

July 20, 2022

UCL Energy Institute  
14 Upper Woburn Place  
London  
WC1H 0NN  
jialeng.guo.20@ucl.ac.uk

Dear Editor of UCL Open Environment,

I am writing regarding the article entitled “BoatNet: Automated Small Boat Composition Detection using Deep Learning on Satellite Imagery” that we are submitting to be considered as a research article by UCL Open Environment.

We confirm that this work is original and has not been published elsewhere nor considered for publication elsewhere.

The article introduces a general overview of the maritime energy-emission problem and the relevance of estimating emission inventories on the shipping sector’s forgotten but critical element: the small boat fleet. The most recent development on emission inventories for the small fleet indicates that the most significant focus is on fishing vessels and vessels that have installed GPS sensors which are not the most common practice for small boats.

Compared to large commercial ships, small boats tend to be more diverse in their construction, machinery and fuel selection: (1) they tend to navigate closer to the coast for shorter periods; (2) their operational profile is more seasonal and geographically limited; and (3) registries of small boats tend to be incomplete or non-existing, especially in developing nations. Therefore, we need to cover several knowledge gaps to have a systematic and effective method to quantify the small boat fleet emission inventories. Furthermore, understanding the proportion of the emissions of this sector allows national policymakers to tailor inclusive technosocial solutions to transition the fleet following their National Determined Contributions.

In this paper, we show that our small boat detection model, BoatNet, can detect, measure and classify small boats even under low-resolution and blurry satellite images, achieving an accuracy of 93.9% with a precision of 74.0%. This is significant because the model can help accurately track the carbon inventory of the small boat globally.

Given the increase in the environment, energy, and climate change issues globally, we believe that the findings presented in our paper will appeal to the specific scientists interested in the planet’s environments and energy use, specifically how to use machine learning to establish and improve a transparent system of carbon emission and energy use caused by human activities.

We have no conflicts of interest to disclose.

If there is any further information you require from me, please do not hesitate to contact me. I look forward to hearing back from you regarding the presubmission at your earliest convenience. Please address all correspondence concerning this manuscript to me at jialeng.guo.20@ucl.ac.uk.

I appreciate your consideration of this manuscript.

Sincerely,  
Jialeng

**Jialeng Guo**  
MSc Energy Systems and Data Analytics (20/21)  
UCL Energy Institute  
The Bartlett School of Environment, Energy and Resources  
University College London

# BoatNet: Automated Small Boat Composition Detection using Deep Learning on Satellite Imagery

 **Guo Jialeng**,  **Santiago Suárez de la Fuente**,  **Tristan Smith**

UCL Energy Institute

The Bartlett School of Environment, Energy and Resources

University College London

{jialeng.guo.20, santiago.fuente.11, tristan.smith}@ucl.ac.uk

**Abstract**—Tracking and measuring national carbon footprints is one of the keys to achieving the ambitious goals set by countries. 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 has begun to look into the role played by small boat fleets in terms of Greenhouse Gases (GHG), but this either relies on high-level techno-activity assumptions or the installation of GPS 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 GHG emissions. This 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 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 GHG emissions in any given region. The data curated and produced in this study is freely available at <https://github.com/theiresearch/BoatNet>.

**Index Terms**—Object Detection, Deep Learning, Machine Learning, Transfer Learning, Small Boats Activity, Climate Change.

## I. INTRODUCTION

### A. 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]. Nowadays, 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 stability of nations and human livelihoods [5], [6]. On the other hand, renewable energy production and distribution tend to lie within the country's boundaries. During the last few years, its price 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, 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]. Still, it is important to note that renewable energy variability and investment requirements are significant challenges to grid stability and energy security.

### B. 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 since it moves around 90% of all goods around the globe [10]. However, its reliance on fossil fuels, coupled with robust economic growth, saw total CO<sub>2</sub> emissions grow from 962 Mt in 2012 to 1,056 Mt in 2018, representing more than 10% of the total global transportation emissions [11]. Further, if nothing is done in the sector, it was projected that by 2050 shipping CO<sub>2</sub> emissions could grow to 1,500 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 24m 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. 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 GPS transponders, 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. 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. Similarly, engine providers are extensive, giving a broad range of fuel consumption curves and emissions.

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, on this matter, not all small boats are powered by an internal combustion engine. They can instead be powered by sail, battery-electric or paddles.

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

Although it is possible to estimate emissions from large vessels using AIS data sent from the 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. [14]). Developed economies such as the UK tend to have a national registry of smaller vessels [15] that provides a sense of their activity level and hence can infer CO<sub>2</sub> 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 [16]. Still, it is not easy to know where they are located and infer their activities. In all, Mexico does not have a regional CO<sub>2</sub> inventory specialised in the small boat fleet; instead, they are aggregated as part of the *maritime and fluvial navigation [IA3d]* class in the national annual emission inventory developed by the Instituto Nacional de Ecología y Cambio Climático (INECC) [17] in a top-down approach based on the IPCC Guidelines [18]. 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 [19] [20] [21]. Furthermore,

in the Gulf of California, there is the largest fishing state (Sonora) in Mexico [22] and the most prominent sports fishing destination (Los Cabos, Baja) [23]. 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 [24], [25].

### C. Bringing Deep Learning to Small Ship Detection in Satellite Images

Bringing deep learning, especially convolutional neural networks, to the field of satellite image recognition is essential. First, the technological barriers due to satellite image recognition gradually decrease with time. As the quality and quantity of global satellite images improve, it is possible to obtain the same or even better detection results with reduced parameters and complexity of the convolutional neural networks. Second, satellite image recognition is not new. For example, scholars at the University of Texas at Austin have been using machine learning and remote sensing imagery to discover undersea shipwrecks successfully [26]. However, finding submarine wrecks does not require algorithms to describe the location and size of the wreck very precisely, as opposed to identifying small boats and their physical characteristics. Therefore, in this study, accurately identifying and measuring the dimensions of small ships is a current challenge and an essential objective of this work.

### D. 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.

## II. RELATED WORK

### A. 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. [27] used a top-down approach to estimate fishing sector emissions in 2011, which reached about 179 Mt CO<sub>2</sub>e, representing 17.1% of the total large fishing ship

emissions in that year [28]. However, their work only distinguished between motorised and non-motorised fishing vessels. Greer et al. [29] took a bottom-up approach to classify the fishing fleet in six different sizes, three below 24 metres long. The findings show that the small fishing boat fleet in 2016 emitted 47 Mt CO<sub>2</sub> about 22.7% of the total fishing fleet. Ferrer et al. [30] used an activity-based method using GPS, landing and fuel used data to estimate the fishing activity around Baja California Peninsula in Mexico. They found that just the small-scale fishing fleet produced 3.4 Mt of CO<sub>2</sub>e 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 CO<sub>2</sub>e, clearly placing into perspective the role of this fleet segment on national inventories [17].

Several authors have proposed using AIS to monitor the carbon emissions of the fleet [31]–[35]. Johansson et al. [36] proposed a new model (FMI-BEAM) to describe leisure boat fleet emissions in the Baltic Sea region with over 3,000 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. [37] 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, which their transponders are not normally installed on board small boat to capture their activity.

Zhang et al. [38] 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.

### B. 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 [39]. Furthermore, there are only two possibilities for this output signal: zero or one, which is very similar to a computer.

In image recognition, a 7x7 image, for example, has 49 elements or cells. If 'X' is inputted to the grid, as shown in Figure 1a, the computer will interpret it as a series of numbers (e.g. zeros and ones) as seen in Figure 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 7x7 matrix filled

with zeros and ones. After feeding the algorithm as much data as 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 [40], LeCun [41] 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 3x3 or 5x5. For instance, if the convolution kernel is 3x3, see Figure 1c, then a convolution operation will be undertaken with the 7x7 "X" matrix (Figure 1b) and the kernel (Figure 1c). The operation result is also known as a feature map (Figure 1d) [42].

The feature map reinforces the features of the convolution kernel. The 3x3 convolution kernel portrayed in Figure 1c) only has three oblique blocks of pixels that are ones. So if the original 7x7 matrix (Figure 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 (Figure 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 5x5 matrix (Figure 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 [43]. For example, after activating the sigmoid function, each element in the feature map would be between 0 and 1, as shown in Figure 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. Since 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

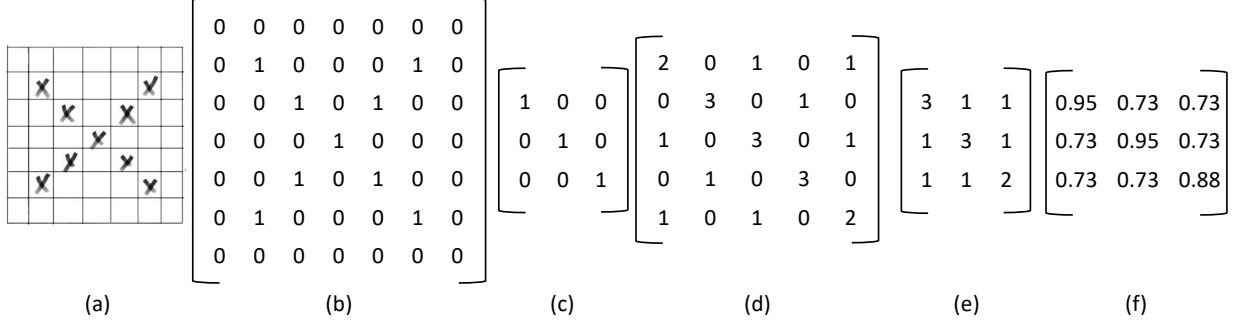


Fig. 1. From left to right: (a) Letter X in a 7x7 image; (b) Letter X in a 7x7 matrix; (c) A 3x3 convolution kernel; (d) A 5x5 feature map; (e) A 3x3 feature map after pooling; (f) A 3x3 feature map after activating with sigmoid function.

possible to determine what the image is.

### C. Convolutional Neural Networks in Image Recognition

The above literature review has demonstrated that 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 convolutional neural networks (CNNs), it may be possible to accurately identify small vessels from open satellite imagery and support understanding of this ship segment.

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 convolutional neural network (DCNN) proposed by Krizhevsky et al. [44], achieved record accuracy in image classification at the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC), making convolutional neural networks the dominant paradigm for image recognition. Next, Girshick et al. [45] introduced Region-Based Convolutional Neural Networks (R-CNN), the first convolutional neural network (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 selective search to pre-extract some candidate regions that are more likely to contain an object in order to avoid computationally costly classification and object searches, which makes it faster and significantly less computationally expensive [45], [46]. Overall, the R-CNN approach is divided into four steps:

- Generate candidate regions.
- Extract features using CNN on the candidate regions.
- Feed the extracted features into a Support Vector Machine (SVM) classifier.
- 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 [47].

Since its inception, R-CNN has undergone several developments and iterations: Fast R-CNN, Faster R-CNN and Mask R-CNN [48]–[50]. The improvement of Fast R-CNN is the design of a pooling layer structure for 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 Region Proposal Network (RPN) [49]. 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 convolutional neural network-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 RCNN model, the YOLO model can better detect smaller objects, i.e. traffic lights at a distance [51], 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 RCNN [51]. 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 [50]. Thus, the Mask R-CNN is essentially an Instance Segmentation algorithm, compared to Semantic Segmentation<sup>1</sup>. Instance Segmentation is a more

<sup>1</sup>In computer vision, image segmentation is the process of partitioning an image into multiple image segments, also known as image regions or image objects.

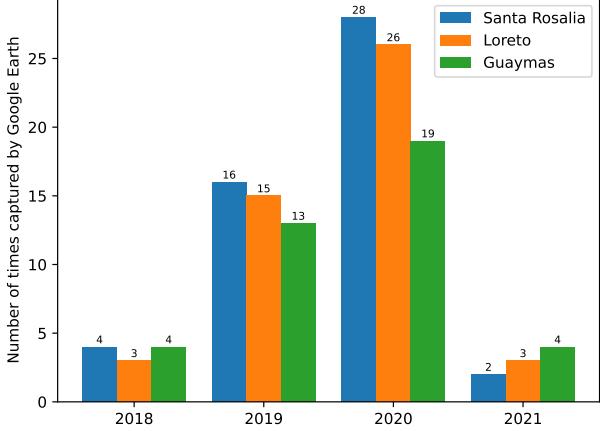


Fig. 2. Number of times the three cities (Santa Rosalia, Loreto, Guaymas) captured by Google Earth Pro from 2018 to 2021.

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 [52] researched the effect of convolutional network depth on its accuracy in large-scale image recognition settings. Their research found that even with small (3x3) convolution filters, significant accuracy is achieved by pushing the depth from 16 to 19 weight layers.

Finally, this study intends to develop the first stages of BoatNet. This image recognition model aims at detecting small 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.

### III. CONVOLUTIONAL NEURAL NETWORK CONFIGURATIONS

#### A. 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 694 high-resolution (4800 pixels x 2908 pixels) images of ships worldwide collected from Google Earth [53]. 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

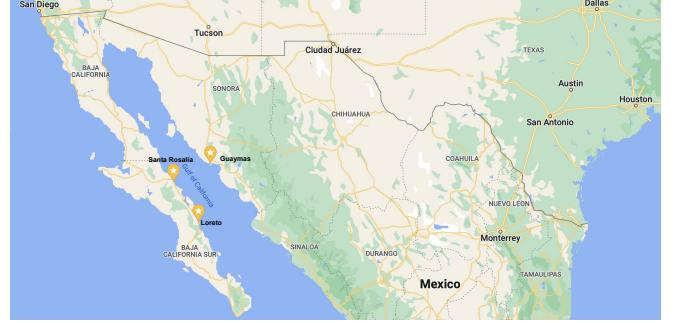


Fig. 3. The geographic locations of the three target cities - Santa Rosalia, Loreto, and Guaymas - in Google Maps.

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 Figure 3 for their geographical locations). The number of times captured by Google Earth Pro [54] is shown in Figure 2 with a database of 583 images with timestamps between 2019 and 2020 for the three Mexican coastal cities.

#### B. Preprocessing

Each satellite image was manually pre-labelled with a highly-precise label box [53]. 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 x 2908 pixels to 416 pixels x 416 pixels, with the file sizes reduced to between 10KB and 40KB [55].

#### C. 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 convolutional neural network (CNN). The detection accuracy was determined by computing the mean probability score from the Gulf's satellite images. In Sec II-C, recent literature and the development of Convolutional Neural Networks (CNNs), including the

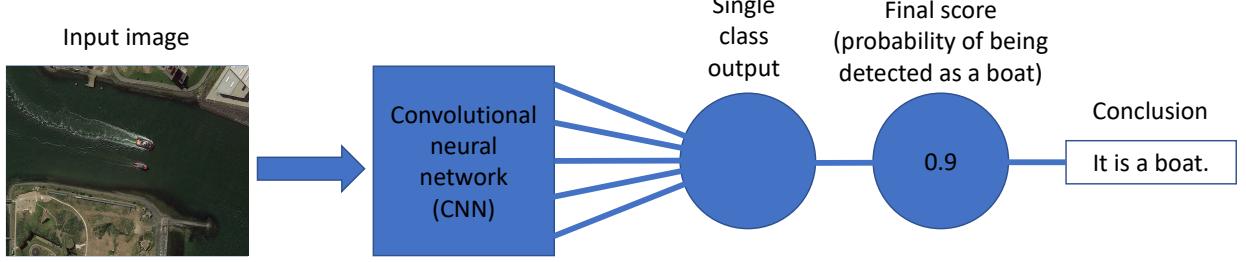


Fig. 4. **Model architecture.** Detection model architecture for obtaining a conclusion from an input satellite image of boats. Images are preprocessed and passed through a CNN. The model's output is a score,  $y \in (0, 1)$ , representing the probability of being detected as a boat.

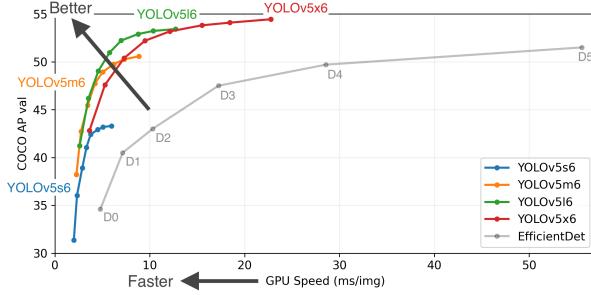


Fig. 5. Average Precision (AP) vs GPU Speed in the 6th generation of YOLOv5 model under COCO data set [56].

YOLO model, were discussed. YOLOv5 has four different categories of models, YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x [56]. They have 7.3 million, 21.4 million, 47.0 million and 87.7 million parameters, respectively. The performance charts can be seen as Figure 5 where the YOLOv5l model can achieve higher average precision with the same faster computing speed. Thus, in this study, the YOLOv5l model was chosen as the model for the training dataset.

Satellite images often contain noise such as shadows cast by water on the sea surface or clouds in the atmosphere, which make the training data inaccurate and often cause problems ensuring the model's correctness. He et al. [57] 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.

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 5x5 image kernel.

#### D. Object Measurement and Classification

Measuring the length of a ship is 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 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. Secondly, since the scale is central to the detection of the small boat fleet, the imagery scale should follow 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 metres based on the above rules. This study used a satellite image of Zurich Lake, Switzerland, on the 16th August 2018 as the standard image for defining the scale (Figure 6). 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.

To validate the length method a length measurement comparison was done between BoatNet and Google Earth Pro's length tool. The vessel in Figure 6 had a YOLO length of 0.43 that, after the scale conversion, represented 55.17 metres. With an eye altitude of 200 metres, 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% to 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

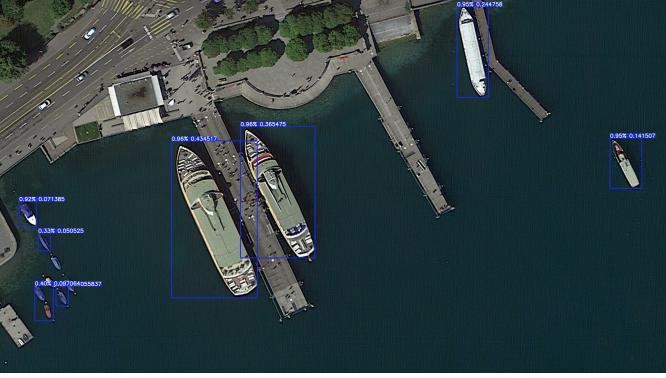


Fig. 6. An image from Google Earth Pro for Zurich Lake on 16 Aug 2018 when eye alt is 200 metres.

image must be the same, so measurements are standardised. For this purpose, all datasets that detect small boats will maintain a resolution of 3840x2160 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 200m 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 a recreational boat, while any other mix of colour would be designated 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.

#### IV. RESULTS

##### A. Train Custom Data: Weights, Biases Logging, Local Logging

As shown in Figure 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)<sup>2</sup> is between 50% 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

<sup>2</sup>IOU is an evaluation metric used to measure the accuracy of an object detector on a particular dataset.

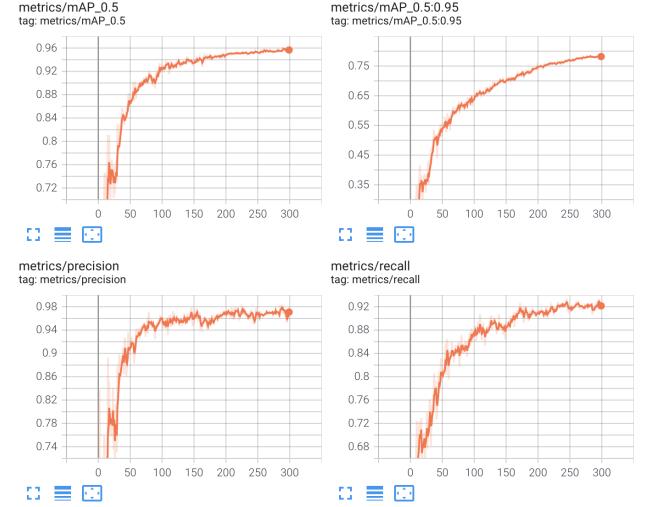


Fig. 7. Average model precision when IOU is larger than 0.5; Average model precision when IOU is between 0.5 and 0.95; The model precision; The model recall rate.

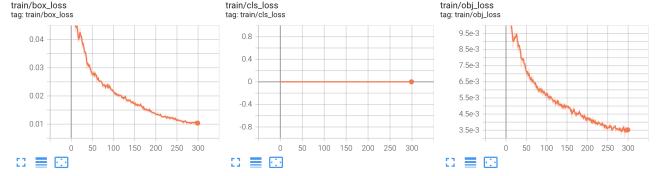


Fig. 8. The box loss rate of the model; The class loss rate of the model; The object loss rate of the model.

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, a reduction of the images' resolution of about 70 times, resulting in images of 416 pixels x 416 pixels. The training could otherwise take two weeks if the images used had a resolution of 4800 pixels x 2908 pixels.

Similarly, as shown in Figure 8, the loss rate of the box can eventually reach 1% as the number of training sessions increases. Since 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 0. 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. Since only one class was set, some were also considered a 100% probability of being boats.

##### B. 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

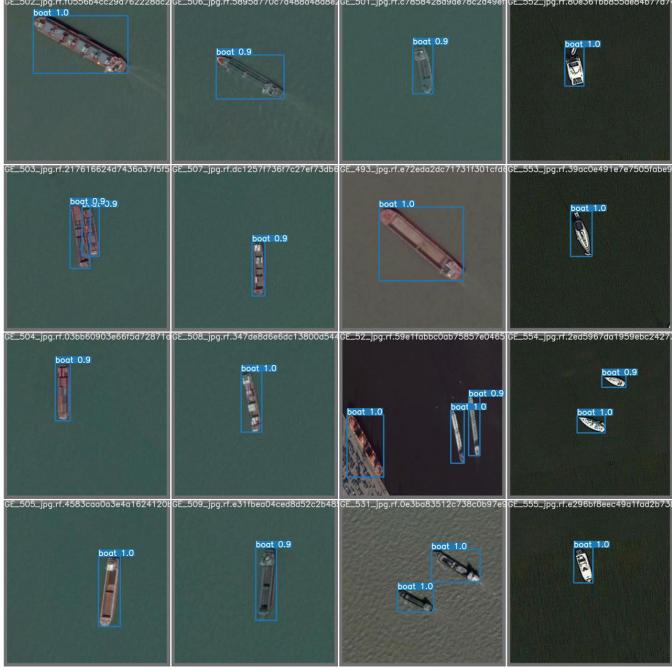


Fig. 9. Test result of a trained model for detecting ships.

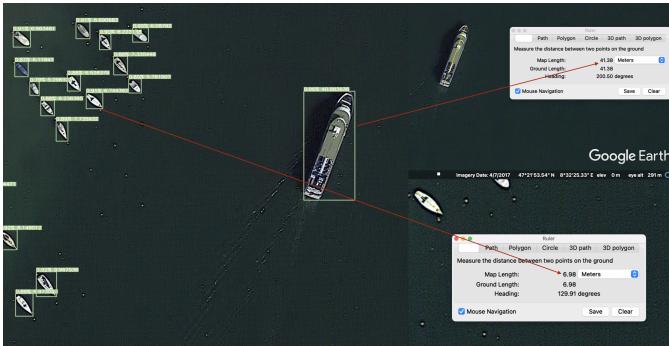


Fig. 10. Compare the length of boats using Google Earth ruler and computer vision algorithm. This example shows the image from Google Earth Pro for Zurich Lake on 16th August 2018 when the eye alt was 200 metres.

Earth Pro measuring tools, Figure 10 shows that the small boat detected measured 6.98 metres using Google Earth Pro, while BoatNet estimated 6.74 metres. The error between them is 3.4%. Google Earth Pro measured the larger ship at 41.38 metres and BoatNet at 40.98 metres. The error between them is less than 1.0%.

As explained in Sec III-C, 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 Sec III-C, 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 5x5 sharpening kernel allowed for a higher recognition rate. However, the following situations still occur:

- 1) Figure 11: When the detailed representation of the image



Fig. 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.

is indigent, i.e., 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 Figures 11, 12, 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 Figure 14 and Figure 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

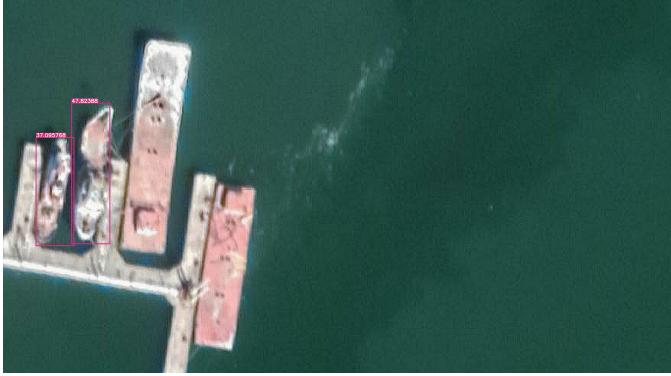


Fig. 12. When the cargo ship is full of cargo, the ship looks like a rectangular jetty from above and loses the normal shape of a ship. The image is from Guaymas, January 2020.



Fig. 13. Models may have difficulty detecting small boats moored on the beach. The image is from Santa Rosalia, February 2021.

3.00 times more than in Loreto.

3) Guaymas had a larger number of large ships as well.

There were more than ten 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.

In fact, according to statistic [58] 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 since, by probability, there would be more economical and leisure activities around larger coastal cities.

#### C. Entertainment and Fishing Boats in the Gulf of California

According to the statement in Sec III-C, 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 I, 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 since the algorithm does not consider, for example, other colours as part of the characteristics of leisure boats. Furthermore, although there are many uncertainties in

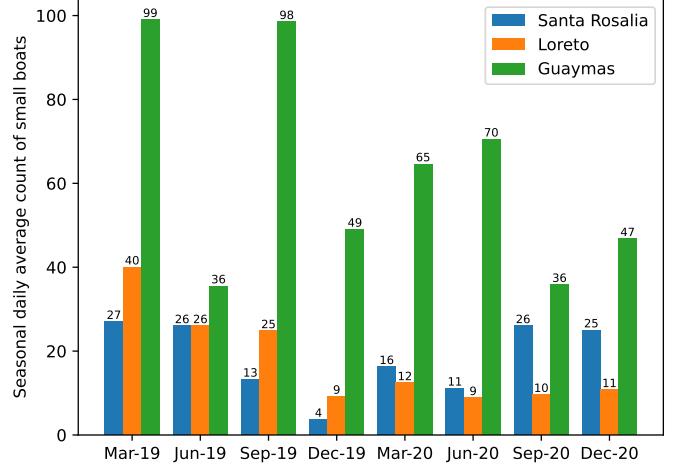


Fig. 14. The seasonal daily average count of small boats in Santa Rosalia, Loreto, and Guaymas between 2019 and 2020.

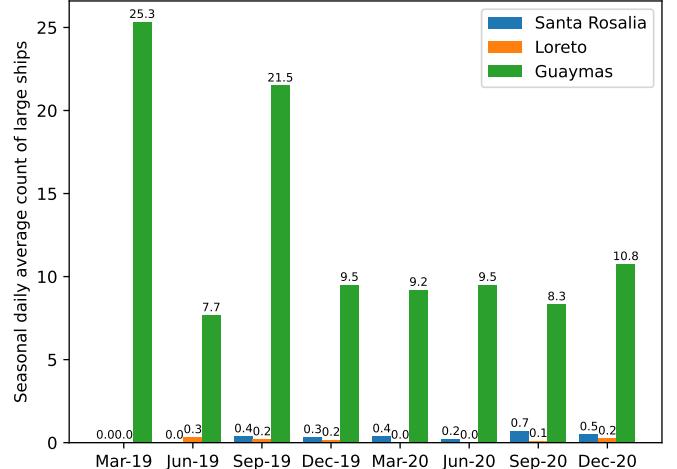


Fig. 15. The seasonal daily average count of large ships in Santa Rosalia, Loreto, and Guaymas between 2019 and 2020.

	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

TABLE I

**DETECTION EXAMPLE. SMALL BOATS, LARGE SHIPS, AND SMALL WHITE BOATS IN GUAYMAS IN 2020.**

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.

## V. 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% and encouraging results considering the high variability of the input images.

Even with the model 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 and ultimately their carbon inventory by adding the emissions produced by the small boat fleet. Further, 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 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  $\times$  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 [59]. 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  $\times$  2908 pixels. Thus, BoatNet was set to learn from resized small images measuring 416 pixels  $\times$  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.

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 [30], typical machinery, fuel characteristics, and emission factors unique to this maritime segment [17].

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.

## VI. ACKNOWLEDGEMENT

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.

## VII. DATA AVAILABILITY STATEMENT

All the data and source codes used in this paper can be viewed at: <https://github.com/theiresearch/BoatNet>, which is under Apache License 2.0.

## VIII. DECLARATIONS AND CONFLICT OF INTERESTS STATEMENT

The authors certify that they have no affiliations with or involvement in any organisation or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

## REFERENCES

- [1] H.-O. Pörtner, D. Roberts, M. Tignor, E. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, and B. Rama, "Climate change 2022: Impacts, adaptation, and vulnerability." *Cambridge University Press*, 2022.
- [2] A. K. Raturi, "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," 2021. [Online]. Available: <https://iea.blob.core.windows.net/assets/4ed140c1-c3f3-4fd9-acae-789a4e14a23c/WorldEnergyOutlook2021.pdf>
- [5] S. C. Bhattacharyya, "Fossil-fuel dependence and vulnerability of electricity generation: Case of selected european countries," *Energy Policy*, vol. 37, no. 6, pp. 2411–2420, 2009.
- [6] DBIS, "Russia-Ukraine and UK energy: factsheet," *Department for Business, Energy & Industrial Strategy, UK Government*, 2022, <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, <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] T. Economics, “UK natural gas - 2022 data - 2020-2021 historical - 2023 forecast - price - quote,” *Trading Economics*, 2022, <https://tradingeconomics.com/commodity/uk-natural-gas>.
- [10] T. R. Walker, O. Adebambo, M. C. D. A. Feijoo, E. Elhaimer, T. Hossain, S. J. Edwards, C. E. Morrison, J. Romo, N. Sharma, S. Taylor *et al.*, “Environmental effects of marine transportation,” in *World Seas: an environmental evaluation*. Elsevier, 2019, pp. 505–530.
- [11] S. Faber, S. Hanayama, S. Zhang, P. Pereda, B. Comer, E. Hauerhof, and H. Kosaka, “Fourth IMO GHG study 2020,” 2020.
- [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.
- [13] UK Government, “Operational standards for small vessels,” <https://www.gov.uk/operational-standards-for-small-vessels>, 2021, accessed: 2021-08-30.
- [14] S. SENER, “Balance nacional de energía 2017,” *España: Dirección general de planeación e información*, 2018.
- [15] UK Ship Register, “UK small ship registration,” <https://www.ukshipregister.co.uk/registration/leisure/>, 2021, accessed: 2021-08-30.
- [16] Mexico National Aquaculture and Fisheries Commission, “Registered vessels,” <https://www.conapesca.gob.mx/wb/cona/embarcaciones-registradas>, 2021, accessed: 2021-08-30.
- [17] INECC, “Inventario nacional de emisiones de gases y compuestos de efecto invernadero (INEGYCEI),” 2020.
- [18] I. Change, “2006 ipcc guidelines for national greenhouse gas inventories,” *Institute for Global Environmental Strategies, Hayama, Kanagawa, Japan*, 2006.
- [19] S. E. Lluch-Cota, E. A. Aragón-Noriega, F. Arreguín-Sánchez, D. Auriolos-Gamboa, J. Jesús Bautista-Romero, R. C. Brusca, R. Cervantes-Duarte, R. Cortés-Altamirano, P. Del-Monte-Luna, A. Esquivel-Herrera, G. Fernández, M. E. Hendrickx, S. Hernández-Vázquez, H. Herrera-Cervantes, M. Kahru, M. Lavin, D. Lluch-Belda, D. B. Lluch-Cota, J. López-Martínez, S. G. Marinone, M. O. Nevárez-Martínez, S. Ortega-García, E. Palacios-Castro, A. Parés-Sierra, G. Ponce-Díaz, M. Ramírez-Rodríguez, C. A. Salinas-Zavala, R. A. Schwartzlose, and A. P. Sierra-Beltrán, “The gulf of california: Review of ecosystem status and sustainability challenges,” *Progress in Oceanography*, vol. 73, no. 1, pp. 1–26, 2007. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0079661107000134>
- [20] A. Munguia-Vega, A. L. Green, A. N. Suarez-Castillo, M. J. Espinosa-Romero, O. Aburto-Oropeza, A. M. Cisneros-Montemayor, G. Cruz-Piñón, G. Danemann, A. Giron-Nava, O. Gonzalez-Cuellar *et al.*, “Ecological guidelines for designing networks of marine reserves in the unique biophysical environment of the gulf of california,” *Reviews in Fish Biology and Fisheries*, vol. 28, no. 4, pp. 749–776, 2018.
- [21] S. Marinone, “Seasonal surface connectivity in the gulf of california,” *Estuarine, Coastal and Shelf Science*, vol. 100, pp. 133–141, 2012, recent advances in biogeochemistry of coastal seas and continental shelves. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0272771412000169>
- [22] L. Meltzer and J. O. Chang, “Export market influence on the development of the pacific shrimp fishery of sonora, mexico,” *Ocean & Coastal Management*, vol. 49, no. 3, pp. 222–235, 2006. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0964569106000329>
- [23] V. Hernández-Trejo, G. P.-D. Germán, D. Lluch-Belda, and L. Beltrán-Morales, “Economic benefits of sport fishing in e relative los cabos, mexico: is the abundance a determinant,” *Journal of Sustainable Tourism*, vol. 161, pp. 165–176, 2012.
- [24] CNANP, “Atlas nacional de las areas naturales protegidas de mexico,” <http://sig.conanp.gob.mx/website/interactivo/anps/>, accessed: 2022-04-05.
- [25] SMARN, “Islas y Áreas protegidas del golfo de california,” <https://www.gob.mx/semarnat/articulos/islas-y-areas-protegidas-del-golfo-de-california-269050>, accessed: 2022-04-05.
- [26] L. Character, A. Ortiz Jr, T. Beach, and S. Luzzadher-Beach, “Archaeologic machine learning for shipwreck detection using lidar and sonar,” *Remote Sensing*, vol. 13, no. 9, p. 1759, 2021.
- [27] R. W. Parker, J. L. Blanchard, C. Gardner, B. S. Green, K. Hartmann, P. H. Tyedmers, and R. A. Watson, “Fuel use and greenhouse gas emissions of world fisheries,” *Nature Climate Change*, vol. 8, no. 4, pp. 333–337, 2018.
- [28] T. Smith, J. Jalkanen, B. Anderson, J. Corbett, J. Faber, S. Hanayama, E. O’keeffe, S. Parker, L. Johansson, L. Aldous *et al.*, “Third IMO GHG Study 2014,” 2015.
- [29] K. Greer, D. Zeller, J. Woroniak, A. Coulter, M. Winchester, M. D. Palomares, and D. Pauly, “Global trends in carbon dioxide (co2) emissions from fuel combustion in marine fisheries from 1950 to 2016,” *Marine Policy*, vol. 107, p. 103382, 2019. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0308597X1730893X>
- [30] E. M. Ferrer, O. Aburto-Oropeza, V. Jimenez-Esquivel, J. J. Cota-Nieto, I. Mascarenas-Osorio, and C. Lopez-Sagastegui, “Mexican small-scale fisheries reveal new insights into low-carbon seafood and “climate-friendly” fisheries management,” *Fisheries*, vol. 46, no. 6, pp. 277–287, 2021.
- [31] M. Traut<sup>a</sup>, A. Bowsa, C. Walsha, and R. Wooda, “Monitoring shipping emissions via ais data? certainly,” *Low Carbon Shipping Conference*, 2013.
- [32] L. Johansson, J.-P. Jalkanen, and J. Kukkonen, “A comprehensive modelling approach for the assessment of global shipping emissions,” in *International Technical Meeting on Air Pollution Modelling and its Application*. Springer, 2016, pp. 367–373.
- [33] A. Mabunda, A. Astito, and S. Hamdoune, “Estimating carbon dioxide and particulate matter emissions from ships using automatic identification system data,” *International Journal of Computer Applications*, vol. 88, pp. 27–31, 2014.
- [34] T. Hensel, C. Ugé, and C. Jahn, “Green shipping: using ais data to assess global emissions,” *Sustainability Management Forum — NachhaltigkeitsManagementForum*, vol. 28, pp. 39–47, 2020.
- [35] W. Han, W. Yang, and S. Gao, “Real-time identification and tracking of emission from vessels based on automatic identification system data,” *2016 13th International Conference on Service Systems and Service Management (ICSSSM)*, pp. 1–6, 2016.
- [36] L. Johansson, J.-P. Jalkanen, E. Fridell, I. Maljutenko, E. Ytreberg, M. Eriksson, E. Roth, and V. Fischer, “Modeling of leisure craft emissions,” in *International Technical Meeting on Air Pollution Modelling and its Application*. Springer, 2018, pp. 205–210.
- [37] C. Ugé, T. Scheidweiler, and C. Jahn, “Estimation of worldwide ship emissions using ais signals,” *2020 European Navigation Conference (ENC)*, pp. 1–10, 2020.
- [38] Y. Zhang, J. Fung, J. Chan, and A. Lau, “The significance of incorporating unidentified vessels into ais-based ship emission inventory,” *Atmospheric Environment*, vol. 203, pp. 102–113, 2019.
- [39] W. S. McCulloch and W. Pitts, “A logical calculus of the ideas immanent in nervous activity,” *The bulletin of mathematical biophysics*, vol. 5, no. 4, pp. 115–133, 1943.
- [40] K. Fukushima and S. Miyake, “Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition,” in *Competition and cooperation in neural nets*. Springer, 1982, pp. 267–285.
- [41] Y. LeCun, Y. Bengio *et al.*, “Convolutional networks for images, speech, and time series,” *The handbook of brain theory and neural networks*, vol. 3361, no. 10, p. 1995, 1995.
- [42] J. Bouvrie, “Notes on convolutional neural networks,” 2006.
- [43] G. Lin and W. Shen, “Research on convolutional neural network based on improved relu piecewise activation function,” *Procedia computer science*, vol. 131, pp. 977–984, 2018.
- [44] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” *Advances in neural information processing systems*, vol. 25, pp. 1097–1105, 2012.
- [45] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 580–587.
- [46] J. R. Uijlings, K. E. Van De Sande, T. Gevers, and A. W. Smeulders, “Selective search for object recognition,” *International journal of computer vision*, vol. 104, no. 2, pp. 154–171, 2013.
- [47] J. Huang, V. Rathod, C. Sun, M. Zhu, A. Korattikara, A. Fathi, I. Fischer, Z. Wojna, Y. Song, S. Guadarrama *et al.*, “Speed/accuracy trade-offs for modern convolutional object detectors,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7310–7311.
- [48] R. Girshick, “Fast r-cnn,” in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1440–1448.
- [49] S. Ren, K. He, R. Girshick, and J. Sun, “Faster r-cnn: Towards real-time object detection with region proposal networks,” *arXiv preprint arXiv:1506.01497*, 2015.

- [50] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2961–2969.
- [51] P. Dwivedi, "Yolov5 compared to Faster RCNN. Who wins?" <https://towardsdatascience.com/yolov5-compared-to-faster-rcnn-who-wins-a771cd6c9fb4>, 2020, accessed: 2021-08-30.
- [52] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *CoRR*, vol. abs/1409.1556, 2015.
- [53] T. Lutherborrow, A. Agoub, and M. Kada, "Ship detection in satellite imagery via convolutional neural networks," 2018.
- [54] R. J. Lisle, "Google earth: a new geological resource," *Geology today*, vol. 22, no. 1, pp. 29–32, 2006.
- [55] J. Nelson, "You might be resizing your images incorrectly," Jan 2020. [Online]. Available: <https://blog.roboflow.com/you-might-be-resizing-your-images-incorrectly/>
- [56] G. Jocher, A. Chaurasia, A. Stoken, J. Borovec, NanoCode012, Y. Kwon, TaoXie, J. Fang, imyhxy, K. Michael, Lorna, A. V, D. Montes, J. Nadar, Laughing, tkianai, yxNONG, P. Skalski, Z. Wang, A. Hogan, C. Fati, L. Mammana, AlexWang1900, D. Patel, D. Yiwei, F. You, J. Hajek, L. Diaconu, and M. T. Minh, "ultralytics/yolov5: v6.1 - TensorRT, TensorFlow Edge TPU and OpenVINO Export and Inference," Feb. 2022. [Online]. Available: <https://doi.org/10.5281/zenodo.6222936>
- [57] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1956–1963, 2009.
- [58] INEGI, "Population and housing census 2020 in mexico," <https://www.inegi.org.mx/app/cpv/2020/resultadosrapidos/>, accessed: 2022-06-11.
- [59] M. Tan and Q. Le, "Efficientnetv2: Smaller models and faster training," in *International Conference on Machine Learning*. PMLR, 2021, pp. 10 096–10 106.