on very large crude carrier design and CO₂ emissions. *Sh Offshore Struct*. 2011;6(4):355–68.
- [20] Dedes EK, Hudson DA, Turnock SR. Assessing the potential of hybrid energy technology to reduce exhaust emissions from global shipping. *Energy Policy*. 2012;40:204–18.
- [21] Marty P, Corrigan P, Gondet A, Chenouard R, Hétet J-F. Modelling of energy flows and fuel consumption on board ships: application to a large modern cruise vessel and comparison with sea monitoring data. In: Proceedings of the 11th International Marine Design Conference, Glasgow, UK. 2012. p. 11–4.
- [22] Zhen X, Wang Y, Liu D. Bio-butanol as a new generation of clean alternative fuel for SI (spark ignition) and CI (compression ignition) engines. *Renew Energy*. 2020;147:2494–521.
- [23] Dwivedi G, Jain S, Sharma M. Impact analysis of biodiesel on engine performance – a review. *Renew Sustain Energy Rev*. 2011;15(9):4633–41.
- [24] Korakianitis T, Namasivayam A, Crookes R. Natural-gas fueled spark-ignition (SI) and compression-ignition (CI) engine performance and emissions. *Prog Energy Combust Sci*. 2011;37(1):89–112.
- [25] Cariou P, Parola F, Notteboom T. Towards low carbon global supply chains: a multi-trade analysis of CO₂ emission reductions in container shipping. *Int J Prod Econ*. 2019;208:17–28.
- [26] Halim RA, Kirstein L, Merk O, Martinez LM. Decarbonization pathways for international maritime transport: a model-based policy impact assessment. *Sustainability*. 2018;10(7):2243.
- [27] Rissman J, Bataille C, Masanet C, Aden N, Morrow WR III, Zhou N, et al. Technologies and policies to decarbonize global industry: review and assessment of mitigation drivers through 2070. *Appl Energy*. 2020;266:114848.
- [28] Buira D, Tovilla J, Farbes J, Jones R, Haley B, Gastelum D. A whole-economy deep decarbonization pathway for Mexico. *Energy Strategy Rev*. 2021;33:100578.
- [29] Adams WM, Jeanrenaud S. Transition to sustainability: towards a humane and diverse world. International Union for Conservation of Nature (IUCN); 2008.
- [30] Anika OC, Nnabuife SG, Bello A, Okoroafor RE, Kuang B, Villa R. Prospects of low and zero-carbon renewable fuels in 1.5-degree net zero emission actualisation by 2050: a critical review. *Carbon Capture Sci Technol*. 2022;5:100072.
- [31] Woo J, Fatima R, Kibert CJ, Newman RE, Tian Y, Srinivasan RS. Applying blockchain technology for building energy performance measurement, reporting, and verification (MRV) and the carbon credit market: a review of the literature. *Build Environ*. 2021;205:108199.
- [32] Psaraffis HN, Kontovas CA. CO₂ emission statistics for the world commercial fleet. *WMU J Marit Aff*. 2009;8(1):1–25.
- [33] Lindstad H, Asbjørnslett BE, Strømman AH. The importance of economies of scale for reductions in greenhouse gas emissions from shipping. *Energy Policy*. 2012;46:386–98.
- [34] Howitt OJ, Revol VG, Smith IJ, Rodger CJ. Carbon emissions from international cruise ship passengers' travel to and from New Zealand. *Energy Policy*. 2010;38(5):2552–60.
- [35] Dalsøren SB, Eide M, Endresen Ø, Mjelde A, Gravir G, Isaksen IS. Update on emissions and environmental impacts from the international fleet of ships: the contribution from major ship types and ports. *Atmos Chem Phys*. 2009;9(6):2171–94.
- [36] Sener S. *Balance nacional de energía 2017. España: Dirección general de planeación e información*; 2018.
- [37] UK Ship Register. UK small ship registration. 2021. [Accessed 30 August 2021]. Available from: <https://www.ukshipregister.co.uk/registration/leisure/>.
- [38] Mexico National Aquaculture and Fisheries Commission. Registered vessels. 2021. [Accessed 30 August 2021]. Available from: https://www.conapesca.gob.mx/wb/cona/embarcaciones_registradas.
- [39] INECC. Inventario nacional de emisiones de gases y compuestos de efecto invernadero (INEGYCEI). 2020.
- [40] Change I. 2006 IPCC guidelines for national greenhouse gas inventories. Hayama, Kanagawa, Japan: Institute for Global Environmental Strategies; 2006.
- [41] Lluç-Cota SE, Aragón-Noriega EA, Arregun-Sánchez F, Aurióles-Gamboa D, Jesús Bautista-Romero J, Brusca RC, et al. The Gulf of California: review of ecosystem status and sustainability challenges. *Prog Oceanogr* [online]. 2007;73(1):1–26. Available from: <https://www.sciencedirect.com/science/article/pii/S0079661107000134>.
- [42] Munguia-Vega A, Green AL, Suarez-Castillo AN, Espinosa-Romero MJ, Aburto-Oropeza O, Cisneros-Montemayor AM, et al. Ecological guidelines for designing networks of marine reserves in the unique biophysical environment of the Gulf of California. *Rev Fish Biol Fish*. 2018;28(4):749–76.

- [43] Marinone S. Seasonal surface connectivity in the gulf of california. *Estuar Coast Shelf Sci* [online]. 2012;100:133–41. Available from: <https://www.sciencedirect.com/science/article/pii/S0272771412000169>.
- [44] Meltzer L, Chang JO. Export market influence on the development of the Pacific shrimp fishery of Sonora, Mexico. *Ocean Coast Manag* [online]. 2006;49(3):222–235. Available from: <https://www.sciencedirect.com/science/article/pii/S0964569106000329>.
- [45] Hernández-Trejo V, Germán GP-D, Lluch-Belda D, Beltrán-Morales L. Economic benefits of sport fishing in Los Cabos, Mexico: is the relative abundance a determinant? *J Sustain Tour*. 2012;161:165–76.
- [46] CNANP. Atlas Interactivo de las Áreas Naturales Protegidas de México. [Accessed 26 April 2023]. Available from: <http://sig.conanp.gob.mx/website/interactivo/atlas/>.
- [47] SMARN. Islas y Áreas Protegidas del Golfo de California. [Accessed 5 April 2022]. Available from: <https://www.gob.mx/semarnat/articulos/islas-y-areas-protegidas-del-golfo-de-california-269050>.
- [48] Manfreda S, McCabe MF, Miller PE, Lucas R, Pajuelo Madrigal V, Mallinis G, et al. On the use of unmanned aerial systems for environmental monitoring. *Remote Sens*. 2018;10(4):641.
- [49] Chang N-B, Bai K, Imen S, Chen C-F, Gao W. Multisensor satellite image fusion and networking for all-weather environmental monitoring. *IEEE Syst J*. 2016;12(2):1341–57.
- [50] Willis KS. Remote sensing change detection for ecological monitoring in united states protected areas. *Biol Conserv*. 2015;182:233–42.
- [51] Pourbabaee B, Roshtkhari MJ, Khorasani K. Deep convolutional neural networks and learning ECG features for screening paroxysmal atrial fibrillation patients. *IEEE Trans Syst Man Cybern Syst*. 2017 Jun 1;48(12):2095–104.
- [52] Kumar MS, Ganesh D, Turukmane AV, Batta U, Sayyadliyakat KK. Deep convolution neural network based solution for detecting plant diseases. *J Pharm Negat Results*. 2022;13(1):464–71.
- [53] Jozdani SE, Johnson BA, Chen D. Comparing deep neural networks, ensemble classifiers, and support vector machine algorithms for object-based urban land use/land cover classification. *Remote Sens*. 2019;11(14):1713.
- [54] Shen D, Wu G, Suk H-I. Deep learning in medical image analysis. *Annu Rev Biomed Eng*. 2017;19:221.
- [55] Andrearczyk V, Whelan PF. Using filter banks in convolutional neural networks for texture classification. *Pattern Recognit Lett*. 2016;84:63–69.
- [56] Kieffer B, Babaie M, Kalra S, Tizhoosh HR. Convolutional neural networks for histopathology image classification: training vs. using pre-trained networks. In: 2017 seventh international conference on image processing theory, tools and applications (IPTA), IEEE; 2017. p. 1–6.
- [57] Sherrah J. Fully convolutional networks for dense semantic labelling of high-resolution aerial imagery. *arXiv preprint arXiv:1606.02585*. 2016.
- [58] Kampffmeyer M, Salberg A-B, Jenssen R. Semantic segmentation of small objects and modeling of uncertainty in urban remote sensing images using deep convolutional neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 2016. p. 1–9.
- [59] Sheykhmousa M, Kerle N, Kuffer M, Ghaffarian S. Post-disaster recovery assessment with machine learning-derived land cover and land use information. *Remote Sens*. 2019;11(10):1174.
- [60] Helber P, Bischke B, Dengel A, Borth D. EuroSAT: a novel dataset and deep learning benchmark for land use and land cover classification. *IEEE J Sel Top Appl Earth Obs Remote Sens*. 2019;12(7):2217–26.
- [61] Talukdar S, Singha P, Mahato S, Pal S, Liou Y-A, Rahman A. Land-use land-cover classification by machine learning classifiers for satellite observations – a review. *Remote Sens*. 2020;12(7):1135.
- [62] Parker RW, Blanchard JL, Gardner C, Green BS, Hartmann K, Tyedmers PH, et al. Fuel use and greenhouse gas emissions of world fisheries. *Nat Clim Change*. 2018;8(4):333–7.
- [63] Smith TW, Jalkanen JP, Anderson BA, Corbett JJ, Faber J, Hanayama S, et al. Third IMO GHG study 2014, International Maritime Organization (IMO), London, UK; 2015. Available from: <https://greenvoyage2050.imo.org/wp-content/uploads/2021/01/third-imo-ghg-study-2014-executive-summary-and-final-report.pdf>.
- [64] Greer K, Zeller D, Woroniak J, Coulter A, Winchester M, Palomares MD, et al. Global trends in carbon dioxide (CO₂) emissions from fuel combustion in marine fisheries from 1950 to 2016. *Marine Policy* [online]. 2019;107:103382. Available from: <https://www.sciencedirect.com/science/article/pii/S0308597X1730893X>.
- [65] Ferrer EM, Aburto-Oropeza O, Jimenez-Esquivel V, Cota-Nieto JJ, Mascarenas-Osorio I, Lopez-Sagastegui C. Mexican small-scale fisheries reveal new insights into low-carbon seafood and ‘climate-friendly’ fisheries management. *Fisheries*. 2021;46(6):277–87.
- [66] Traut M, Bows A, Walsh C, Wood R. Monitoring shipping emissions via AIS data? Certainly. In: Low Carbon Shipping Conference 2013, London; 2013.
- [67] Johansson L, Jalkanen J-P, Kukkonen J. A comprehensive modelling approach for the assessment of global shipping emissions. In: International technical meeting on air pollution modelling and its application. Springer; 2016. p. 367–73.
- [68] Mabunda A, Astito A, Hamdoune S. Estimating carbon dioxide and particulate matter emissions from ships using automatic identification system data. *Int J Comput Appl*. 2014;88:27–31.
- [69] Hensel T, Ugé C, Jahn C. Green shipping: using AIS data to assess global emissions. *Sustainability Management Forum | NachhaltigkeitsManagementForum*. 2020;28:39–47.
- [70] Han W, Yang W, Gao S. Real-time identification and tracking of emission from vessels based on automatic identification system data. In: 2016 13th international conference on service systems and service management (ICSSSM). 2016. p. 1–6.
- [71] Johansson L, Jalkanen J-P, Fridell E, Maljutenko I, Ytreberg E, Eriksson M, et al. Modeling of leisure craft emissions. In: International technical meeting on air pollution modelling and its application. Springer; 2018. p. 205–10.
- [72] Ugé C, Scheidweiler T, Jahn C. Estimation of worldwide ship emissions using AIS signals. In: 2020 European Navigation Conference (ENC). 2020. p. 1–10.
- [73] Zhang Y, Fung J, Chan J, Lau A. The significance of incorporating unidentified vessels into AIS-based ship emission inventory. *Atmos Environ*. 2019;203:102–13.
- [74] McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. *Bull Math Biophys*. 1943;5(4):115–33.
- [75] Fukushima K, Miyake S. Neocognitron: a self-organizing neural network model for a mechanism of visual pattern recognition. In: Competition and cooperation in neural nets. Springer; 1982. p. 267–85.



UCLPRESS

Extra information

UCL Open: Environment is an open scholarship publication, all previous versions and open peer review reports can be found online in the *UCL Open: Environment* Preprint server at ucl.scienceopen.com

- [76] LeCun Y, Bengio Y. Convolutional networks for images, speech, and time series. In: The handbook of brain theory and neural networks. 1995. p. 1995. Available from: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=e26cc4a1c717653f323715d751c8dea7461aa105>.
- [77] Bouvrie J. Notes on convolutional neural networks. 2006. Available from: http://web.mit.edu/jvb/www/papers/cnn_tutorial.pdf.
- [78] Lin G, Shen W. Research on convolutional neural network based on improved relu piecewise activation function. *Procedia Comput Sci*. 2018;131:977–984.
- [79] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst*. 2012;25:1097–105.
- [80] Girshick R, Donahue J, Darrell T, Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2014. p. 580–87.
- [81] Uijlings JR, Van De Sande KE, Gevers T, Smeulders AW. Selective search for object recognition. *Int J Comput Vis*. 2013;104(2):154–71.
- [82] Huang J, Rathod V, Sun C, Zhu M, Korattikara A, Fathi A, et al. Speed/accuracy trade-offs for modern convolutional object detectors. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2017. p. 7310–1.
- [83] Girshick R. Fast R-CNN. In: Proceedings of the IEEE international conference on computer vision. 2015. p. 1440–8.
- [84] Ren S, He K, Girshick R, Sun J. Faster R-CNN: towards real-time object detection with region proposal networks. *arXiv preprint arXiv:1506.01497*, 2015.
- [85] He K, Gkioxari G, Dollár P, Girshick R. Mask R-CNN. In: Proceedings of the IEEE international conference on computer vision. 2017. p. 2961–9.
- [86] Dwivedi P. Yolov5 compared to faster RCNN. Who wins? 2020. [Accessed 30 August 2021]. Available from: <https://towardsdatascience.com/yolov5-compared-to-faster-rcnn-who-wins-a771cd6c9fb4>.
- [87] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *CoRR*. 2015;abs/1409.1556.
- [88] Li Y, Huang H, Chen Q, Fan Q, Quan H. Research on a product quality monitoring method based on multi scale PP-YOLO. *IEEE Access*. 2021;9:80373–7.
- [89] Pham M-T, Courtrai L, Friguet C, Lefèvre S, Baussard A. Yolo-fine: one-stage detector of small objects under various backgrounds in remote sensing images. *Remote Sens*. 2020;12(15):2501.
- [90] Do J, Ferreira VC, Bobarshad H, Torabzadehkashi M, Rezaei S, Heydarigorji A, et al. Cost-effective, energy-efficient, and scalable storage computing for large-scale AI applications. *ACM Trans Storage (TOS)*. 2020;16(4): 1–37.
- [91] Lisle RJ. Google Earth: a new geological resource. *Geology Today*. 2006;22(1):29–32.
- [92] Lutherborrow T, Agoub A, Kada M. Ship detection in satellite imagery via convolutional neural networks. 2018. Available from: <https://www.semanticscholar.org/paper/SHIP-DETECTION-IN-SATELLITE-IMAGERY-VIA-NEURAL-Lutherborrow-Agoub/2928b63bd05c4e83a4b95b47857d010d70bfaac7>.
- [93] Nelson J. You might be resizing your images incorrectly. Jan 2020 [online]. Available from: <https://blog.roboflow.com/you-might-be-resizing-your-images-incorrectly/>.
- [94] Jocher G, Chaurasia A, Stoken A, Borovec J, NanoCode012, Kwon Y, et al. Ultralytics/YOLOv5: v6.1 – TensorRT, TensorFlow Edge TPU and OpenVINO Export and Inference [online]. Feb 2022. Available from: <https://doi.org/10.5281/zenodo.6222936>.
- [95] He K, Sun J, Tang X. Single image haze removal using dark channel prior. In: 2009 IEEE conference on computer vision and pattern recognition. 2009. p. 1956–63.
- [96] INEGI. Population and housing census 2020 in Mexico. [Accessed 11 June 2022]. Available from: <https://www.inegi.org.mx/app/cpv/2020/resultadosrapidos/>.
- [97] Tan M, Le Q. Efficientnetv2: smaller models and faster training. In: International conference on machine learning. PMLR; 2021. p. 10096–106.