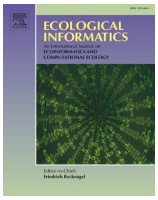
Ecological Informatics 78 (2023) 102285

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf



Intelligent airborne monitoring of irregularly shaped man-made marine objects using statistical Machine Learning techniques

Kaya Kuru [a,*], Stuart Clough [b,1], Darren Ansell [a], John McCarthy [c,2], Stephanie McGovern [c,2]

[a] School of Engineering and Computing, University of Central Lancashire, Fylde Rd, Preston, Lancashire, PR12HE, UK
[b] APEM Inc., 2603 NW 13th Street, 402, Gainesville, FL 32609-2835, USA
[c] APEM Ltd., The Embankment Business Park, Stockport SK4 3GN, UK

ARTICLE INFO

Keywords:
Marine ecosystems
Marine man-made objects
Statistical ML
HSV colour space
Object detection
ROC curve
Aerial surveys

ABSTRACT

The marine economy has historically been highly diversified and prolific due to the fact that the Earth's oceans comprise two-thirds of its total surface area. As technology advances, leading enterprises and ecological organisations are building and mobilising new devices supported by cutting-edge marine mechatronics solutions to explore and harness this challenging environment. Automated tracking of these types of industries and the marine life around them can help us figure out what's causing the current changes in species numbers, predict what could happen in the future, and create the right policies to help reduce the environmental impact and make the planet more sustainable. The objective of this study is to create a new platform for the automated detection of irregularly shaped man-made marine objects (ISMMMOs) in large datasets derived from marine aerial survey imagery. In this context, a novel nonparametric methodology, which harbours several hybrid statistical Machine Learning (ML) methods, was developed to automatically segment ISMMMOs on the sea surface in large surveys. This methodology was validated on a wide range of marine domains, providing robust empirical proof of concept. This approach enables the detection of ISMMMOs automatically, without any prior training, with accuracy (*ACC*), Matthews correlation coefficient (*MCC*), negative predictive value (*NPV*), positive predictive value (*PPV*), specificity (*Sp*) and sensitivity(*Se*) over 0.95. The outlined methodology can be utilised for a variety of purposes, but it's especially useful for researchers and policymakers who want to keep an eye on how the maritime industry is deploying and make sure the right policies are in place to meet regulatory and legal requirements to promote maritime tech innovation and shape what the future looks like for the marine ecosystem. For the first time in the literature, a method, the so-called ISMMMOD, has been developed to automate the detection of all types of ISMMMOs by statistical ML techniques that require no prior training, which will pioneer the monitoring of human footprint in the marine ecosystem.

1. Introduction

The process of marine spatial planning is highly contentious due to the presence of multiple stakeholders, often with conflicting objectives and values (Elrick-Barr et al., 2022). The maritime economy has historically been highly diversified and prolific due to the fact that the Earth's oceans comprise two-thirds of its total surface area. As technology advances, leading enterprises and ecological organisations are building and mobilising new devices supported by cutting-edge marine mechatronics solutions (Shi et al., 2017) within the framework of Automation of Everything (AoE) (Kuru and Yetgin, 2019) to explore and harness this challenging environment. More specifically, robotic vehicles, autonomous vehicles, and surface vessels have been deployed for the offshore industries and deep-sea archaeology, ocean engineering projects, rescue operations and environmental measurements for the last several decades. For instance, the Argo program, an international collaboration, has deployed approximately 3900 instruments in the world's oceans to facilitate the collection of data for climatological and oceanographic studies. (Riser et al., 2016). Besides, artificial structures such as gas, oil and deep seabed mining platforms, offshore renewable energy harvesting technologies such as oil and gas installations, wind farms and wave energy converters, fish farms, ships, boats and yachts for

* Corresponding author.
E-mail address: kkuru@uclan.ac.uk (K. Kuru).
[1] https://apem-inc.com
[2] https://www.apemltd.co.uk

https://doi.org/10.1016/j.ecoinf.2023.102285
Received 26 June 2023; Received in revised form 28 August 2023; Accepted 31 August 2023
Available online 9 September 2023

transportation, autonomous marine vehicles from unmanned ships to smaller vessels are becoming inevitable components of the offshore environment. For instance, in recent years, the offshore wind industry has seen a remarkable expansion, with an annual rate of growth of 25%, for constructing offshore energy islands to meet the reduction of gas emission targets (Zhang et al., 2021). These may conflict with nature conservation objectives, such as habitat loss or species endangerment. In other words, this rapidly expanding industry, which allows for extensive, ongoing human influence in the marine domain, has the potential to have a significant impact on the marine environment, particularly on the marine floor, on turtles, fish, and birds. For instance, the population of monitored seabirds, which account for about 19% of the all seabird populations, declined by almost 70% from 1950 to 2010 (Paleczny et al., 2015), resulting in a net loss of almost 3 billion (%29) birds since 1970 (Rosenberg et al., 2019). The decline of bird populations serves as a stark reminder of the need for immediate action to mitigate threats to the eventual decline of avifauna and the resulting degradation of ecosystem health, functionality, and services (Rosenberg et al., 2019). Intervention into nature is a natural consequence of human activities, however, when managed effectively, these interventions can be beneficial not only to the environment, but also to the ongoing development of civilisation (Sànchez-Marrè et al., 2004). To better understand the planet and to ensure effective conservation planning, it is essential to have a comprehensive understanding of the species, habitats, and sites that require protection. Unfortunately, for the majority of species and regions, comprehensive quantitative knowledge is not yet available (Bibby et al., 1998). One of the key objectives of the development and utilisation of ecological models and applications is to influence the ecological policy practices, outputs and results in a beneficial manner (McIntosh et al., 2011). There is an urgent need to monitor the environmental upheavals, impacts and possible trends with environmental time series analysis, models and tools as the footprint of human activities increases with the rapid development of the industry. In this manner, modelling, automated detection, location and real-time monitoring of industrial sites and ecosystems around them can help uncover the current and potential future effects on nature. Furthermore, the insights observed and models developed based on these insights may help researchers and policymakers to monitor this diverse ecosystem along with the associated maritime industries and thereby help to determine the legal and regulatory requirements for reducing the ecological footprint concerning immediate foreseeable environmental problems.

There are numerous studies in the literature to detect underwater man-made objects (MMOs) within a limited region of interest (RoI) using underwater imagery, robots or sonography. For instance, Abu et al. (Abu and Diamant, 2022) proposes a contour-based features analysis method to discern underwater MMOs from natural environment considering that contours of MMOs' are supposed to be smoother than natural objects. There are a limited number of studies in the literature to detect specific types of surface marine MMOs using supervised Machine Learning (ML) and Deep Learning (DL) approaches that require prior training in the marine ecosystem. For instance, Han et al. (Han et al., 2022) proposed a DL technique titled LCSE-ResNet to detect, classify and locate vessels and oil platforms based on remote optical imagery, by which all other MMOs are excluded. There are no studies in the literature that investigates the detection of all types of surface marine MMOs, which makes this research the first study of its kind. Most irregularly shaped man-made marine objects (ISMMMOs) are made of materials such as metal, treated wood, fibreglass, PVC plastic, glass, or concrete and they have different types of irregular shapes and colours. Hence, it is infeasible to apply: i) a template matching technique based on a specific object to input as a template, and ii) a supervised ML approach based on a prior knowledge/similar datasets to train similar objects and then detect these objects automatically. Moreover, the current clustering algorithms used to group visual datasets are not capable of accomplishing this task with a high degree of precision (Kuru et al., 2013; Kuru and Khan, 2018), particularly for objects with indefinite shapes. Therefore, a new method is needed to realise this objective. On one hand, automatic detection of ISMMMOs is not easy based on two main reasons which pose a considerable challenge: i) the rapidly changing background depending on the camera, water turbidity, weather, wind, wave speed and period, sun glint and density of clouds, and ii) various non-definitive morphologies of MMOs. On the other hand, the characteristics of ISMMMOs differ from the natural environment and other natural objects within this ecosystem regarding the composition, features of the surface, saturation of light and colourfulness relative to the brightness to which an area radiate a varying amount of light.

Many studies aim to detect marine natural objects in sea areas using stationary land-based fixed cameras, in particular, sea animals: detection of animals in deep-sea video (Mehrnejad et al., 2013), detection of sharks using multispectral imaging (Lopez et al., 2014), and detection of killer whales using infrared spectrum (Graber, 2011). Furthermore, aerial surveys from a helicopter or small aircraft have been conducted for many years to detect, locate and monitor specific marine animals using human-based visual observations. Although there are several studies for the detection of MMOs such as ship (Saur et al., 2011), specific objects (e.g., boats, humans) on the ocean surface using infrared cameras (Leira et al., 2015). To the best of our understanding, there is no study that aims to detect all kinds of ISMMMOs automatically with unsupervised approaches using standard advanced cameras and aerial surveys, in particular, from the perspective of ecology. Aerial surveys provide a cost-effective way to collect environmental information over large areas in a short amount of time; however, they may not provide reliable data if not conducted correctly (Davis et al., 2022). Long-term data using standardised and well-structured approaches are the best way to measure change in ecology; unfortunately, this data is not available for most biogeographical regions (Clements and Robinson, 2022) due to the cost of data processing with intensive human intervention. In this sense, this study mainly aims to fill this gap in the scientific literature either by processing the collected data in an automated way, with no human intervention, to separate several hundreds of ISMMMOs from large surveys, or by processing images as they are streamed from the airborne camera systems to monitor ISMMMOs with their geospatial locations immediately with a novel approach using statistical ML techniques and HSV colour mode.

In a conventional marine survey program, there may be a large number of images, e.g., around a million, collected over a period of one year to be analysed for a particular site, and it is labour-intensive to categorise the data into two groups: positive images containing man-made material and negative images without man-made material. In fact, many of the surveys that APEM Ltd.[3] has acquired indicate that >95% of the aerial survey images do not accommodate any targeted object (Kuru et al., 2023). According to some research on visual perception, humans perceive only a small portion of an environment or scene in detail under typical viewing conditions (Noe et al., 2000), which may result in discarding other details that should be taken into account. Although the elements that influence how a scene is perceived are not yet known, it appears that focus is a significant factor (Noe et al., 2000). Within this context, detection of ISMMMOs in large-scale images within very large surveys is a non-trivial task and labour-intensive. Therefore, the utilisation of automated intelligent computer systems to automate this work would be highly advantageous in order to facilitate the development of efficient environmental models with real-world inputs.

To the best of our knowledge, this study, for the first time, explicitly investigates the automatic detection of offshore ISMMMOs to assist researchers, environmentalists, and policymakers in monitoring and managing the various applications of the maritime industry and to provide guidance on the necessary regulations and legal requirements.

[3] APEM Ltd. is an environmental company and proposes novel solutions for environmental problems (https://www.apemltd.co.uk).

In order to illustrate the novelty of this research, specific contributions are listed below.

1. A novel methodology, the so-called ISMMMOD, that detects and splits ISMMMOs automatically in large-scale images in typical very large marine surveys is built.
2. The ISMMMOD is developed using the HSV colour space and statistical analysis of histograms of the channels in this space based on the ROC (receiver operating characteristic) curve analysis. The techniques in the methodology differ man-made-built structures from natural maritime habitats (i.e., waves, sea animals, birds, seawater) in various aspects, in particular, composition, features of the surface and saturation of light.

The rest of this document is structured as follows: The methodology is revealed in Section 2. The datasets on which the methodology is built and tested are explored in Section 3. A summary of the findings is provided in Section 4. Discussions are outlined in Section 5. Section 6 draws a conclusion as well as future potential works. Finally, the limitations of the study are disclosed in Section 7.

2. Methodology

All types of mobile and stationary human activities – human footprint – are required to be monitored on a regular basis and most of these activities involve the use of non-uniform, human built structures in multiple shapes during the exploration and exploitation of these tough marine ecosystem. Detecting these non-uniform structures on their highly dynamic background entails the development of a new technique that is not based on pre-trained uniform object classifiers, but based on the features independent from their shapes. In this respect, we would like to reveal the features that are different from the maritime ecosystem by which a new detection method is aimed to be developed. Built structures differ from the natural maritime habitats and their creatures in various aspects, in particular, composition, features of the surface and saturation of light. The saturation level of ISMMMOs significantly varies from their surroundings (i.e., waves, sea animals, birds, seawater). More explicitly, the saturation level of these ISMMMOs is more intense than

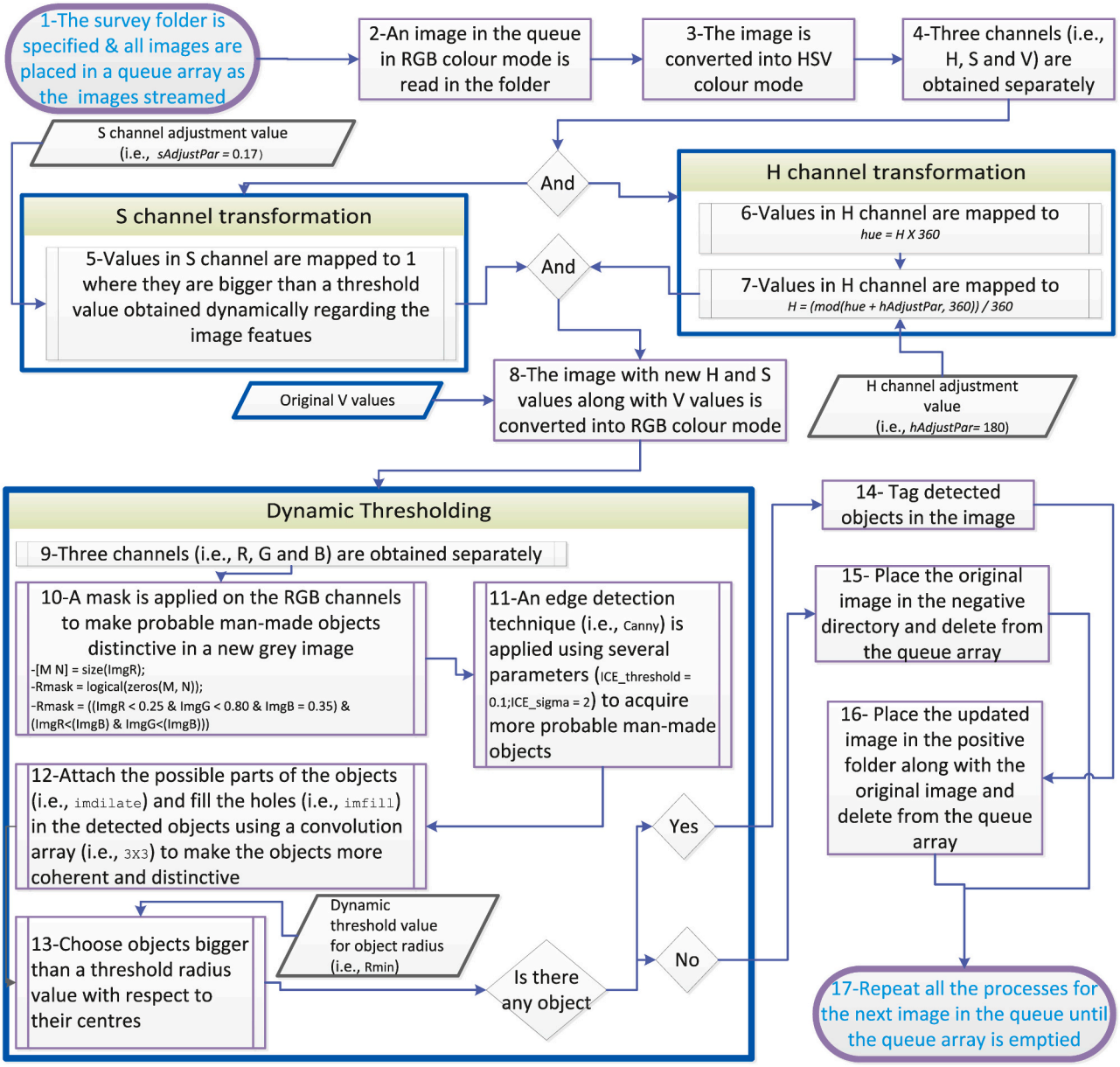


Fig. 1. Overall methodology. The images as they stream are automatically placed in the queue array to be processed in an automated way.

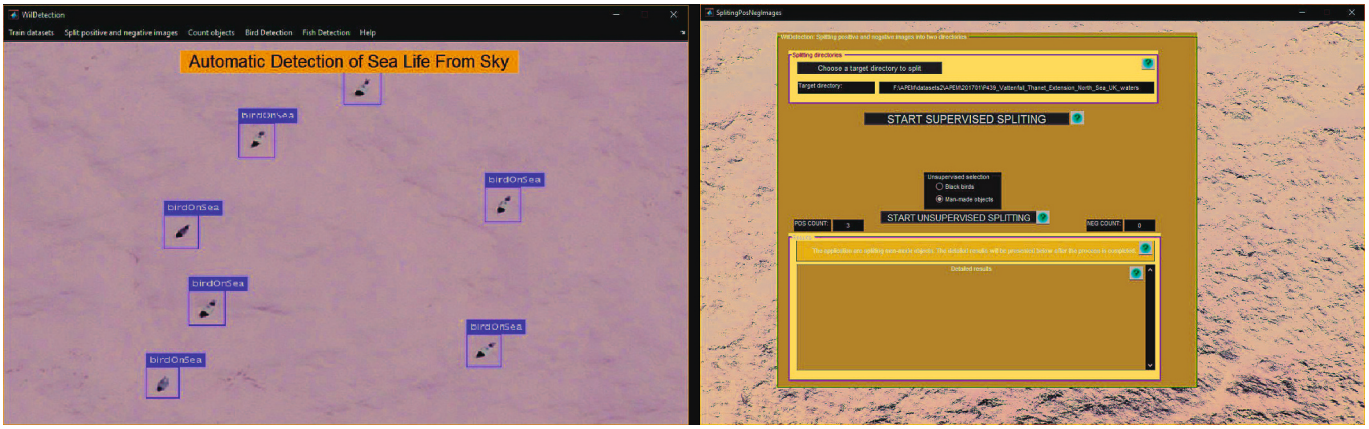


Fig. 2. User interfaces. Left: the main platform developed for multi purpose environmental applications. Right: man-made object detection and splitting interface that can be opened from the main platform.

that of the natural marine life, and in this study, more saturated sections in images considering this distinguishing feature are made distinct to detect these artificial objects. More specifically, the methodology is based on the HSV colour space (elaborated in Section 2.1) and statistical analysis of histograms of the channels in this space (elaborated in Section 2.2). The essential phases of the technique and its automated execution are depicted in Fig. 1. The dynamic thresholding in the implementation of the methodology is presented in Algorithm 1 and the automated implementation of the overall methodology is presented in Algorithms 2 and 3 which are placed in Appendix A. The execution of the methodology is exemplified for the images in Figs. 11a, 12a, 13a, 14a, 15a. The techniques in this research was built with Matlab R2020a. The interface is shown in Fig. 2. Generally speaking, in the proposed approach, the aerial image in RGB colour space is converted into HSV colour space and then the converted image is split into three components (i.e., channels), namely *H*, *S* and *V* that are designed to approximate the human vision. The result is a 3D matrix with elements of Hue, Saturation and Value. In the next step, the histograms of these components are computed as illustrated in the first rows of Figs. 11b, 12b, 13b, 14b, 15b. Then, the dynamically calculated threshold value is applied to *S* component along with shifting the *H* channel. At the end, the morphological operations, namely, masking, filter, and smoothing are carried out to extract the required area by suppressing the irrelevant parts mentioned in Section 2.3 such as glinting regions. The statistical terms used throughout the paper are explained regarding the scope of this paper in Table 3 for the readers who are not familiar with these commonly used terms for the statistical analysis related to the confusion matrix.

Before revealing the methodology on data samples in detail, we would like to explore the basic concepts of the HSV colour space in Section 2.1 to shed light on the developed techniques. Then, the phases of the methodology (Fig. 1) are disclosed on the sample images acquired from various image surveys. The dynamic thresholding phase for *S* channel is explained in Section 2.2. The phases of the masking and dilation are presented in Section 2.3.

2.1. HSV and its applications in the methodology

The main colour models are RGB, HSV, CIELAB, CMYK, and XYZ. The colour models different from the RGB are employed to realise different objectives because several fundamental issues can not be addressed using the additive RGB colour mode for image segmentation such that it is not viable to get the luminance of the image regarding human perception. For instance, the CIELAB colour space that is close to the human visual perception is applied to H&E stained microscopical images to correct the Kohler illumination problem in microscopical images (Kuru, 2014). Likewise, HSV provides a close representation of human visual perception of colour in cylindrical-coordinate representations as illustrated in Fig. 4 whereas the RGB colour mode represents the colours processing in the human biological visual system (Loesdau et al., 2014). HSV stands for i) the hue that corresponds to the angle (from the red at 0°, to the green at 120° and the blue at 240°, and then back to red again at 360°), more explicitly, moving from red to yellow to green to cyan to blue to magenta and back to red, ii) the saturation that corresponds to the distance from the axis (i.e., radius), the brightness of the colour, and iii) the value indicating the luminance or intensity.

In HSV, the component, hue, has the most control over the colour information compared to the other components in terms of determining the colour information whereas the saturation designates the colourfulness relative to the brightness based on the amount of light it appears to absorb and how much light it seems to be emitting. The saturation characteristics of ISMMMOs are significantly different from those of the sea background and maritime animals, as explained earlier. Therefore, we process the chromatic hue and saturation components to reveal the artificial objects not belonging to the natural marine environment. First, the hue component is shifted by 180° to suppress the blueish background into reddish (Fig. 4) as shown in the examples in Figs. 11c, 12c, 13c, 14c, 15c and in the technical reports in the supplements. Second, more saturated sections of the image are made more distinctive as explained in Section 2.2.

2.2. Dynamic thresholding in S channel

It is observed that the closer the values of histogram *S* are to the centre, with respect to the distribution of histograms, the likelier the pixels are of representing the background and natural marine life, and vice versa the more likelier they represent ISMMMOs wherever these values get away from the axis meaning that saturation is greater. However, there is no specific value that makes this separation distinct based on the different features of the images acquired in different circumstances, mainly different lighting times of the day, month, season, and type of camera. Furthermore, the distribution of the histogram values plays a major role in representing the characteristics of the image regarding the colourfulness relative to the brightness to which to which an area radiate a varying amount of light as explained in Section 1. The objective is to separate more saturated regions from less saturated ones to determine if there is an unnatural object. All threshold values and necessary parameters need to be determined based on the distribution and features of datasets in many surveys without any user intervention due to the maritime dynamics and image capturing techniques. It is noteworthy to emphasise that the saturation values are almost normally distributed with a Gaussian function as displayed in Eq. 1. The exact

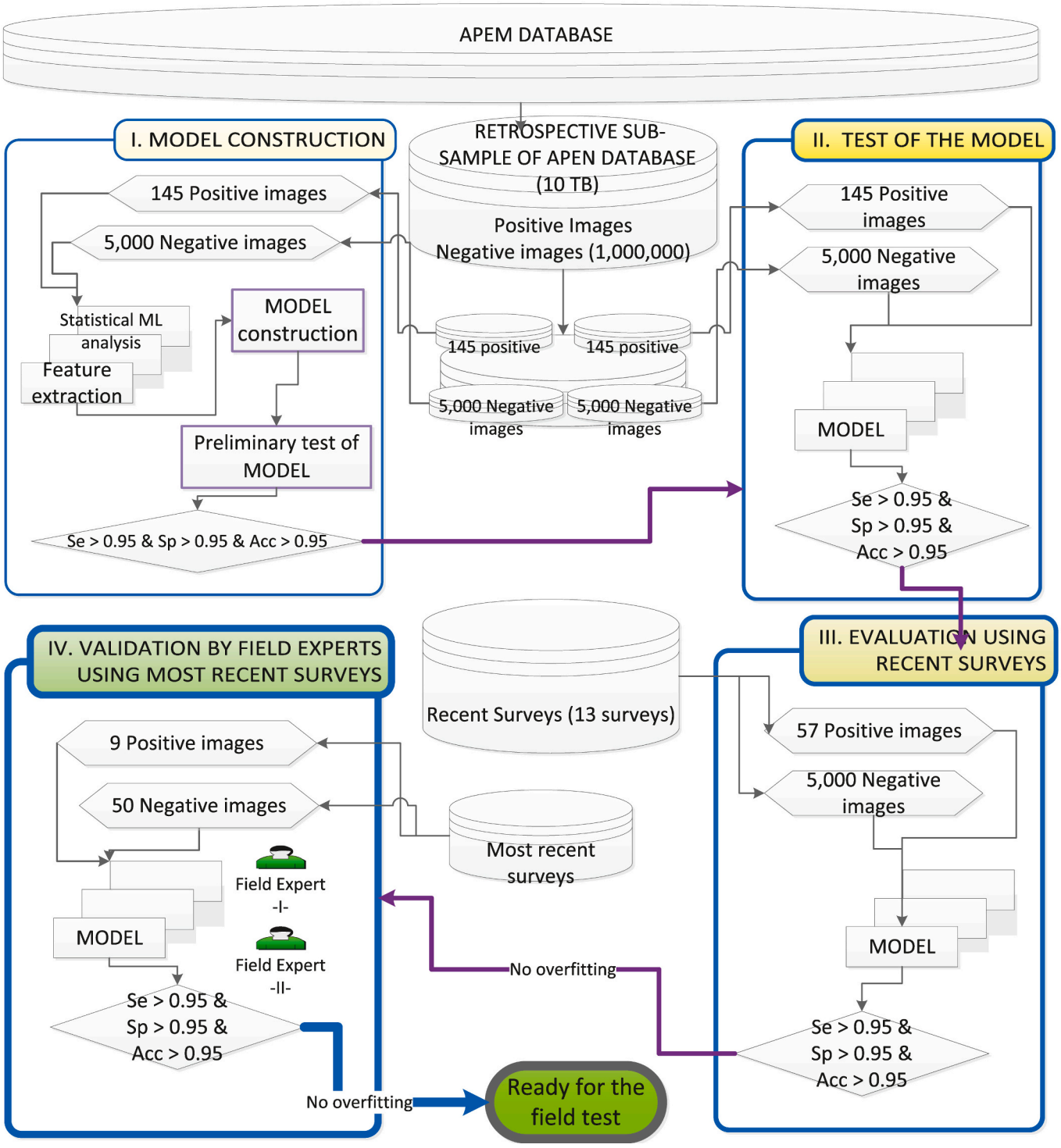


Fig. 3. Use of datasets during model construction, testing, evaluation, and validation of the model.

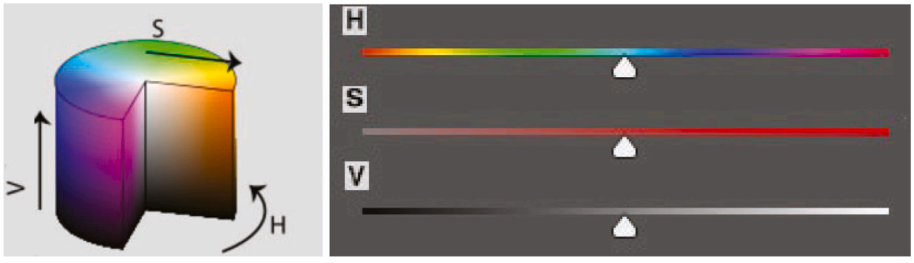


Fig. 4. Model of the HSV colour space.
The left image is the courtesy of the author in (Rosebrock, 2017).

distribution of data points using this Gaussian function is presented in Fig. 5 with respect to the σ.

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \quad (1)$$

In the first instance, a viable threshold value that separates more saturated regions from less saturated ones is found using 145 images with ISMMMOs and 5000 images with no ISMMMOs which were acquired from 22 aerial surveys as shown Fig. 3 I. A ROC curve is an ideal figure to observe how the classification model performs at various classification cut-off points using TPR (True Positive Rate) and FPR (False Positive Rate) (1-TNR) (Table 3). Hence, a ROC curve is established using a large set of threshold values, i.e., cut-off points (i.e., 17) for the purpose of determining the optimal cut-off point, which is the point of the curve nearest to the upper left-hand corner. The results are shown in Table 4. The optimum cut-off point, 0.17, is found, which is between the cut-off points of 0.15 and 0.20 as displayed in Fig. 6, and this results in 0.856 (i.e., TP = 124) and 0.817 (i.e., TN = 817) and 0.80 for sensitivity (*Se*), specificity (*Sp*) and accuracy (i.e., ACC) respectively. However, these outcomes are far away from our objectives in terms of separating images with ISMMMOs from others within large-scale surveys with higher accuracy rates. In other words, in order to achieve the desired separation (i.e., (*Se*) > 0.95, *Sp* > 0.95, and ACC > 0.95), a curve that is much closer to the top left-hand side of the ROC figure is required where the area under the ROC curve (i.e., AUC) increases, which is a desirable outcome for a test.

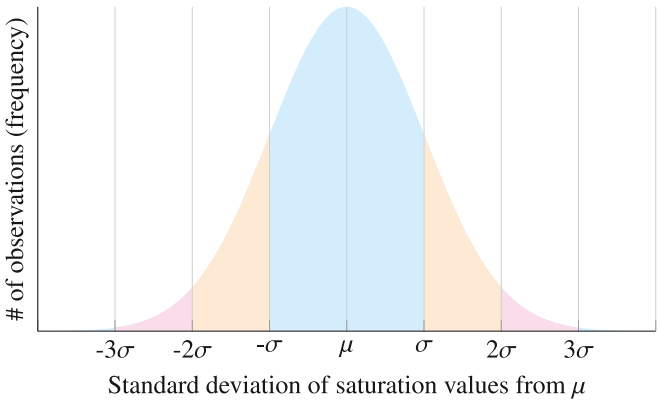


Fig. 5. Generation of the probability distribution using the Gaussian function in Eq. 1 and representation of the standard deviation (σ) of saturation values from μ. Areas: blue (one σ of μ): 0.6826; orange: 0.2718; magenta: 0.0428; sides (right of 3σ and left of -3σ): 0.0027. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

they are classified based on their μ values and iii) those classes are analysed separately to find out the best cut-off points for each class. The sections on which the ROC curves are analysed are depicted in Figs. 7, 8, 9, 10 and in Tables 5, 6, 7 and 8 based on the distribution of the histogram using the statistical analysis of the μ and σ values where the cut-off points on the ROC curves are specified based on the times of σ in the both directions of μ (Fig. 5).

Algorithm 1: HSV colour space adjustment function where hAdjust = 0.17 and sAdjustPar = 180.

```
FUNCTION newImage = HSVadjustManMade(img, hAdjust,sAdjustMask):
– >convert the image from RGB colour space into HSV space;
hsvVAL = rgb2hsv(img);
hue = 360*hsvVAL(:,:,1);
h = hsvVAL(:,:,1); s = hsvVAL(:,:,2); v = hsvVAL(:,:,3);
meanS = mean2(s);
stdS = std2(s);
– >perform the dynamic thresholding for S channel;
if meanS > 0.50 then
    sAdjust = meanS - 2*stdS;
else if meanS > 0.25 then
    sAdjust = meanS - stdS/2;
else if meanS > sAdjustMask then
    sAdjust = meanS;
else
    sAdjust = meanS + 4*stdS;
s (s > sAdjust) = 1;
– >perform the shifting of H channel;
h = (mod(hue + hAdjust, 360)) / 360;
– >acquire the updated RGB image from HSV space;
hsvVAL(:, :, 1) = h; hsvVAL(:, :, 2) = s; hsvVAL(:, :, 3) = v;
newImage = hsv2rgb(hsvVAL);
```

The saturation varies significantly, in particular, from one survey to another based on the changing conditions as mentioned above and demonstrated in the technical reports in the supplements with many examples.[4] Therefore, the designs of various ROC curves are based on the several most important sections of the histogram concerning the distribution of the saturation, and *Se* and *Sp* values by determining the required number of dynamic cut-off points for increasing the *Se* and *Sp* values significantly. Technically speaking, i) the mean values (μ) and standard deviations (σ) are acquired following the histogram of the *S* components are obtained from those 145 images mentioned earlier, ii)

The number of the cut-off points for each class is specified based on the distribution of the histogram values. This analysis is mainly carried out to find out i) if there is an evident saturated region in the image that distinctively differs from the other majority regions regarding the features of saturation and most importantly ii) what the best cut-off points making this distinction resulting in higher *Se* and *Sp* values are. The histogram values based on the obtained best cut-off points are transformed to the most outer side of the radius in *S* channel and set to 1 to make the most saturated sections more distinct, in other words, the probable ISMMMOs visible using the masking and dilation techniques mentioned in Section 2.3. Several examples are presented in Figs. 11, 12, 13, 14 and 15. The observed best cut-off points regarding the analysed sections along with their *Se* and *Sp* values in those ROC curves are

[4] The reports from 1 to 7 titled as MarineObjects_Man-made_Supplement are for ISMMMOs and the reports from 1 to 5 titled as MarineObjects_Man-made_Supplement_Blank are for blank images.

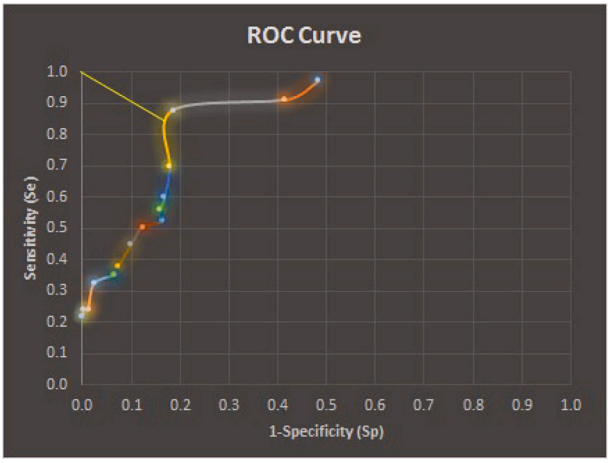


Fig. 6. ROC Curve for Table 4 based on TPR (y-axis) and FPR (x-axis) at 17 classification thresholds: The best cut-off point is 0.17 that is closest to the upper left corner of the curve between the cut-off points 0.15 and 0.20.

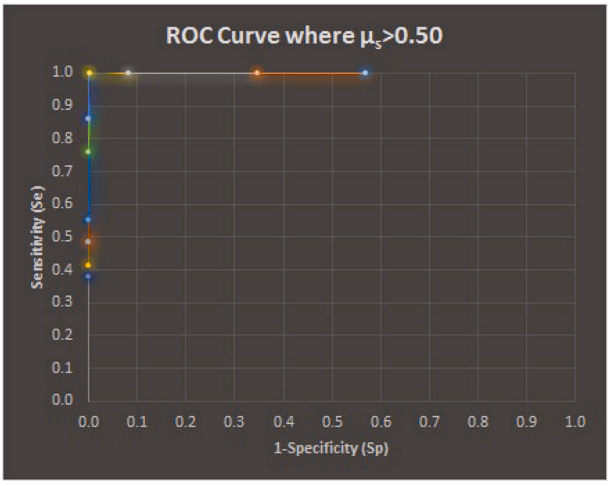


Fig. 7. ROC Curve for Table 5: The best cut-off point is μ - 2σ.

summarised in Table 9. The implementation of dynamic thresholding is presented in Algorithm 1 and exemplified in Figs. 11d, 12d, 13d, 14d, 15d with several examples along with H shifting whose new histograms are presented in the second rows of Figs. 11b, 12b, 13b, 14b, 15b.

The methodology was developed using the characteristics and distribution of 22 surveys with around 3 million large-scale images that have been acquired in the various geographical regions, and in the various time zones and seasons. The images with no ISMMMOs were exploited to obtain the general characteristic of the background whereas the images with ISMMMOs were used to determine the general characteristics of ISMMMOs. Both features are merged in the methodology to distinguish the ISMMMOs from its background and consequently discern the positive images from the negative images for further analysis.

2.3. Masking and dilation

Two masks are applied on the image acquired from the dynamic thresholding technique on H and S channels mentioned above, one of which is for detecting the blueish part and the other one is for removing the unwanted background parts from the image. First mask (i.e.,$((ImgR < 0.25 \& ImgG < 0.80 \& ImgB = 1) \& (ImgR < (ImgB) \& ImgG < (ImgB)))$) makes the blueish sections visible by suppressing all other sections, in particular, reddish parts that dominantly indicate the background of sea as depicted in Figs. 11e, 12e, 13e, 14e, 15e and in our

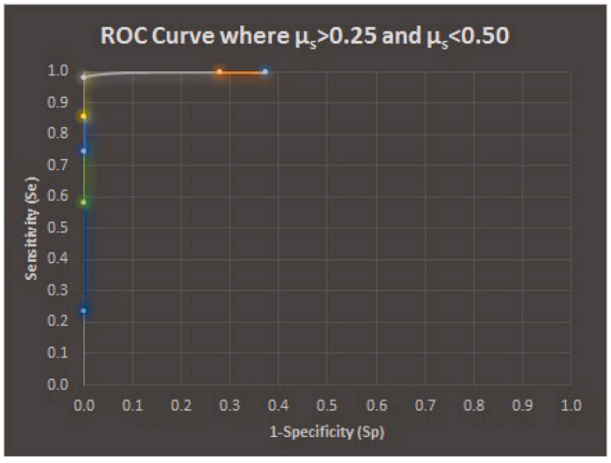


Fig. 8. ROC Curve for Table 6: The best cut-off point is μ - σ/2.

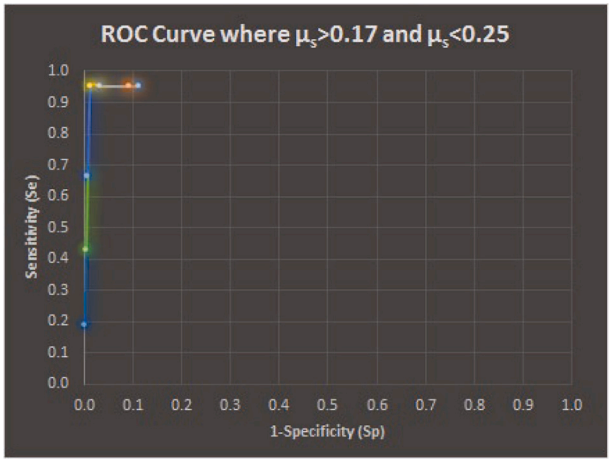


Fig. 9. ROC Curve for Table 7: The best cut-off point is μ.

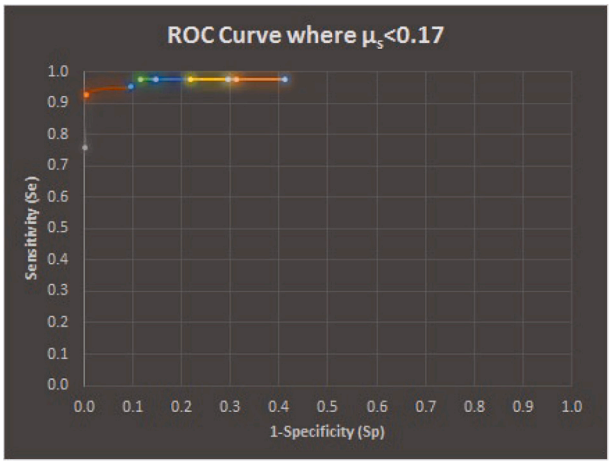
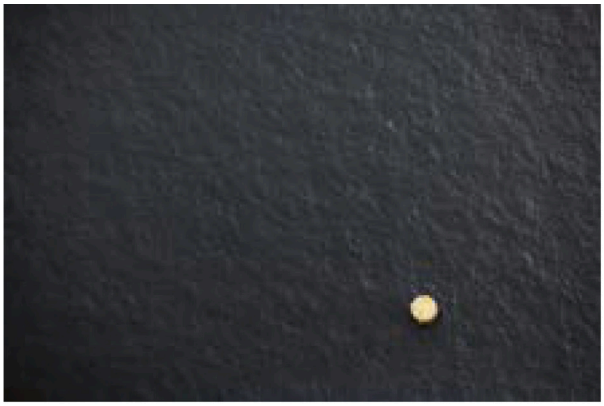


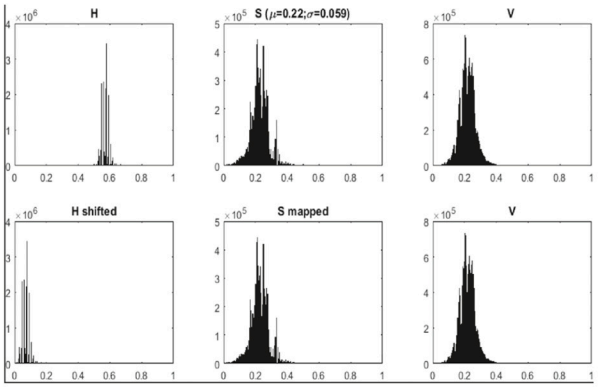
Fig. 10. ROC Curve for μ_s where <0.17 (Table 8): The best cut-off point is μ + 4σ.

technical reports. After applying this mask, the obtained image is dilated and holes are filled to make ISMMMOs coherent as shown in Figs. 11f, 12f, 13f, 14f, 15f. This process is mainly performed to gain the complete white areas of objects that are not obtained with the proposed technique as elaborated in Sections 5 and 7.

There might be several small unwanted dots that are not a part of the



(a) Original image



(b) Histogram of H, S, V channels and dynamic thresholding: first row original; second row converted

(c) Conversion with H channel values

(d) Conversion with new H and S channel values

(e) Use of threshold and mask

(f) Filling objects with closed boundaries

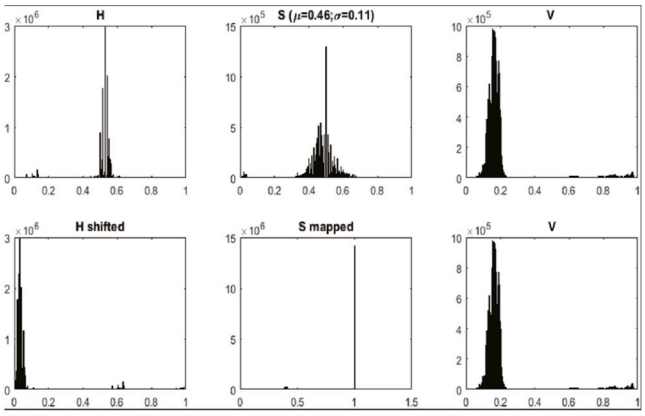
Fig. 11. Stationary example 1: man-made object detection.

ISMMMOs after applying the first mask, usually a process of glinting sections after the HSV processing phase. Around 20% of the blank images come up with similar small dots usually after dilation and filling holes in the image as depicted in Fig. 16h). Several examples for these types of processed images can be reached from our technical report (e.g., examples 4, 6, 7, 15, 17, 19, 20 in MarineObjects_Man-made_Technical_Blank_1.pdf) in the supplements. These small sections are much smaller than the ISMMMOs and are discarded by applying a size mask technique. In the last phase of the implementation, the images with detected ISMMMOs are labelled as positive images and placed in a separate directory by the application for further analysis.

3. Experimental design

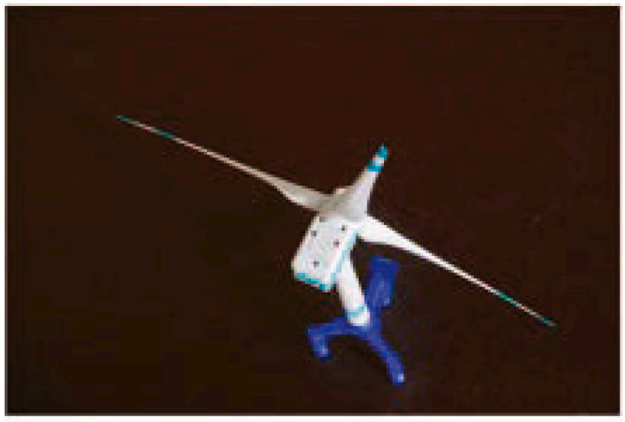
Offshore digital wildlife surveys for the offshore renewables sector are performed by APEM, capturing high-quality images year-round in all light conditions and up to four different sea states. The data is recorded using a wide range of advanced, high-resolution photogrammetry sensor technologies, including 35 mm and medium format sensors from a variety of manufacturers, in either multiple camera or a single camera configurations, subject to the scope of the project. These high-tech cameras, enabling a very high resolution ranging from 35MP to 50MP, are mounted in a tiny twin-engine aircraft (e.g., Fig. 17) on a route where all areas of interest are monitored with geospatial data (i.e., latitude, longitude, and altitude). It is noteworthy to emphasise that we have followed the standardised way of constructing applications for real-world uses with the development phases of i) build the model using a dataset and move to the second phase if the test results are satisfactory ii) test/evaluate the model using another dataset completely different from the first dataset to observe if the test results are satisfactory without

(a) Original image

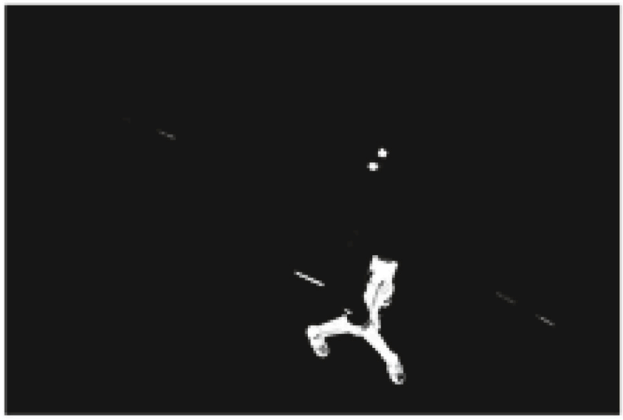


(b) Histogram of H, S, V channels and dynamic thresholding: first row original; second row converted

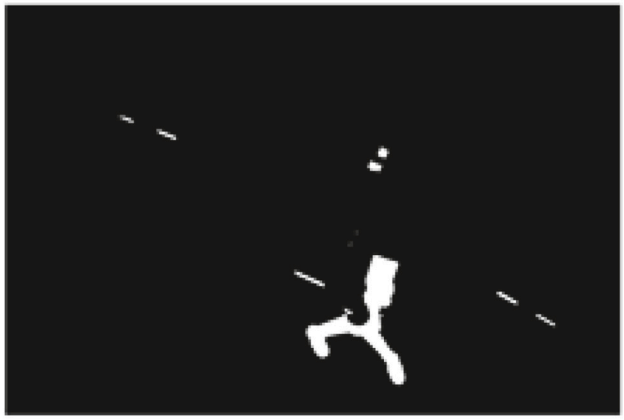
(c) Conversion with H channel values



(d) Conversion with new H and S channel values



(e) Use of threshold and mask



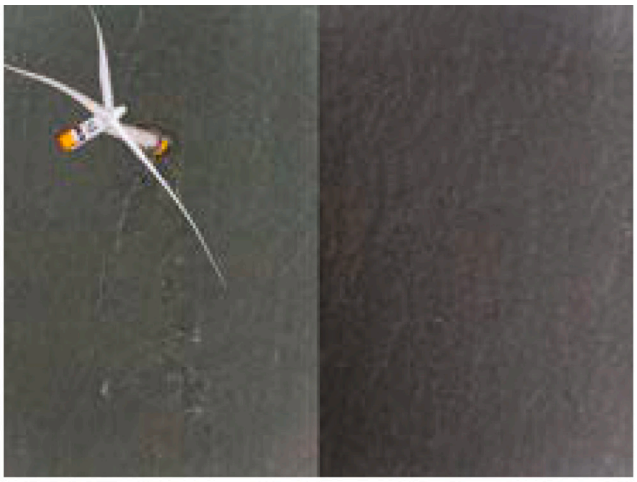
(f) Filling objects with closed boundaries

Fig. 12. Stationary example 2: man-made object detection.

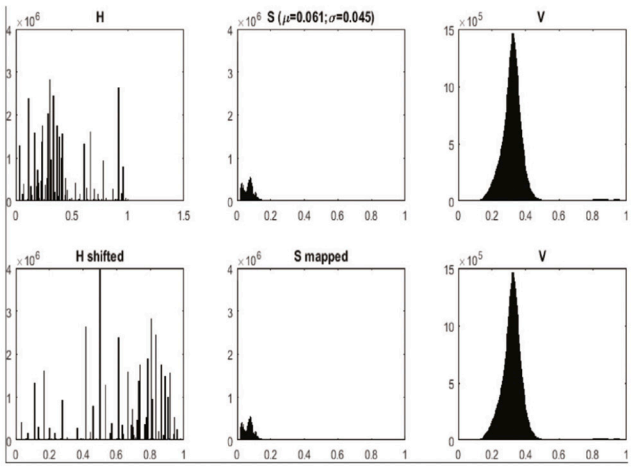
overfitting, and finally iii) let field experts evaluate the model with a completely new dataset independent from the first and second datasets. The model can be deployed if it passes these three phases successfully. These phases are outlined in Fig. 3. The obtained results as well as their evaluation are provided in the following section. The experimental design of data utilisation and data processing phases with their targeted objectives are outlined in Table 3 regarding the APEM's database. The viability of the methodology was ensured in 4 phases.

Phase I. Model construction (Fig. 3 I): The proposed methodology was established using 145 images with ISMMMOs and 5000 images with no ISMMMOs acquired from the 22 surveys between 2014 and 2017, with around 250 samples from each survey. The sub-samples of these surveys have around 3 million large-scale images that have been obtained from the various areas of the world in all seasons and numerous time zones. This large number of surveys enabled us to identify the broad features and parameters of aerial surveys and apply these parameters to make our methodology robust. All the steps of the model construction phase are explored in the sections above in detail. Phase II was conducted after the successful execution of Phase I by realising the targeted objectives, which is elaborated as follows.

Phase II. Test of the model (Fig. 3 II): In addition to the dataset used for the establishment of the methodology, a test dataset was prepared. This set was composed of 55 images with ISMMMOs and 5000 images with no ISMMMOs, The test results are displayed in Table 10 A. We

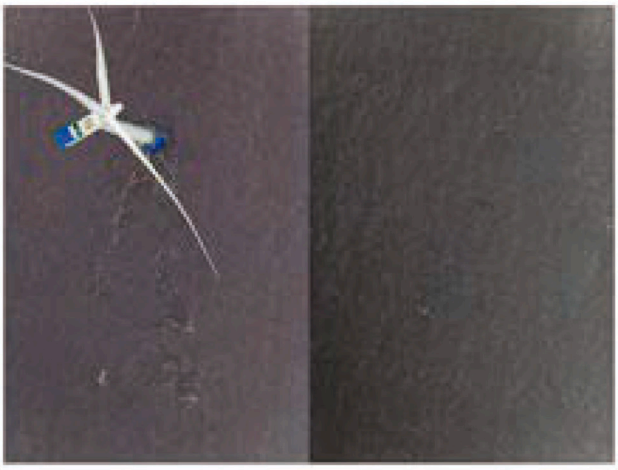


(a) Original image



(b) Histogram of H, S, V channels and dynamic thresholding: first row original; second row converted

(c) Conversion with H channel values



(d) Conversion with new H and S channel values

(e) Use of threshold and mask

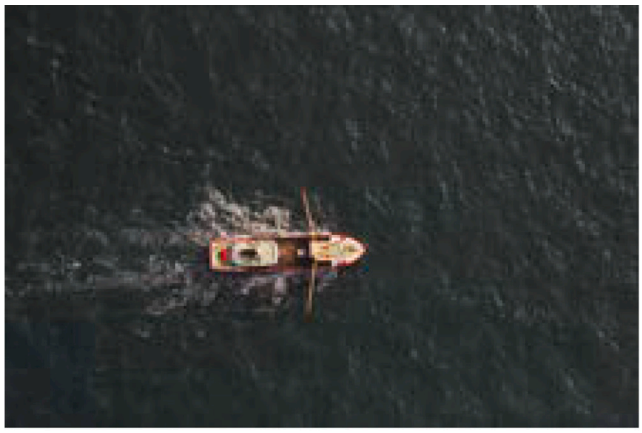
(f) Filling objects with closed boundaries

Fig. 13. Stationary example 3: man-made object detection.

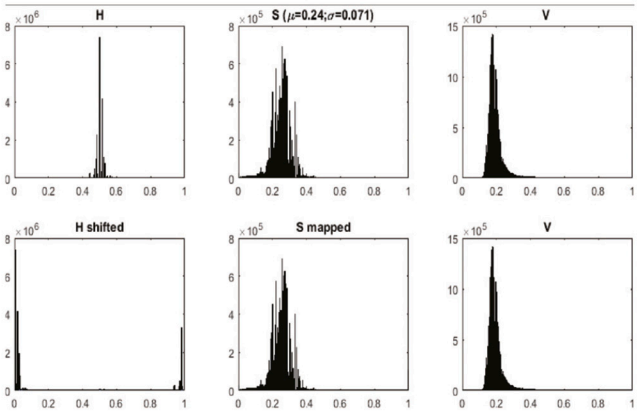
moved to the next phase to evaluate the system using independent datasets after the satisfactory results (*Se*, *Sp*, *PPV*, *NPV*, and $ACC > 0.95$) obtained in this phase.

Phase III. Evaluation using recent surveys (Fig. 3 III): A dataset was prepared to evaluate the eligibility of the methodology. This set consists of 57 images with ISMMMOs and 5000 images with no ISMMMOs. This set is not included in the dataset used for the establishment of the methodology to observe if the methodology works as desired for other independent datasets. The test results are displayed in Table 10 B. We moved to the next phase to verify the system with field experts using other independent datasets after the satisfactory results (*Se*, *Sp*, *PPV*, *NPV*, and $ACC > 0.95$) obtained in this phase.

Phase IV. Validation by field experts using the most recent surveys (Fig. 3 IV): Furthermore, in an independent verification dataset,



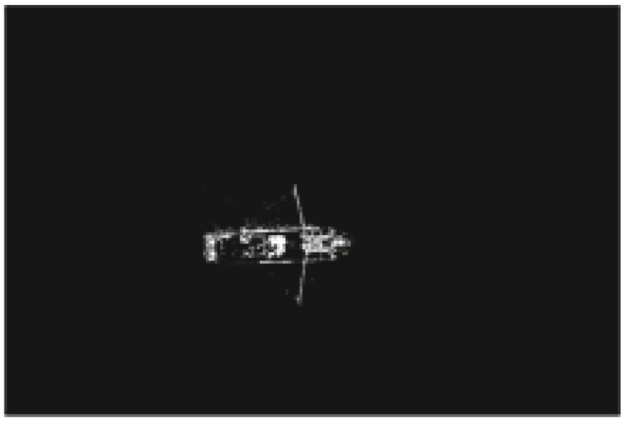
(a) Original image



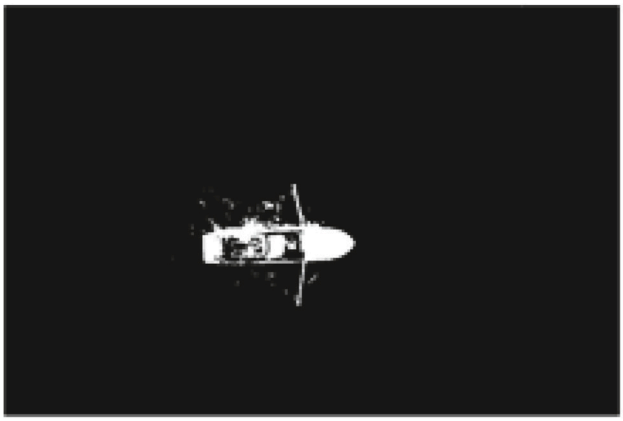
(b) Histogram of H, S, V channels and dynamic thresholding: first row original; second row converted

(c) Conversion with H channel values

(d) Conversion with new H and S channel values



(e) Use of threshold and mask



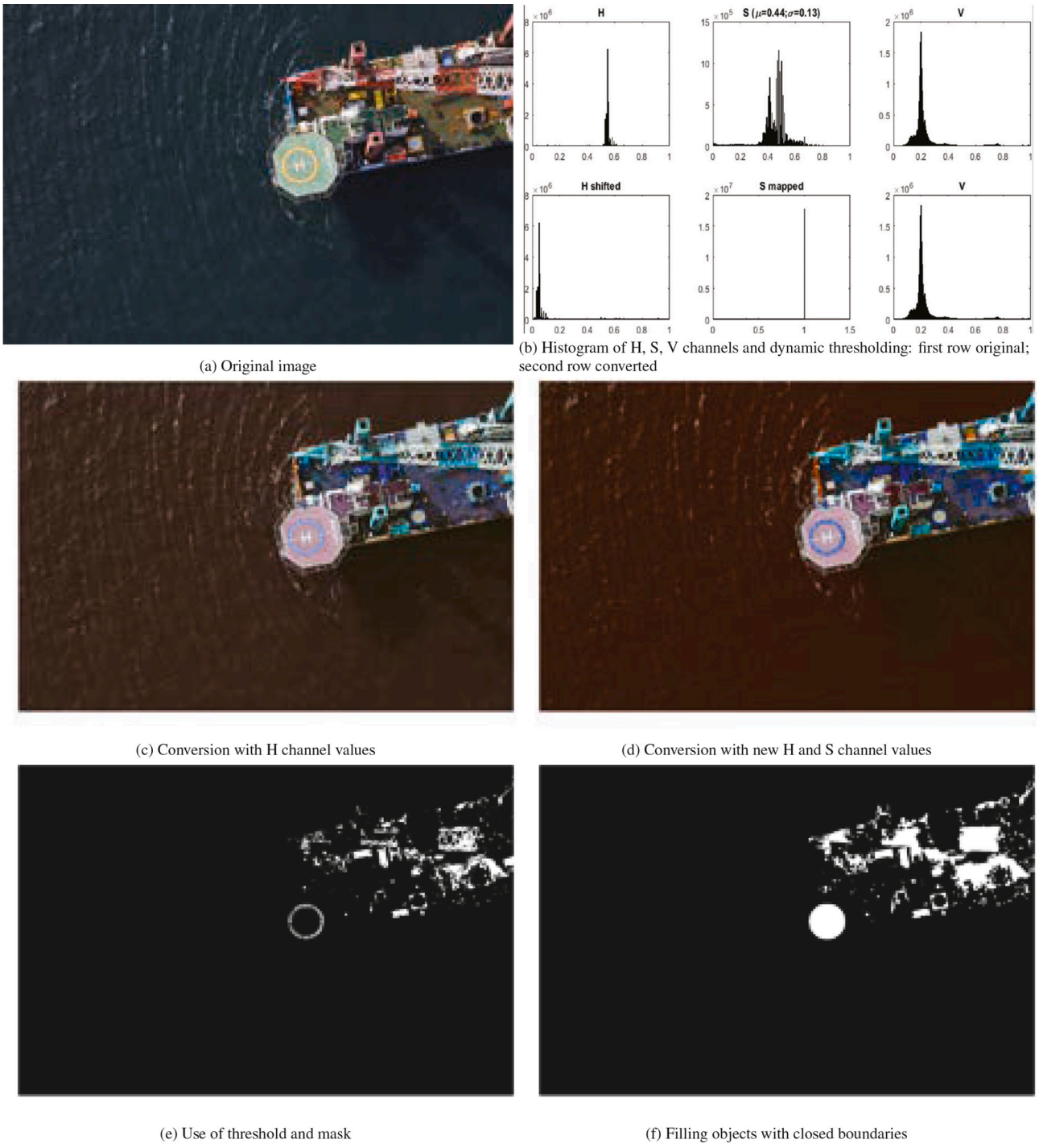
(f) Filling objects with closed boundaries

Fig. 14. Moving example 1: man-made object detection.

9 more images with ISMMMOs and 50 images with no ISMMMOs in different surveys from the surveys on which the methodology was established were provided by APEM for affirming the viability of the system to observe if the methodology can work as desired for any aerial datasets. Two field experts from APEM Ltd. confirmed that the established system can meet their needs to detect ISMMMOs while performing surveys. The test results are displayed in Table 10 C. The results (*Se*, *Sp*, *PPV*, *NPV*, and $ACC > 0.95$) obtained in this phase were found to be highly satisfactory by the field experts.

4. Results

The effectiveness of the proposed methodology in detecting images with ISMMMOs is demonstrated by several experiments performed on many aerial survey images as elaborated in Section 3. The results of these experiments are outlined in Table 10 and they are summarised in Table 11. The numerous tangible outcomes of these successful results are demonstrated in the supplementary technical reports of the paper and in Figs. 11f, 12f, 13f, 14f, 15f. With this approach, ISMMMOs can be captured with *Se*, *Sp*, *PPV*, *NPV*, and *ACC* values over 0.95. More specifically, 140 images out of 145, 55 images out of 57 and 9 images out of 9 in the test, evaluation and validation phases (Fig. 3)are tagged as the



(a) Original image

(b) Histogram of H, S, V channels and dynamic thresholding: first row original; second row converted

(c) Conversion with H channel values

(d) Conversion with new H and S channel values

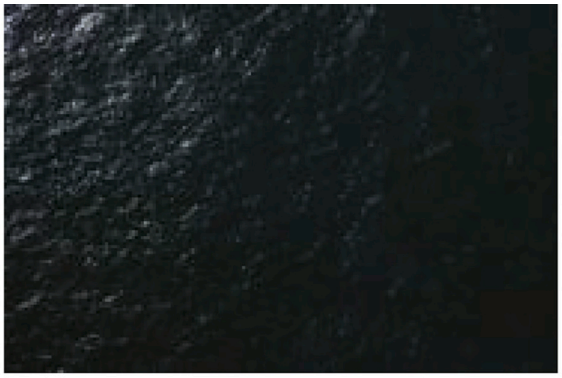
(e) Use of threshold and mask

(f) Filling objects with closed boundaries

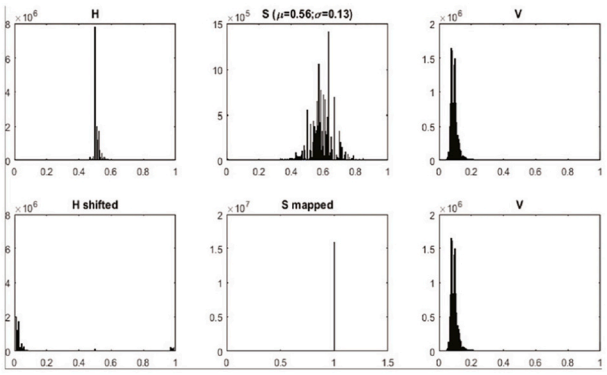
Fig. 15. Moving example 2: man-made object detection.

images with ISMMMOs successfully with a high *Se* over the targeted value (> 0.95) in the research, which indicates that the methodology is strong in separating positive images from negative ones in situations where it is preferable to not miss positive images. The particular results of these phases are averaged at the bottom column in Table 11. All the averages are higher than 0.95, which indicates that the required phases were completed successfully and the application is ready for real-world deployment (Fig. 3). It is noteworthy to emphasise that Precision (*Pr*), i.e., Positive Predictive Value – $PPV = Pr = \mathrm{TP} / (\mathrm{TP} + \mathrm{FP})$ – is 0.9813 for the average. This high value demonstrates that the model is highly successful in assigning "Positive" images to the "Positive" class while the 'Negative" images are assigned to the "Negative" class effectively with a NPV of 0.9995. Most importantly, we calculated Matthews Correlation Coefficient (MCC) due to an unbalanced number in the classes where the number of negative values were high, which may yield misleading ACC values. The MCC, ranging from -1 to 1, was found to be 0.971, which indicates that the model is very close to a perfect prediction (i.e., 1).

During the implementation of the methodology, testing, and evaluation, it was observed that the ISMMMOs with completely white features (i.e., $R = 255$, $G = 255$, $B = 255$) had difficulties being detected by our methodology since they have the same characteristics as waves with respect to HSV conversion, where zero is assigned to the S component during the conversion from RGB to HSV mode (Table 1, Table 2). In this respect, it is worth noting that the images with ISMMMOs that could not



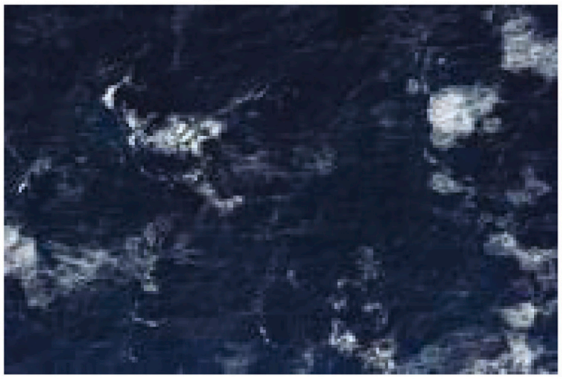
(a) Original image



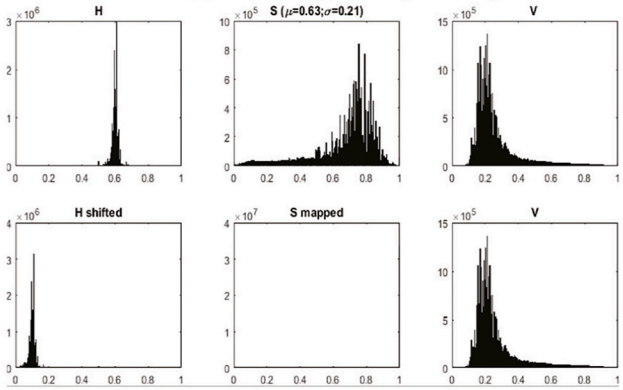
(b) Histogram of H, S, V channels:first row original;second row converted

(c) HSV conversion

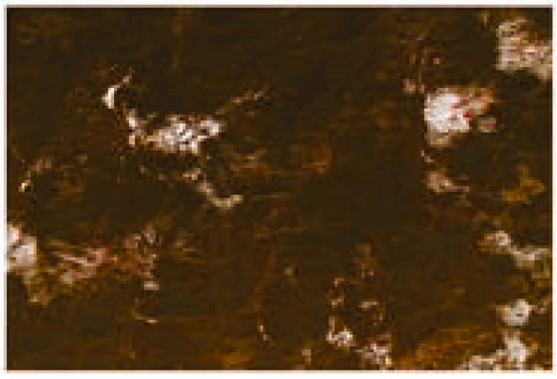
(d) Thresholding, masking and filling



(e) Original image



(f) Histogram of H, S, V channels:first row original;second row converted



(g) HSV conversion

(h) Thresholding, masking and filling

Fig. 16. Examples for blank images with no man-made objects.

Fig. 17. APEM aircraft during an aerial survey.

be detected during the design and development phases (Fig. 3) are these types of images. Examples of these images are presented in Appendix B (Fig. 18). We refer the readers to Fig. 13 to observe how the white parts of the wind turbine cannot be detected adequately.

5. Discussion

Gibert et al. (Gibert et al., 2018) defined Data Science as a multidisciplinary field that is a combination of data analysis, data processing techniques, and domain knowledge that transforms data into comprehensible and actionable insights relevant to making informed decisions. Within this context, the objective of this study is to create a new environmental platform for the monitoring of the maritime environment by combining domain knowledge and data scientists in a productive collaboration and perform the detection of mobile and stationary ISMMMOs in an automated manner with their geospatial coordinate system. Changes in the marine ecosystem, such as habitat loss or population decreases in marine organisms, may not be readily foreseeable and it requires long-term studies to reveal the environmental changes and impacts on the ecosystem and consequently to determine the required policies accordingly. Studies in marine environments, especially far offshore, are comparatively costly and require the employment of new automatic techniques and merge of different studies for field researchers. In this sense, this study intends to help authorities and researchers with the automatic detection of offshore ISMMMOs using an advanced platform to fill some of this gap.

The robustness of the platform was validated on a wide range of aerial maritime domains, providing a high level of empirical proof of concept with successful results (Table 11). Strictly speaking, the experimental results show that the proposed approach is efficient and effective for the detection and the segmentation of ISMMMOs in large scale aerial images. More specifically, the dynamic thresholding approach employed in the methodology increases *Se* from 0.85 to 0.97 and *Sp* from 0.82 to 0.99 when compared to the static optimum threshold value as displayed in Table 4. This increase is statistically significant ($p<0.01$) by rejecting the null hypothesis (i.e., there is no significant difference between two results) using a paired-samples *t*-test. The ISMMMOs not detected by the methodology are all complete white objects. This issue is specified in Section 7 as a limitation of the study. Furthermore, the evaluation and validation results using the new datasets (Table 10) that were not in the surveys used during the establishment of the methodology (Fig. 3) demonstrate that the methodology can work effectively on any aerial survey with high accuracy rates. In other words, during the evaluation phase, 55 out of 57 images with ISMMMOs were put into the positive folder and 4998 out of 5000 images with no ISMMMOs were placed into the negative folder. During the validation by field experts, 9 out of 9 images with ISMMMOs were put into the positive folder with all objects detected successfully and 50 out of 50 images with no ISMMMOs were placed into the negative folder successfully. It must be noted that the developed methodology neither classifies the detected ISMMMOs into groups nor determines the recognition of them, such as "ship", "wind turbine" etc. Particular classification tools need to be developed to group the ISMMMOs that are placed in the positive

Table 1
Formulas for converting from RGB to HSV colour space.

Channel	Formula	Condition / Output	
Rnew	= Red / 255	Output: 0 < Rnew < 1	Red value mapped between 0 and 1
Gnew	= Green / 255	Output: 0 < Gnew < 1	Green value mapped between 0 and 1
Bnew	= Blue / 255	Output: 0 < Bnew < 1	Blue value mapped between 0 and 1
MaxRGB	= Max (Rnew, Gnew, Bnew)	Output: maximum assigned	Maximum of three components found
MinRGB	= Min (Rnew, Gnew, Bnew)	Output: minimum assigned	Minimum of three components found
Hue (H)	= 0	If Rnew = Gnew = Bnew	Zero assigned
	= 60 * ((Gnew - Bnew) / (MaxRGB - MinRGB)) + 0	If max is Rnew	
	= 60 * ((Bnew - Rnew) / (MaxRGB - MinRGB)) + 2	If max is Gnew	Add 360 to Hue if H < 0
	= 60 * ((Rnew - Gnew) / (MaxRGB - MinRGB)) + 4	If max is Bnew	
Saturation (S)	= 0	If Rnew = Gnew = Bnew	Zero assigned
	= (MaxRGB - MinRGB) / MaxRGB	Else	If all components different
Value (V)	= MaxRGB	Output: maximum assigned	Maximum of three components assigned

Table 2
Formulas for converting from HSV to RGB colour space.

Channel	Formula	Condition / Output	
Chroma (Chr)	= S * V	Output: Colourfulness	Intermediate colour purity calculated
MidX value	$= Chr(1 - \lvert(H/60) mod 2 - 1\rvert)$	Output: Mid value	Intermediate conversion value calculated
MidM value	= V - Chr	Output: Mid parameter	Intermediate parameter calculated
(MidR, MidG, MidB)	= (Chr, MidX, 0)	If $0 \leq H < 60$	Intermediate red, green and blue calculated based on Hue
	= (MidX, Chr, 0)	If $60 \leq H < 120$	Intermediate red, green and blue calculated based on Hue
	= (0, Chr, MidX)	If $120 \leq H < 180$	Intermediate red, green and blue calculated based on Hue
	= (0, MidX, Chr)	If $180 \leq H < 240$	Intermediate red, green and blue calculated based on Hue
	= (MidX, 0, Chr)	If $240 \leq H < 300$	Intermediate red, green and blue calculated based on Hue
	= (Chr, 0, MidX)	If $300 \leq H < 360$	Intermediate red, green and blue calculated based on Hue
Rback	= (MidR + MidM) × 255	Output: Targeted red value	Red value calculated using intermediate parameter
Gback	= (MidG + MidM) × 255	Output: Targeted green value	Green value calculated using intermediate parameter
Bback	= (MidX + MidM) * 255	Output: Targeted blue value	Blue value calculated using intermediate parameter

Table 3
Main statistical terms and calculations used throughout the paper.

#	Abbreviation	Description	Detail
1	P	Positive	An image with ISMMMO
2	N	Negative	An image with no ISMMMO
3	TP	True Positive	An image with ISMMMO is tagged as "image with ISMMMO"
4	TN	True Negative	An image without ISMMMO is tagged as "image without ISMMMO"
5	FP	False Positive (False Alarm)	An image without ISMMMO is tagged as "image with ISMMMO"
6	FN	False Negative	An image with ISMMMO is tagged as "image without ISMMMO"
7	Se	Sensitivity	True Positive Rate (TPR) = Se = TP / (TP + FN) How strong is the test in detecting images with ISMMMO correctly.
8	Sp	Specificity	True Negative Rate (TNR) = Sp = TN / (TN + FP) How strong is the test in detecting images without ISMMMO correctly.
9	PPV	Positive Predictive Value	Precision (Pr) = PPV = TP / (TP + FP) How strong is the test in assigning images with ISMMMO to Positive class.
10	NPV	Negative Predictive Value	NPV = TN / (TN / FN) How strong is the test in assigning images without ISMMMO to Negative class.
11	ACC	Accuracy	ACC = (TN + TP) / (FP + FN + TP + TN) Overall correct assignment rate of the test.
12	MCC	Matthews Correlation Coefficient	$MCC = (TN * TP - FN * FP) / \sqrt{((TP+FN)(TP+FP)(TN+FN)(TN+FP))}$ Quality of a test concerning the unbalance in classes
13	μ	Mean	Arithmetic average of a set of observed values.
14	σ	Standard deviation	Measurement of variation, dispersion from the average, within a set of observed values.

Table 4
Finding the optimum point using 17 cut-off points. 145 images with ISMMMOs.

Cut-off	*TP*	*FN*	*TN*	*FP*	*Se*	*Sp*	*1-Sp*
0.05	141	4	517	483	0.972	0.517	0.483
0.10	132	13	585	415	0.910	0.585	0.415
0.15	127	18	814	186	0.876	0.814	0.186
0.20	101	44	821	179	0.697	0.821	0.179
0.25	87	58	832	168	0.600	0.832	0.168
0.30	81	64	841	159	0.559	0.841	0.159
0.35	76	69	835	165	0.524	0.835	0.165
0.40	73	72	877	123	0.503	0.877	0.123
0.45	65	80	901	99	0.448	0.901	0.099
0.50	55	90	927	73	0.379	0.927	0.073
0.55	51	94	932	68	0.352	0.932	0.068
0.60	51	94	936	64	0.352	0.936	0.064
0.65	47	98	977	23	0.324	0.977	0.023
0.70	35	110	988	12	0.241	0.988	0.012
0.75	35	110	997	3	0.241	0.997	0.003
0.80	32	113	1000	0	0.221	1.000	0.000
0.85	32	113	1000	0	0.221	1.000	0.000

Table 5
Statistical analysis using 11 cut-off points: 29 images with ISMMMOs where $\mu_s > 0.50$.

Cut-off	*TP*	*FN*	*TN*	*FP*	*Se*	*Sp*	*1-Sp*
μ **- 5**σ	29	0	432	568	1.000	0.398	0.602
μ **- 4**σ	29	0	653	347	1.000	0.653	0.347
μ **- 3**σ	29	0	917	83	1.000	0.917	0.083
μ - 2σ	29	0	997	3	1.000	0.997	0.003
μ **-** σ	25	4	999	1	0.862	0.999	0.001
μ	22	7	1000	0	0.759	1.000	0.000
$\mu + \sigma$	16	13	1000	0	0.552	1.000	0.000
$\mu +$ **2**σ	14	15	1000	0	0.483	1.000	0.000
$\mu +$ **3**σ	14	15	1000	0	0.483	1.000	0.000
$\mu +$ **4**σ	12	17	1000	0	0.414	1.000	0.000
μ **- 5**σ	11	18	1000	0	0.379	1.000	0.000

directory by the proposed technique in this study, which is proposed as a future work in Section 6. Bespoke semi-supervised ML approaches (e.g., SelfMatch in (Xing et al., 2022a)) can be a good candidate for addressing this type of research question by extracting features from labelled data and comparing them with the features that are obtained from detected ISMMMOs based on the semantic information (e.g., (Xiao et al., 2022) and a feature/distance-based matching scheme (e.g., (Xing et al., 2022b)) considering various pose compositions (e.g., (Çalışkan, 2023)).

The methodology not only distinguishes ISMMMOs from the blank background (sea canvas with waves in many different shapes), but also

Table 6
Statistical analysis using 7 cut-off points. 55 images with ISMMMOs where $\mu_s > 0.25$ and < 0.50.

Cut-off	*TP*	*FN*	*TN*	*FP*	*Se*	*Sp*	*1-Sp*
μ **- 2**σ	55	0	627	373	1.000	0.627	0.373
μ **-** σ	55	0	721	279	1.000	0.721	0.279
μ - $\sigma/2$	54	1	998	2	0.982	0.998	0.002
μ	47	8	1000	0	0.855	1.000	0.000
$\mu + \sigma/2$	41	14	1000	0	0.745	1.000	0.000
$\mu + \sigma$	32	23	1000	0	0.582	1.000	0.000
$\mu +$ **2**σ	13	42	1000	0	0.236	1.000	0.000

Table 7
Statistical analysis using 7 cut-off points. 21 images with ISMMMOs where $\mu_s > 0.17$ and < 0.25.

Cut-off	*TP*	*FN*	*TN*	*FP*	*Se*	*Sp*	*1-Sp*
μ **- 2**σ	20	1	887	113	0.952	0.887	0.113
μ **-** σ	20	1	908	92	0.952	0.908	0.092
μ - $\sigma/2$	20	1	967	33	0.952	0.967	0.033
μ	20	1	987	13	0.952	0.987	0.013
$\mu + \sigma/2$	14	7	993	7	0.667	0.993	0.007
$\mu + \sigma$	9	12	995	5	0.429	0.995	0.005
$\mu +$ **2**σ	4	17	998	2	0.190	0.998	0.002

Table 8
Statistical analysis using 10 cut-off points. 40 images with ISMMMOs where $\mu_s < 0.17$.

Cut-off	*TP*	*FN*	*TN*	*FP*	*Se*	*Sp*	*1-Sp*
μ **-** σ	40	1	587	413	0.976	0.587	0.413
μ - $\sigma/2$	40	1	685	315	0.976	0.685	0.315
μ	40	1	704	296	0.976	0.704	0.296
$\mu + \sigma/2$	40	1	781	219	0.976	0.781	0.219
$\mu + \sigma$	40	1	851	149	0.976	0.851	0.149
$\mu +$ **2**σ	40	1	883	117	0.976	0.883	0.117
$\mu +$ **3**σ	39	2	903	97	0.951	0.903	0.097
$\mu +$ **4**σ	38	3	994	6	0.927	0.994	0.006
$\mu +$ **5**σ	31	10	997	3	0.756	0.997	0.003
$\mu +$ **6**σ	27	14	997	3	0.659	0.997	0.003

from other objects (e.g., different types of flying birds, sitting birds, big mammals (e.g., whales, dolphins), sharks, turtles, rays) in images with various shapes and characteristics successfully for which several examples can be reached in the technical report (e.g., Fig. 3 in MarineObjects_Man-made_Supplement_1.pdf) in the supplements. The results suggest that the saturation of maritime natural objects is significantly different from ISMMMOs. The processing time for each image

Table 9
Dynamic threshold points for S channel based on μ and σ acquired from the statistical analysis of the images using several cut-off points as depicted in Tables 5, 6, 7, 8. The most closest pointi to the upper left corner of the ROC curve indicate the ideal cut-off points as shown in Figs. 7, 8, 9, 10.

Mean of the S channel (μ)	threshold	objective	Se	Sp	example
> 0.50	$\mu - 2x\sigma$	Almost all values are mapped to 1	1.00	0.997	Fig. 16b
> 0.25	$\mu - \sigma/2$	Most of the values are mapped to 1	0.982	0.987	Fig. 12b
> sAdjustMask (i.e, 0.17)	μ	Almost half of the values are mapped to 1	0.927	0.994	Fig. 14b
otherwise (i.e., < 0.17)	$\mu + 4x\sigma$	Most of the values are not mapped to 1, left as is	0.96	0.95	Fig. 13b

Table 10
Detailed confusion matrix of the classifiers outlined in Table 11.

		A. Test Results (UCLAN)			B. Evaluation (UCLAN)			C. Validation (APEM)		
		Actual Class			Actual Class			Actual Class		
		Positive	Negative	%	Positive	Negative	%	Positive	Negative	%
Pred	Positive	140 (TP)	3 (FP)	0.9790 (PPV)	55 (TP)	2 (FP)	0.9649 (PPV)	9 (TP)	0 (FP)	1 (PPV)
	Negative	5 (FN)	4997 (TN)	0.9990 (NPV)	2 (FN)	4998 (TN)	0.9996 (NPV)	0 (FN)	50 (TN)	1 (NPV)
	%	0.9655 (Se)	0.9994 (Sp)	**0.9984 (ACC)**	0.9649 (Se)	0.9996 (Sp)	**0.9992 (ACC)**	1 (Se)	1 (Sp)	**1 (ACC)**

Table 11
Test, evaluation and validation results in summary detailed in Table 10.

Phase	*Positive*	Negative	*TP*	*FN*	*TN*	*FP*	*Se*	*SP*	*PPV*	*NPV*	*ACC*	*Location*	*Check*
Test	**145**	**5000**	140	5	4997	3	0.966	0.997	0.9790	0.9990	0.9984	UCLAN	✓
Evaluation	**57**	**5000**	55	2	4998	2	0.965	0.999	0.9649	0.9996	0.9992	UCLAN	✓✓
Validation	**9**	**50**	9	0	50	0	1	1	1	1	1	APEM	✓✓✓
Verification	**211**	**10,050**	**204**	**7**	**10,045**	**5**	**0.977**	**0.9987**	**0.9813**	**0.9995**	**0.9992**	**Average**	✓✓✓✓

varies from 7 s to 16 s, depending on the image size and the number of objects in the image and their sizes, which is a very fast processing time for high-pixels-per-image (HPP) images up to 50 MB based on the camera system that is explained in Section 3. The overall computational complexity of the developed algorithms is $O(n\ log\ n)$. It is important to point out that the supervised DL and ML approaches, designed by us in our previous work in (Kuru et al., 2023), that runs on similar images in the same surveys, can detect specific marine small natural objects (e.g., birds) in a few seconds (i.e., between 2 and 4 s). In this sense, we can conclude that DL and ML techniques slightly outperform the proposed non-supervised technique developed in this study considering the processing time.

The current rate of global environmental alteration necessitates the quantification of impacts in species abundance in order to evaluate the effects on the ecosystem. To assess the extent of the decline, effective long-term surveillance of populations and trends is required, which is rarely the case for most species (Rosenberg et al., 2019). Environmental models work better when they are based on the findings of more up-to-date data analysis on specific domains. It is essential to continuously monitor species and ISMMMOs in an automated manner cost-efficiently, which necessitates the utilisation of sophisticated equipment with effective intelligent surveillance methods. In this regard, WILDetect, which is a new non-parametric platform by utilising a combination of supervised ML and Reinforcement Learning (RL) methods, was built in our previous work in (Kuru et al., 2023) to carry out automated wildlife censuses in highly dynamic marine environments. With similar automated platforms, one of which is proposed in this research for detecting ISMMMOs, existing labour-intensive and costly censuses performed over long periods of time can be replaced by cost-efficient and highly automated computerised systems and they can be repeated automatically in regular, shorter periods. In this way, the environmental models, equipped with near-real-time outcomes for both marine wildlife and man-made presence, can foretell future trends with more realistic projections based on human footprint, which, in turn, help mitigate the potential damaging effects of human footprint.

6. Conclusions and future work

A novel methodology, the so-called ISMMMOD, that detects and splits ISMMMOs automatically in large-scale images in typical large marine surveys is built. The ISMMMOD is developed using the HSV colour space and statistical analysis of histograms of the channels in this space based on the ROC curve analysis. The techniques in the methodology differ man-made structures from natural maritime habitats (i.e., waves, sea animals, birds, seawater) in various aspects, in particular, composition, features of the surface and saturation of light. The large number of surveys, that were conducted in the various geographical regions and in the various time zones and seasons, on which our methodology was built, represent the key features of aerial surveys, which made our approach powerful and resilient in detecting ISMMMOs with very high accuracy rates. The successful results obtained in this research (Table 11) is an indication that using an automated computer-based system could be an effective alternative to labour-intensive approaches. The approach built in this study can be employed for several reasons, in particular, will provide researchers and policymakers with the ability to monitor maritime industries and ensure their proper deployment through the implementation of a suitable legal and regulatory framework that takes into account the changing dynamics of marine ecosystems. Additionally, this study will direct the researchers who would like to establish similar systems using unsupervised approaches.

The proposed method was tested on large-scale aerial images acquired by aeroplanes and we would like to observe the results of our method on satellite images wherever datasets are available, which may reduce the cost significantly regarding the detection of ISMMMOs and may provide a real-time and quick evaluation of ISMMMOs in marine ecosystems. This study may direct other studies about the automatic classification of marine ISMMMOs. We will be developing other novel nonparametric approaches to detect maritime life (e.g., different types of flying birds, sitting birds, big mammals (e.g., whales, dolphins), sharks, turtles, rays) automatically in large number of images in surveys

using supervised approaches (e.g., (Kuru et al., 2023)) to help evaluate the maritime industry and natural ecosystem together within well-prepared models. We aim to incorporate the built methodology with camera systems used in aeroplanes and unmanned aerial vehicles (UAVs) and to employ it in real-time rescue missions on high seas and open oceans as a future work, in particular, after aeroplane crashes and maritime accidents to find wreckages and survivals.

7. Limitations of the study

Complete white ISMMMOs as displayed in Appendix B (Fig. 18) have the same characteristics as waves ($R = 255$, $G = 255$, $B = 255$) with respect to HSV conversion, in particular, saturation. As it can be readily noticed in Table 1, zero is assigned to the hue and saturation during the conversion from RGB to HSV colour space when the values for the RGB colour space is 255, 255, 255 for the three channels. Our techniques perform successfully where the hue and saturation values are distinctive (i.e., $H > 0$ and $S > 0$) and therefore, these types of objects can not be detected using the approach built in this study.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix A. Pseudo codes of the methodology based on the Matlab syntax

Algorithm 2: Main methodology titled startSplittingObjectsUnsupervised: Phases of the operations to detect man-made objects in images.

Data: The target directory of a survey with images.
Result: Two directories, one of which is for images with man-made objects, and the other is for other images.

```
– >Variables;
infoDetail=''; steps = 1; k=1; posImageCount = 0; name = 'Human object detection';
set(handles.edtGeneral, 'String', ''); set(handles.edtDetail, 'String', ''); set(h,'Name','WilDetection: Splitting images using unsupervised technique'); set(h,
  'Position', [1090 50 355 50]) set(handles.edtGeneral, 'String', strcat(name,' is on process...')); set(handles.edtDetail, 'String', 'The information will be
  displayed here after the process above is completed.');
– >Create NEG and POS folder;
[parentFolder deepestFolder] = fileparts(imageDir); negSubFolder = sprintf('%s/NEG-%s', imageDir, deepestFolder); posSubFolder =
  sprintf('%s/POS-%s', imageDir, deepestFolder);
– > Create the folders if they do not exist.;
if exist(negSubFolder, 'dir') then
    mkdir(negSubFolder);
if exist(posSubFolder, 'dir') then
    mkdir(posSubFolder);
– > Progress bar to show the progress of the process;
h = waitbar(0.1,'Please wait...','CreateCancelBtn',... 'setappdata(gcbf,"canceling",1); delete(gcbf);')
– > Start splitting;
[Result,posImageCount] = startSplittingManMadeObjects(imageDir,posSubFolder,handles,h);
infoDetail = strcat(name,' has been completed; Total count of POS images is', ' ', num2str(posImageCount));
c_str1=[infoDetail];
set(handles.edtDetail, 'String', c_str); set(h,'Name','WilDetection: Splitting process of man-made object detection is COMPLETED');
waitbar(k / steps,h,sprintf('COMPLETED'));
ImDir = dir([imageDir, '.jpg']);
totNegImageCount = length(ImDir(not([ImDir.isdir])));
set(handles.edtNegCount, 'String', num2str(totNegImageCount));
set(handles.edtGeneral, 'String', ' All tasks have been completed');
result = 'Splitting images has been completed.';
```

Algorithm 3: Function titled startSplittingManMadeObjects:Main methodology: Phases of the operations to detect man-made objects in images.

```
Data: The target directory of a survey with images (imageDir)
Result: Two directories, one of which is for images with man-made objects, and the other is for other images.
- >Variables;
Rmin = 12.5; Rmax = 1000; curCount = 0; counter = 0; posImageCount = 0; hAdjust = 180; sAdjustPer = 0.25; files=dir(strcat(imageDir,'\*.jpg')); path = strcat(imageDir,'\'); steps = numel(files);
foreach k=1:numel(files) do
    file_name=files(k).name; image_name=strcat(path,file_name);
    - > Get positive image name and change it by adding _ at the end to move changed and not changed ones into same directory;
    [pathstr,name,ext] = fileparts(file_name); name = strcat(name,'_'); new_file_name = strcat(name,ext);
    Irgb=imread(image_name);
    figure; imshow(Irgb);
    index = k; index_image = k;
    newImage = HSVadjustManMade(Irgb, hAdjust,sAdjustPer);
    imshow(newImage);
    ImgR = newImage(:, :, 1); ImgG = newImage(:, :, 2); ImgB = newImage(:, :, 3); [M N] = size(ImgR);
    - >Get location of pure red pixels;
    Rmask = logical(zeros(M, N)); Rmask = ((ImgR < 0.25 & ImgG < 0.80 & ImgB > 0.35) & (ImgR < (ImgB) & ImgG < (ImgB)));
    - >Replace the pixels with some other colour;
    ImgR(Rmask) = 255; ImgG(Rmask) = 255; ImgB(Rmask) = 255; ImgR(Rmask==0) = 0; ImgG(Rmask==0) = 0; ImgB(Rmask==0) = 0;
    - >Combine into a new image;
    Img_new(:, :, 1) = ImgR; Img_new(:, :, 2) = ImgG; Img_new(:, :, 3) = ImgB;
    figure; imshow(Img_new);
    - >create indexed image from binary;
    BW = im2bw(Img_new,0.05);
    - >loop over ICE_threshold;
    ICE_threshold = 0.1; ICE_sigma = 2; img_edge = edge(BW, 'canny', ICE_threshold, ICE_sigma);
    - >create 3x3 array of 1s for dilate mask;
    SE = ones(3);
    - >dilate image to create closed boundary for birds with incomplete boundaries defined;
    img_dilate = imdilate(img_edge, SE); img_dilate2 = imdilate(img_dilate, SE);
    - >fill objects with closed boundaries;
    img_fill2 = imfill(img_dilate2,'holes');
    L = bwlabeln(img_fill2);
    - >get the centroid of each object to use as seeds for local neighbourhood definitions;
    stats = regionprops ('table', L, 'Area', 'BoundingBox', 'Centroid', 'MajorAxisLength','MinorAxisLength'); stats1 = regionprops (L, 'Area', 'BoundingBox', 'Centroid', 'MajorAxisLength','MinorAxisLength'); statsDetected = stats1;
    - >Get centers and radii of the circles;
    centers = stats.Centroid; centers diameters = mean([stats.MajorAxisLength stats.MinorAxisLength],2); radii = diameters/2; radiiDedected = diameters/2; centersDetected = centers;
    - >count the number of objects in the image;
    no_objects = size(stats, 1);
    if no_objects < 11 then
        Rmin = 7.5;
    else if no_objects > 10 && no_objects < 101 then
        Rmin = 12.5;
    else
        Rmin = 20.5;
    imshow(Irgb);
    for object = 1:no_objects do
        if (diameters(object) > Rmin*2) && (diameters(object) < Rmax*2) && (stats1(object).MajorAxisLength < Rmax*2) && (stats1(object).MinorAxisLength > Rmin*2) then
            - > Save image x and y coordinates into variable 'seed';
            posImageCount = posImageCount + 1; seed(object,1) = round(stats1(object).Centroid(1)); seed(object,2) = round(stats1(object).Centroid(2));
            - > Signify the object;
            hold on; plot(stats1(object).Centroid(1), stats1(object).Centroid(2), 'g+'); hold off;
    if (posImageCount > 0) then
        curCount = str2double(get(handles.edtPosCount,'String'))+1; set(handles.edtPosCount, 'String', num2str(curCount));
        - > Save the updated image with the signified objects;
        saveHighResolution(posSubFolder,new_file_name); - > remove the original image from processing folder;
        movefile(image_name,posSubFolder); delete(image_name)
    close All; clear All;
    - > Set the variable to 0 for next image;
    posImageCount = 0;
    - > Update the progress;
    waitbar(k / steps,h,'Man-made object detection is procesing...');
result = 'Man-made objects are tagged on the image';
```

Appendix B. Examples for objects not detected by the proposed approach

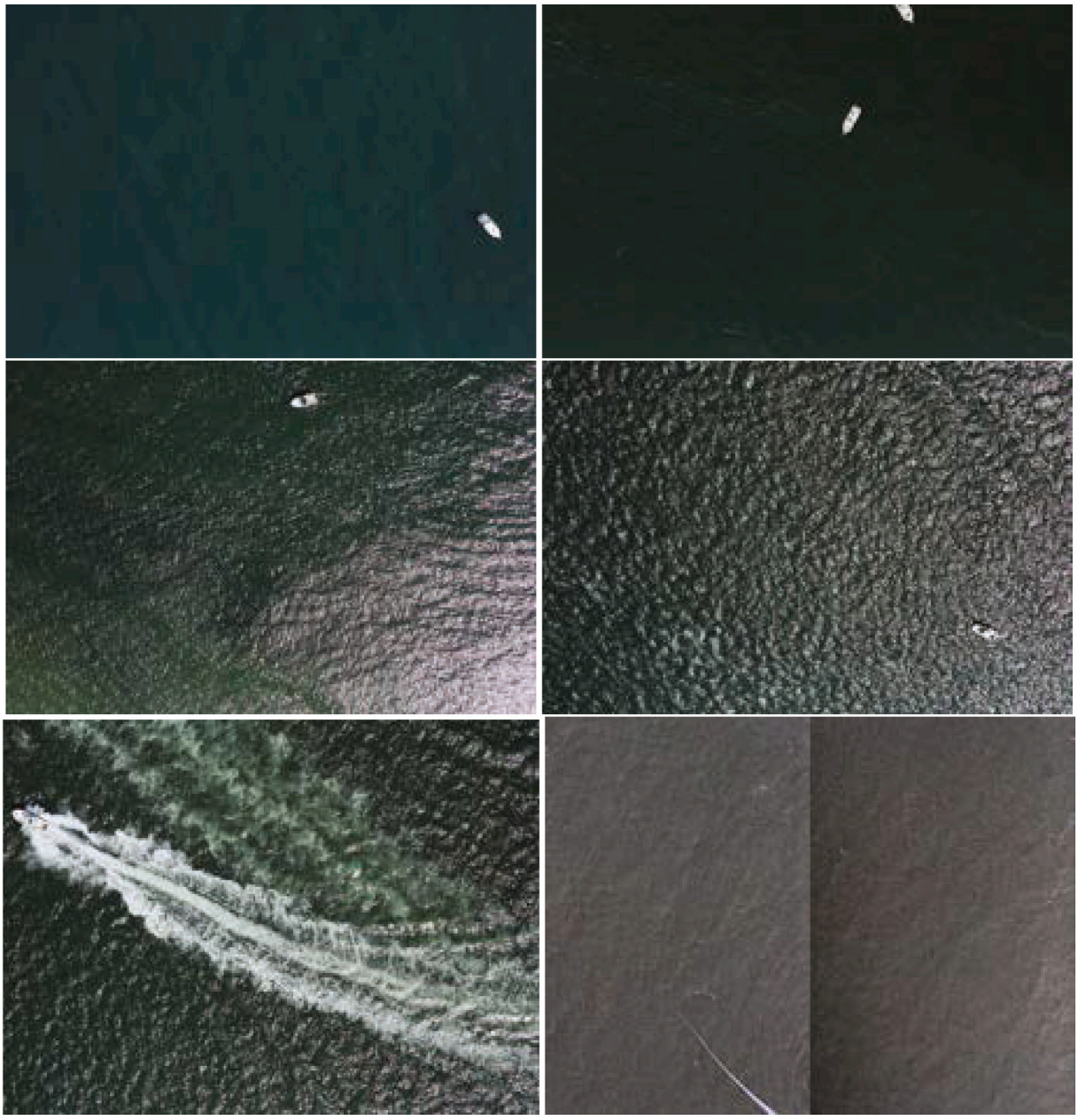


Fig. 18. Examples for objects not detected:

Appendix C. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecoinf.2023.102285.

References

Abu, A., Diamant, R., 2022. Feature set for classification of man-made underwater objects in optical and sas data. IEEE Sensors J. 22 (6), 6027–6041. http://dx.doi.org/10.1109/JSEN.2022.3148530.

Bibby, C., Jones, M., Marsden, S., 1998. Expedition Field Techniques: Bird Surveys. Royal Geographical Society, London.

Çalışkan, A., 2023. Detecting human activity types from 3d posture data using deep learning models. Biomed. Signal Process. Control 81, 104479. http://dx.doi.org/10.1016/j.bspc.2022.104479.

Clements, N., Robinson, W., 2022. A re-survey of winter bird communities in the Oregon coast range, USA, initially surveyed in 1968-1970. Biodiers. Data J. http://dx.doi.org/10.3897/arphapreprints.e91575. Aug.

Davis, K.L., Silverman, E.D., Sussman, A.L., Wilson, R.R., Zipkin, E.F., 2022. Errors in aerial survey count data: identifying pitfalls and solutions. Ecol. Evol. 12 (3), e8733 http://dx.doi.org/10.1002/ece3.8733.

Elrick-Barr, C.E., Zimmerhackel, J.S., Hill, G., Clifton, J., Ackermann, F., Burton, M., Harvey, E.S., 2022. Man-made structures in the marine environment: a review of stakeholders' social and economic values and perceptions. Environ. Sci. Pol. 129, 12–18. http://dx.doi.org/10.1016/j.envsci.2021.12.006.

Gibert, K., Horsburgh, J.S., Athanasiadis, I.N., Holmes, G., 2018. Environmental data science. Environ. Model Softw. 106, 4–12 special Issue on Environmental Data

Science. Applications to Air quality and Water cycle. https://doi.org/10.1016/j.envsoft.2018.04.005.
Graber, J., 2011. Land-Based Infrared Imagery for Marine Mammal Detection. Master's thesis,. University of Washington, USA.
Han, Z., Xing, J., Wang, X., Xue, F., Fan, J., 2022. A robust lcse-resnet for marine man-made target classification based on optical remote sensing imagery. Int. J. Artif. Intell. Tools 31 (06), 2240022. http://dx.doi.org/10.1142/S021821302240022X.
Kuru, K., 2014. Optimization and enhancement of h&e stained microscopical images by applying bilinear interpolation method on lab color mode. Theor. Biol. Med. Mod. 11 (1), 9. http://dx.doi.org/10.1186/1742-4682-11-9.
Kuru, K., Khan, W., 2018. Novel hybrid object-based non-parametric clustering approach for grouping similar objects in specific visual domains. Appl. Soft Comput. 62, 667–701. http://dx.doi.org/10.1016/j.asoc.2017.11.007.
Kuru, K., Yetgin, H., 2019. Transformation to advanced mechatronics systems within new industrial revolution: a novel framework in automation of everything (aoe). IEEE Access 7, 41395–41415. http://dx.doi.org/10.1109/ACCESS.2019.2907809.
Kuru, K., Girgin, S., Arda, K., Bozlar, U., 2013. A novel report generation approach for medical applications: the sisds methodology and its applications. Int. J. Med. Inform. 82 (5), 435–447. http://dx.doi.org/10.1016/j.ijmedinf.2012.05.019.
Kuru, K., Clough, S., Ansell, D., McCarthy, J., McGovern, S., 2023. Wildetect: an intelligent platform to perform airborne wildlife census automatically in the marine ecosystem using an ensemble of learning techniques and computer vision. Expert Syst. Appl. 231, 120574 http://dx.doi.org/10.1016/j.eswa.2023.120574.
Leira, F.S., Johansen, T.A., Fossen, T.I., 2015. Automatic detection, classification and tracking of objects in the ocean surface from uavs using a thermal camera. In: 2015 IEEE Aerospace Conference, pp. 1–10. http://dx.doi.org/10.1109/AERO.2015.7119238.
Loesdau, M., Chabrier, S., Gabillon, A., 2014. Hue and saturation in the RGB color space. In: Elmoataz, A., Lezoray, O., Nouboud, F., Mammass, D. (Eds.), Image and Signal Processing: 6th International Conference, ICISP 2014, Cherbourg, France, June 30–July 2, 2014. Proceedings. Springer International Publishing, Cham, pp. 203–212. http://dx.doi.org/10.1007/978-3-319-07998-1_23.
Lopez, J., Schoonmaker, J., Saggese, S., 2014. Automated detection of marine animals using multispectral imaging. In: 2014 Oceans - St. John's, pp. 1–6. http://dx.doi.org/10.1109/OCEANS.2014.7003132.
McIntosh, B.S., Alexandrov, G., Matthews, K., Mysiak, J., van Ittersum, M., 2011. Preface: thematic issue on the assessment and evaluation of environmental models and software. Environ. Model Softw. 26 (3), 245–246 thematic issue on the assessment and evaluation of environmental models and software. https://doi.org/10.1016/j.envsoft.2010.08.008.
Mehrnejad, M., Albu, A.B., Capson, D., Hoeberechts, M., 2013. Detection of stationary animals in deep-sea video. In: 2013 OCEANS - San Diego, pp. 1–5. http://dx.doi.org/10.23919/OCEANS.2013.6741095.
Noe, A., Pessoa, L., Thompson, E., 2000. Beyond the grand illusion: what change blindness really teaches us about vision. Vis. Cogn. 7 (2), 93–106.
Paleczny, M., Hammill, E., Karpouzi, V., Pauly, D., 2015. Population trend of the world's monitored seabirds, 1950-2010. PLoS One 10 (6), e0129342. http://dx.doi.org/10.1371/journal.pone.0129342.
Riser, S.C., Freeland, H.J., Roemmich, D., Wijffels, S., Troisi, A., Belbéoch, M., Gilbert, D., Xu, J., Pouliquen, S., Thresher, A., Traon, P.-Y.L., Maze, G., Klein, B., Ravichandran, M., Grant, F., Poulain, P.-M., Suga, T., Lim, B., Sterl, A., Sutton, P., Mork, K.-A., Velez-Belchí, P.J., Ansorge, I., King, B., Turton, J., Baringer, M., Jayne, S.R., 2016. Fifteen years of ocean observations with the global argo array. Nat. Clim. Chang. 6 (1), 145–153.
Rosebrock, A., 2017. Face alignment with opencv and python [cited 01.13.2022]. URL. http://www.pyimagesearch.com/2017/05/22/face-alignment-with-opencv-and-python/.
Rosenberg, K.V., Dokter, A.M., Blancher, P.J., Sauer, J.R., Smith, A.C., Smith, P.A., Stanton, J.C., Panjabi, A., Helft, L., Parr, M., Marra, P.P., 2019. Decline of the north american avifauna. Science 366 (6461), 120–124. http://dx.doi.org/10.1126/science.aaw1313.
Sànchez-Marrè, M., Cortés, U., Comas, J., 2004. Environmental sciences and artificial intelligence. Environ. Model Softw. 19 (9), 761–762 environmental Sciences and Artificial Intelligence. https://doi.org/10.1016/j.envsoft.2003.08.009.
Saur, G., Estable, S., Zielinski, K., Knabe, S., Teutsch, M., Gabel, M., 2011. Detection and classification of man-made offshore objects in terrasar-x and rapideye imagery: Selected results of the demarine-deko project. In: Oceans 2011 IEEE - Spain, pp. 1–10. http://dx.doi.org/10.1109/Oceans-Spain.2011.6003596.
Shi, Y., Shen, C., Fang, H., Li, H., 2017. Advanced control in marine mechatronic systems: a survey. IEEE/ASME Trans. Mech. 22 (3), 1121–1131. http://dx.doi.org/10.1109/TMECH.2017.2660528.
Xiao, Z., Zhang, H., Tong, H., Xu, X., 2022. An efficient temporal network with dual self-distillation for electroencephalography signal classification. In: 2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 1759–1762. http://dx.doi.org/10.1109/BIBM55620.2022.9995049.
Xing, H., Xiao, Z., Zhan, D., Luo, S., Dai, P., Li, K., 2022a. Selfmatch: robust semisupervised time-series classification with self-distillation. Int. J. Intell. Syst. 37 (11), 8583–8610. http://dx.doi.org/10.1002/int.22957.
Xing, H., Xiao, Z., Qu, R., Zhu, Z., Zhao, B., 2022b. An efficient federated distillation learning system for multitask time series classification. IEEE Trans. Instrum. Meas. 71, 1–12. http://dx.doi.org/10.1109/TIM.2022.3201203.
Zhang, T., Tian, B., Sengupta, D., Zhang, L., Si, Y., 2021. Global offshore wind turbine dataset. Sci. Data 8 (1), 191.