The potential of Jellytoring 2.0 smart tool as a global jellyfish monitoring platform

Ana Ruiz-Frau1 | Miguel Martin-Abadal2 | Charlotte L. Jennings2 | Yolanda Gonzalez-Cid2 | Hilmar Hinz1

1Institut Mediterrani d’Estudis Avançats, IMEDEA (CSIC-UIB), Esporles, Spain
2Department of Mathematics and Computer Science, Systems Robotics and Vision group (SRV), Universitat de les Illes Balears, Palma, Spain

Correspondence
Ana Ruiz-Frau, Institut Mediterrani d’Estudis Avançats, IMEDEA (CSIC-UIB), 07190 Esporles, Spain. Email: anaruiz@imedea.uib-csic.es

Funding information
Consellería d’Educació Cultura i Universitats (CCAA Illes Balears), Grant/Award Number: RyC 2013 14729; Direcció General de Política Universitària i Recerca, Govern Illes Balears, Grant/Award Number: PRD2018/34; European Union Erasmus+ Funding; H2020 Marie Skłodowska-Curie Actions, Grant/Award Number: 655475; Juan de la Cierva - Incorporación, Grant/Award Number: UJIC2019-040836; Ministerio de Ciencia e Innovación, Grant/Award Number: PID2020-115332RB-C33 and PLEC2021-007525; Ministerio de Economía y Competitividad, Grant/Award Number: DPI2017-86372-C3-3-R

Abstract
Despite the recent recognition of jellyfish as an important component of marine ecosystems and existing concerns on their potential population increase, they are rarely monitored at the appropriate spatial and temporal scales. Traditional jellyfish monitoring techniques are costly and generally restrict the spatial–temporal resolution limiting the quantity and quality of monitoring data. We introduce Jellytoring 2.0, an automatic recognition tool for jellyfish species based on convolutional neural networks (CNN). We trained Jellytoring 2.0 to identify 15 jellyfish species with a global distribution. Our aim is to offer Jellytoring 2.0 as an open-access tool to serve as the backbone for a system that promotes the creation of large-scale and long-term jellyfish monitoring data. Results reveal that Jellytoring 2.0 performed well in the identification of the 15 species with average precision values ranging between 90% and 99% for most of the species. Four of the species presented slightly lower values (75%–80%). Our system was trained on a relatively small dataset, implying that additional integration of image data will further improve the performance of the CNN. We show how the application of CNNs to image data can deliver a tool that will enable the cost-effective collection of jellyfish data on larger spatial and temporal scales. For Jellytoring 2.0 to become a truly global automatic identification system, we ask scientists and nonscientists to actively contribute with jellyfish image data to extend the number of species it can identify.

KEYWORDS
artificial intelligence AI, automatic detection

TAXONOMY CLASSIFICATION
Ecosystem ecology

1 INTRODUCTION

Recent research has highlighted the importance of jellyfish in marine ecosystems (Hays et al., 2018; Lamb et al., 2019; Milisenda et al., 2014). In addition, recently recorded increases in jellyfish have been linked to global change scenarios such as high fishing pressure (Richardson et al., 2009) and global warming (Brotz et al., 2012). However, the paucity of consistent long-term temporal and spatial data on jellyfish is such that there is considerable uncertainty over whether and why jellyfish populations might be increasing (Condon...
et al., 2012, 2013; Pitt et al., 2018; Sanz-Martin et al., 2016). The logistical difficulties in sampling and monitoring these organisms in time and space are mostly responsible for this deficient level of information, hindering our understanding of jellyfish population dynamics (Lamb et al., 2019).

Generally, jellyfish are monitored by visual observations from the shore and/or boats, with data logging approaches ranging from quantitative, to presence/absence and relative abundance indices (Condon et al., 2013). Data generated in this way generally tend to have a limited spatial and temporal resolution due to the high sampling effort and labor intensity involved and their associated cost. Furthermore, human observer data collection often requires special training and is susceptible to observer bias.

Here, we explore the application of convolutional neural networks (CNNs), a type of deep learning methodology, for the identification of different species of jellyfish. The successful application of this methodology could serve as a basis for the cost-effective acquisition of data in a continuous standardized way, with the possibility of moving toward a global scale methodology. Identification bias associated with the observer would be reduced, since networks are trained by a multitude of observers and a high volume of images which would level out any potential biases. Recent studies have successfully applied CNNs for the identification of different jellyfish species (Gauci et al., 2020; Han et al., 2021; Martin-Abadal et al., 2020). The majority focused on the local scale (Gauci et al., 2020; Martin-Abadal et al., 2020), while one study used jellyfish images from different geographical areas without a spatially explicit specification (Han et al., 2021). Yet, if the objective is to significantly advance and improve the quality and quantity of jellyfish data and spatial and temporal coverage, the first step is to develop a cost-effective method that allows for efficient data collection at large scales.

Here, we demonstrate and evaluate the use of the Jellytoring deep learning neural network as the basis for a global jellyfish identification system.

2 | MATERIALS AND METHODS

The present work extends on the technological advances previously achieved by the authors at setting the basis for an automatic jellyfish identification system. Jellytoring, a convolutional neural network, was developed and trained for the identification and quantification of three common jellyfish species in the Western Mediterranean area with successful results (Martin-Abadal et al., 2020). The tool, however, has the potential to be further expanded to identify a wide range of jellyfish species at a global scale.

2.1 | Convolutional neural networks

Convolutional neural networks (CNN) are a type of deep learning approach that allows the automatic extraction of information within an image. Neural networks are inspired on how the neurons in a mammalian brain respond to stimuli (Goodfellow et al., 2016). Training examples are given to the network so that it learns how to complete a task (Goodfellow et al., 2016). Here, we used underwater images containing different species of jellyfish to train a CNN to detect and classify them. Once the CNN has “learned,” it can detect and classify new jellyfish images, assigning confidence values to each detection. Finally, the species with the highest confidence value is associated with the detection.

2.2 | Image acquisition to train the CNN

Jellytoring was originally trained to identify three common Mediterranean species, namely Pelagia noctiluca, Cotylorhiza tuberculata, and Rhizostoma pulmo. To expand the capabilities of Jellytoring, hereafter Jellytoring 2.0, and make its application global, underwater video recordings from a range of jellyfish species were used to obtain additional images. The search platform YouTube was used to scan for publicly available jellyfish underwater videos. Searches were performed using the keyword “jellyfish” subsequently following a snowballing technique. A total of 226 videos corresponding to 15 jellyfish species were identified. A list of the jellyfish species and their respective number of videos are detailed in Table 1. Relevant videos were downloaded, and still images were extracted from each video. On average, a frame was extracted every 5 s. To log the presence of the different jellyfish species, an annotation file was generated using LabelImg (Tzutalin, 2018). For each of the frames, a bounding box around every jellyfish occurrence was drawn and was classified according to the jellyfish species. An .xml file was generated containing the position and classification of the occurring instance within the image. Overall, a total of 3808 occurrences were recorded corresponding to 15 jellyfish species (Table 1).

In an initial step, Jellytoring 2.0 was trained grouping all jellyfish species into one model. Performance however was not entirely satisfactory (Tables S1 and S2) as for similar-looking species (e.g., Rhizostoma pulmo vs. Rhizostoma luteum), the network incorrectly identified them in up to 30% of the instances. To reduce the number of misidentifications and to improve the network’s performance, a different approach was adopted and the CNN was trained to generate different models for different geographical areas.

WoRMS database (www.marinenspecies.org) was used to assess the spatial distribution of the jellyfish species, which were assigned to the different areas of the world where they had been recorded (North Pacific, South Pacific, North Atlantic, South Atlantic, Baltic Sea, Mediterranean Sea, Indian Ocean, South China Sea, Arctic Ocean, and Southern Ocean). To define the number of subareas, a multidimensional scale (MDS) plot was produced using a similarity matrix based on Euclidean distances on the presence/absence of the species for the different oceanic regions. Four distinct groupings emerged, namely the Atlantic/Mediterranean area, the Pacific area, the Arctic/Baltic area, and the Indian Ocean/South China Sea (Figure 1). Therefore, the CNN was trained to generate four different models for the four subareas. Each model was trained using data
from that specific area. This approach ensured that the network could not get confused between two similarly looking species, which had been uniquely recorded in different geographical areas.

### 2.3 Framework network selection

A variety of deep learning frameworks based on CNN can be used to extract instance information from images (Dai et al., 2016; Girshick, 2015; He et al., 2016; Lin et al., 2017). Here, our objective was to select a framework with the capability of detecting and classifying a range of jellyfish species, without the need to obtain pixel-wise segmentation of the detected instances, nor any extra feature that could slow down the process. Therefore, considering the requirements of the present application, the Faster R-CNN framework was selected. Considering the slow movement of jellyfish, an architecture with high detection performance was deemed suitable. We selected the Faster R-CNN-based implementation of the Inception ResNet v2 architecture (Szegedy et al., 2016). Full details of the framework and network selection are detailed in Martin-Abadal et al. (2020). Inception ResNet v2 is a very deep CNN with over 450 layers that can efficiently learn to identify objects within images, outputting instance bounding boxes and classifying them into one of the specified classes with a confidence percentage associated with it.

### 2.4 Jellytoring workflow

As a first step, a set of images containing jellyfish is used as input for a frozen version of a deep object detection neural network-trained model. During inference, the network starts the process of jellyfish...
detection. Subsequently, detection is optimized by a nonmaxima suppression (nms) algorithm (Neubeck & Van Gool, 2006). The final predictions are obtained by deleting instances with an associated confidence lower than a selected threshold value (Cthr). Figure 2 shows a representation of identification bounding boxes for several of the jellyfish species.

2.5 | Evaluation metrics

To evaluate the performance of a trained model, the bounding-box predictions generated as a result of the test dataset are classified as either a true positive, TP; or a false positive, FP. To do so, we use the Intersection over Union (IoU) metric. IoU is defined as the area of the intersection between a predicted and a ground-truth bounding box, divided by the union area of these bounding boxes (Equation 1).

\[
\text{IoU} = \frac{A_{\text{intersection}}}{A_{\text{union}}} \quad (1)
\]

A prediction is classified as TP if the IoU value with any ground-truth bounding box is greater than a threshold IoU and the predicted species class (i.e., species of interest) matches the class specified in the ground-truth box. When these conditions are not met, the prediction is classified as a FP (Equation 2). Ground-truth instances that do not have an IoU greater than the threshold IoU with any prediction are considered as undetected instances (false negatives, FN). Following the criteria applied in the PASCAL VOC challenge (Everingham et al., 2010), the IoU threshold was set at 0.5 (thrIoU = 0.5).

\[
\text{Prediction} = \begin{cases} 
\text{TP}, & \text{if IoU} \geq \text{thr}_{\text{IoU}} \& \text{Class}_{\text{pred}} = \text{Class}_{\text{gt}} \\
\text{FP}, & \text{otherwise} 
\end{cases} \quad (2)
\]

Once the number of TP, FP and FN is obtained, the average precision (AP) is calculated (Zhu, 2004); AP is one the most frequently used metrics in object detection applications. This metric is defined as the area under the max (precision)-recall curve. Once the AP is obtained for each class, a mean average precision (mAP) for all classes is computed. In addition, the confidence score generated by the network for each prediction can be used to fine-tune the network. The confidence score is used to find a confidence value that, when deleting predictions with low confidence values, more FP than TP are deleted thus, improving the detection performance of the network. To do this, a threshold sweep on the confidence prediction from 0% to 100% was performed in 1% steps (C_thr). For each step, the predictions with an associated confidence level lower than the C_thr were removed; and the Precision, Recall, and F1-score metrics were calculated from the TP, FP, and FN values.

Precision is calculated as the percentage between correctly identified instances and all the predictions made by the model (Equation 3). Recall refers to the percentage of TP detections with respect to all jellyfish instances present on the test data (Equation 4). Finally, the F1-score combines the Precision and Recall metrics (Equation 5).

\[
\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{F}} \times 100 \quad (3)
\]

\[
\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100 \quad (4)
\]

\[
\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}
\]
Furthermore, a multiclass confusion matrix is generated using all classes TP and FP counts. A multiclass confusion matrix is a \( N \times N \) matrix, where \( N \) is the number of studied classes. In our case, each row of the matrix corresponds to a jellyfish species and indicates, in the matrix columns, the percentage of the detections made by the neural network that correspond to each species of jellyfish.

For each of the four models corresponding to the different geographical areas, a 5-k-fold validation was performed (Geisser, 1975). Through this method, the dataset is split into five equally sized subsets and the network is trained five times, each time using a different subset as the test data (20% of the dataset) and the remaining four subsets as training data (80% of the dataset). The final results (i.e., AP, mAP, C_thr, Recall, Precision, and F1-score) are thus an average of the five trainings. This method reduces the variability of the results, making them less dependent on the selected test and training sets and therefore obtaining a more accurate performance estimation.

3 | RESULTS

For each one of the four geographical regions, the AP, mAP, Cthr, and F1-score metrics are presented, along with a multiclass confusion matrix. The confusion matrix allows us to identify those pairs of classes where the network has difficulties in telling them apart.

TABLE 2 Performance metrics for the Atlantic/Mediterranean regional model

| Species       | AP   | mAP  | C_THR | REC  | PREC  | F1-SCORE |
|---------------|------|------|-------|------|-------|----------|
| A. aurita     | 81.2%|      |       |      |       |          |
| C. branchi    | 77.3%|      |       |      |       |          |
| C. hysoscella | 79.6%|      |       |      |       |          |
| C. quinquecirra | 91.3%| | | | | |
| C. tuberculata | 91.1%| | | | | |
| C. capillata  | 94.7%| 87.7%| 67.4%| 80.1%| 88.0%| 83.9%    |
| C. lamarckii  | 93.5%| | | | | |
| P. noctiluca  | 74.6%| | | | | |
| R. luteum     | 79.8%| | | | | |
| R. pulmo      | 96.9%| | | | | |
| S. meleagris  | 93.1%| | | | | |
| T. ohboya     | 98.9%| | | | | |

Abbreviations: AP, average precision; C_THR, confidence threshold; mAP, mean average precision; PREC, precision; REC, recall.

3.1 | Atlantic–Mediterranean region

The spatial distribution assessment indicated that 12 of the 15 jellyfish species had been recorded in the Atlantic/Mediterranean area. The AP values for approximately 60% of those species ranged between 90% and 99%. The rest of the species, that is, A. aurita, C. branchi, R. luteum, and P. noctiluca presented values ranging between 75% and 80%. The mAP for all classes was 87.4%. After performing the confidence threshold sweep, an F1-score of 83.9% was obtained when applying a confidence threshold of 67.4% (Table 2). The multiclass confusion matrix indicates that the network does not tend to get confused between species. The exception was the capacity of the network to distinguish between R. pulmo and R. luteum as the network confused them in 17.9% of the cases. Two main reasons are attributable to this weakness: on the one hand, both species belong to the same genus and are very similarly looking (Figure 3). On the other hand, the training data for R. pulmo were thrice that of R. luteum, biasing the network toward R. pulmo (Table 3).

3.2 | Pacific region

Seven of the 15 jellyfish species had been detected in the Pacific area. The network model for this region had a high performance as the AP values for the majority of the species lay between 90% and 99%, while only two species (P. noctiluca and A. aurita) had values

\[
F1 \text{- score} = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \times 100
\]

FIGURE 3 Comparison between the R. luteum (left) and R. pulmo (right) jellyfish species.
comprised between 78% and 83%. The mAP had a value of 90.1% and a F1-score of 83.5% after applying an optimal confidence threshold of 44.8% (Table 4). The multiclass confusion matrix indicates that the network performed very well in distinguishing between species, as all values but one were above 98% with the exception of N. nomurai which was correctly identified in 96.7% of the cases. Only occasionally (3.3% of the cases) was misidentified as C. capillata (Table 5).

### 3.3 | Arctic–Baltic region

Four of the 15 species were recorded in the Arctic/Baltic region. All AP values for this region were above 80%. The values for C. capillata and C. lamarckii were both over 97%, while the AP values for A. aurita and C. fuscens were in the lower eighties (81.5% and 83.2%). The average AP between all the species was slightly over 90%. After applying the optimal confidence threshold of 55.4%, the F1-score achieved a value of 82.2% (Table 6). The performance of the network model was very high as it was able to distinguish between the different species. Even very similarly looking species belonging to the same genus (i.e., C. capillata and C. lamarckii) were correctly identified in 98% of the cases (Table 7).

### 3.4 | Indian Ocean–South China Sea Region

The model for the Indian Ocean and the South China Sea was developed with the four jellyfish species recorded in the region. The AP values for two of the species, N. nomurai and C. quinquecirrha, were very high, ranging between 96.7% and 97.5%. AP values for A. aurita and P. noctiluca were slightly lower, 82.4% and 79.2%, respectively. The mAP had a value of 89%, while the F1-score was 80.8% after applying a 64.8% confidence threshold (Table 8). As with the other models, the network had a high performance in distinguishing between species, as all values were above 99.5% (Table 9).

Overall, the performance metrics are good and consistent over the four regions. Furthermore, the network does not tend to mistake a jellyfish species for a wrong one, demonstrating the effectiveness of the implemented strategy of dividing the species into four regions.

### 4 | DISCUSSION

Thus far, jellyfish have proven difficult to monitor, due to the high associated cost of logistics and the need of humans to analyze visual data (direct observations or video recordings), limiting both the spatial and temporal scales of monitoring programs (but see Condon et al., 2015). The results of this paper demonstrate that the application of CNNs to image data has the potential to deliver a tool that enables the cost-effective collection of jellyfish data on larger spatial and temporal scales. In our initial publication, we described the successful automatic classification and quantification of three jellyfish species common in the Mediterranean (Martin-Abadal et al., 2020).
The results were highly promising, but the question remained of whether the CNN would be able to distinguish a much larger number of species and thus serve as a backbone for the development of a global tool. Here, we demonstrate that the algorithms developed are robust and that most species can be distinguished by the system with high precision and an acceptable level of confusion. The tool
in its original unmodified version showed weaknesses when applied on a larger number of species and a restructuring of the approach architecture was necessary to achieve acceptable levels of error in the classifications. Four separate world subareas models were developed to avoid confusion between similarly looking species, which did not overlap in their geographical distributions. Through this modification, errors in misidentifications were reduced; nevertheless, some species confusions persisted. We note that in our attempt to construct a global CNN, we expected the performance of the CNN to improve with increasing number of images. However, when increasing image numbers for the global CNN, precision did not increase as would be expected. While the overall number of images was higher, images for each individual species might not have been sufficient. When the CNN precision does not improve sufficiently with image number to run at such a scale, the networks can be divided into submodels, corresponding to different geographical areas. This was our approach. The development of the CNN as a single model would be ideal but is only justified when increasing the number of images improves the CNN. We expect the performance of the CNN to improve with increasing numbers of images for each species. In the event that the performance of the CNN does not improve sufficiently to run as a single model with worldwide application, the network can still be left divided into the existing (or modified) submodels, corresponding to different geographical areas. The development of the CNN as a single model would be the ideal objective but the aim is to have a system that is able to facilitate the identification of jellyfish species, which is still achieved through the CNN working as separate submodels.

Within the Atlantic–Mediterranean subarea model, the most notable case was for *Rhizostoma pulmo* and *Rhizostoma luteum* as the network confused them in 17.9% of the cases. This may be related to the very similar morphology of the two species or to the fact that for *R. pulmo* we had thrice the amount of training data available, making the network more biased toward this species. In the other subarea models, confusion percentages were much lower (max. 3.3%). Overall, the average precision values (APs) for most of the species varied between 90% and 99%. However, there were some species, such as *A. aurita*, *C. branchi*, *R. luteum*, or *P. noctiluca*, that showed lower AP values (75%-80%). These lower AP values may be related to multiple factors, ranging from insufficient training data to high similarities in body morphology between different species. For instance, *P. noctiluca* is a translucent organism whose body can adopt significantly different configurations, due to the movement of their tentacles; or in the case of *A. aurita*, it tends to appear in big blooms, causing overlapping in the images and making it hard to identify single organisms. While the classification weaknesses for some species may be improved with the integration of additional training data, for others there may be morphological or behavioral aspects that will make it difficult to increase the overall classification precision. Nevertheless, in general, the incorporation of additional training data is expected to significantly improve the detection and correct classification of objects (Goodfellow et al., 2016).

The comparison of the mean average precision values of Jellytoring 2.0 with other studies dealing with a similar number of species (i.e., 10), shows that Jellytoring 2.0 outperforms previous attempts. While the mAP Jellytoring 2.0 ranged from 88% to 90%, Han et al. (2021) showed an mAP of 75%. By contrast, when compared to studies that automatically identified a lower number of species (i.e., 5), Jellytoring 2.0 performance was lower, Gaucci et al. (2020) mAP = 96%. However, algorithms trained to identify a lower number of jellyfish species tend to perform better, as shown with the original version of Jellytoring (mAP = 95%).

With this paper, we want to offer the scientific and nonscientific community the basis for an automatic jellyfish identification system that could form the first step toward the creation of large-scale and long-term monitoring datasets, as well as real-time monitoring tools. Although not included as part of this exercise due to insufficient image data for each of the species considered here, Jellytoring is capable of performing real-time quantification of the identified jellyfish species (Martin-Abadal et al., 2020). This feature allows for the development of real-time monitoring tools in coastal management scenarios. The rapid development of cheap hardware video-capturing devices and software processing units allows for the integration of CNNs to generate in situ identification and abundance data. Data captured and generated through these assemblages can subsequently be sent via communication systems to land-based hubs where data can be further processed and visualized in near real-time (Figure 4). Such monitoring systems can have a wide range of applications as early warning systems for commercial operation and coastal management and can also provide invaluable data for science helping to close some of the remaining gaps in jellyfish ecology.

It is with this in mind and to extend the scope of the Jellytoring 2.0 tool that we ask the scientific community to actively participate in the provision of image data for both new species not yet included in the application as well as for species already covered. To facilitate the provisioning of this data we have created a web application (jellytoring.uib.es, under development) that will allow uploading photos and video files, as well as the positioning of bounding boxes and classification labels. For the effective and accurate performance of the tool, large numbers of training images are required, especially if the

| Species          | *A. aurita* | *C. quinquecirrh* | *N. nomurai* | *P. noctiluca* |
|------------------|-------------|------------------|--------------|---------------|
| *A. aurita*      | 99.9%       | -                | 0.1%         | -             |
| *C. quinquecirrh*| -           | 100.0%           | -            | -             |
| *N. nomurai*     | -           | -                | 100.0%       | -             |
| *P. noctiluca*   | -           | 0.5%             | -            | 99.5%         |

TABLE 9 Multiclass confusion matrix for the species recorded in the Indian Ocean/South China Sea region.
tool is to become a global application. While currently our website only accepts the uploading of still images, still images from videos can be easily extracted by the user using dedicated free software such as MPEG. This type of software is capable of rapidly converting entire video clips to images, or to extract still frames in a time-frequency set by the user. The results provided by Jellytoring 2.0 will be in form of a data file that will show the time and video frame an individual object was detected, its classification by the network and a confidence and confusion indicator of that classification. From this raw data, the user will be able to easily calculate various derived parameters such as overall abundances within the video footage, abundances per time interval or diversity. The results of this study have been very encouraging but only with the involvement of the wider scientific community and interested citizens can this application grow into a global application.

**AUTHOR CONTRIBUTIONS**

**Ana Ruiz-Frau:** Conceptualization (lead); data curation (lead); formal analysis (lead); funding acquisition (lead); investigation (lead); methodology (equal); project administration (lead); resources (lead); software (supporting); supervision (supporting); validation (supporting); visualization (supporting); writing – original draft (supporting); writing – review and editing (supporting).

**Miguel Martin-Abadal:** Conceptualization (supporting); data curation (lead); formal analysis (lead); funding acquisition (supporting); investigation (supporting); methodology (lead); project administration (supporting); resources (supporting); software (lead); supervision (equal); validation (equal); visualization (lead); writing – original draft (equal); writing – review and editing (equal).

**Charlotte Louise Jennings:** Conceptualization (supporting); data curation (equal); formal analysis (equal); funding acquisition (supporting); investigation (supporting); methodology (supporting); project administration (supporting); software (supporting); supervision (supporting); validation (supporting); visualization (supporting); writing – original draft (supporting); writing – review and editing (supporting).

**Yolanda Gonzalez-Cid:** Conceptualization (supporting); data curation (supporting); formal analysis (lead); funding acquisition (supporting); investigation (equal); methodology (equal); project administration (supporting); resources (equal); software (equal); supervision (equal); validation (equal); visualization (equal); writing – original draft (equal); writing – review and editing (equal).

**Hilmar Hinz:** Conceptualization (lead); data curation (equal); formal analysis (equal); funding acquisition (supporting); investigation (equal); methodology (equal); project administration (supporting); resources (supporting); software (supporting); supervision (equal); validation (equal); visualization (equal); writing – original draft (lead); writing – review and editing (lead).

**ACKNOWLEDGMENTS**

During the conception, implementation and writing of the study A.R.F were supported by a Marie-Sklodowska-Curie Individual Fellowship (JellyPacts project number 655475) and a Juan de la Cierva-Incorporación (IJC2019-040836-I/AEI/10.13039/501100011033) from the Spanish Government. M.M.A. was supported by Ministry of Economy and Competitiveness (AEI, FEDER, UE), under contract DPI2017-86372-C3-3-R. C.L.J. was supported by the EU funded Erasmus+ program. Y.G.C. was supported by Ministry of Economy and Competitiveness (AEI, FEDER, UE) under contract DPI2017-86372-C3-3-R; Ministry of Science and Innovation (AEI, FEDER, UE) under contracts PID2020-115332RB-C33 and PLEC2021-007525 and by the Comunitat Autonoma de les Illes Balears through the Direcció General de Poltica Universitaria i Recerca with funds from the Tourist Stay Tax Law (PRD2018/34). H.H. was partially supported through a Ramón y Cajal Fellowship financed by the Ministerio
de Economía y Competitividad de España and the Conselleria d’Educació, Cultura i Universitats Comunitat Autónoma de las Islas Baleares (RyC 2013 14729). The authors would like to thank Ruben Bustos Romero for his help developing the Jellytoring 2.0 web application platform and Aenne Douwes for her contributions during the early stages of the project. We also thank the contributions of two anonymous reviewers that contributed to achieving a clearer manuscript.

DATA AVAILABILITY STATEMENT
The dataset and code used is deposited in a GitHub repository https://github.com/srv/jf_object_detection/tree/regions.

ORCID
Ana Ruiz-Frau https://orcid.org/0000-0002-1317-2827

REFERENCES
Brotz, L., Cheung, W. W. L., Kleisner, K., Pakhomov, E., & Pauly, D. (2012). Increasing jellyfish populations: Trends in large marine ecosystems. Hydrobiologia, 690, 3–20. https://doi.org/10.1007/s10750-012-1039-7
Condon, R. H., Duarte, C. M., Pitt, K. A., Robinson, K. L., Lucas, C. H., Sutherland, K. R., Mianzan, H. W., Bogeberg, M., Purcell, J. E., Decker, M. B., Uye, S.-I., Madin, L. P., Brodeur, R. D., Haddock, S. H. D., Malej, A., Parry, G. D., Eriksen, E., Quinones, J., Acha, M., ... Graham, W. M. (2013). Recurrent jellyfish blooms are a consequence of global oscillations. Proceedings of the National Academy of Sciences of the United States of America, 110(3), 1000–1005.
Condon, R. H., Graham, W. M., Duarte, C. M., Pitt, K. A., Lucas, C. H., Haddock, S. H. D., Sutherland, K. R., Robinson, K. L., Dawson, M. N., Decker, M. B., Mills, C. E., Purcell, J. E., Malej, A., Mianzan, H., Uye, S.-I., Gelchich, S., & Madin, L. P. (2012). Questioning the rise of gelatinous zooplankton in the world’s oceans. Bioscience, 62(2), 160–169.
Condon, R. H., Lucas, C. H., Duarte, C. M., Pitt, K. A., Haddock, S. H. D., Madin, L. P., Brodeur, R. D., Sutherland, K. R., Mianzan, H. W., Purcell, J. E., Decker, M. B., Uye, S.-I., Malej, A., Bogeberg, M., Everett, J. T., Gibbons, M., Gonzalez, H., Hay, S., Hensche, N., ... Regetz, J. (2015). Jedi: Jellyfish database initiative. Biological and Chemical Oceanography Data Management Office (BCO-DMO), 3. https://doi.org/10.1575/1912/7119
Dai, J., Li, Y., He, K., & Sun, J. (2016). R-fcn: Object detection via region-based fully convolutional networks. NIPS.
Everingham, M., Van Gool, L., Williams, C. K. I., Winn, J., & Zisserman, A. (2010). The PASCAL Visual Object Classes (VOC) challenge. International Journal of Computer Vision, 88(2), 303–338.
Gauci, A., Deidun, A., & Abela, J. (2020). Automating jellyfish species recognition through faster region-based convolution neural networks. Applied Sciences, 10, 8257. https://doi.org/10.3390/app10228257
Geisser, S. (1975). The predictive sample reuse method with applications. Journal of the American Statistical Association, 70(350), 320–328.
Girshick, R. (2015). Fast R-CNN. 2015 IEEE International Conference on Computer Vision (ICCV) (pp. 1440–1448).
Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.
Han, Y., Chang, Q., Ding, S., Gao, M., Zhang, B., & Li, S. (2021). Research on multiple jellyfish classification and detection based on deep learning. Multimedia Tools and Applications, 81, 19429–19444. https://doi.org/10.1007/s11042-021-11307-y
Hays, G. C., Doyle, T. K., & Houghton, J. D. (2018). A paradigm shift in the trophic importance of jellyfish? Trends in Ecology and Evolution, 33, 874–884. https://doi.org/10.1016/j.tree.2018.09.001
He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 770–778).
Lamb, P. D., Hunter, E., Pinnegar, J. K., Doyle, T. K., Creer, S., Taylor, M. I., & Coll, M. (2019). Inclusion of jellyfish in 30+ years of Ecopath with Ecosim models. ICES Journal of Marine Science, 76, 1941–1950. https://doi.org/10.1093/icesjms/fsz165
Lin, T., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2017). Feature pyramid networks for object detection. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 936–944). https://doi.org/10.1109/CVPR.2017.106
Martin-Abadal, M., Ruiz-Frau, A., Hinz, H., & Gonzalez-Cid, Y. (2020). Jellytoring: Real-time jellyfish monitoring based on deep learning object detection. Sensors, 20, 1708. https://doi.org/10.3390/s20061708
Milisenda, G., Rosa, S., Fuentes, V. L., Boero, F., Guglielmo, L., Purcell, J. E., & Piraino, S. (2014). Jellyfish as prey: Frequency of predation and selective foraging of Boops boops (vertebrata, actinopterygii) on the mauve stinger Pelagia noctiluca (cnidaria, scyphozoa). PLoS One, 9, e94600. https://doi.org/10.1371/journal.pone.0094600
Neubeck, A., & Van Gool, L. (2006). Efficient non-maximum suppression. Proceedings of International Conference on Pattern Recognition, 3, 850–855.
Pitt, K. A., Lucas, C. H., Condon, R. H., Duarte, C. M., & Stewart-Koster, B. (2018). Claims that anthropogenic stressors facilitate jellyfish blooms have been amplified beyond the available evidence: A systematic review. Frontiers in Marine Science, 5, 451. https://doi.org/10.3389/fmars.2018.000451
Richardson, A. J., Bakun, A., Hays, G. C., & Gibbons, M. J. (2009). The jellyfish joyride: Causes, consequences and management responses to a more gelatinous future. Trends in Ecology & Evolution, 24(6), 312–322.
Sanz-Martin, M., Pitt, K. A., Condon, R. H., Lucas, C. H., de Santana, C. N., & Duarte, C. M. (2016). Flawed citation practices facilitate the unsubstantiated perception of a global trend toward increased jellyfish blooms. Global Ecology and Biogeography, 25, 1039–1049. https://doi.org/10.1111/geb.12474
Szegedy, C., Ioffe, S., & Vanhoucke, V. (2016). Inception-v4, inception-resnet and the impact of residual connections on learning. AAAI Conference on Artificial Intelligence.
Tutzalin, D. (2018). Labelimg. https://github.com/tutzalin/labelimg
Zhu, M. (2004). Recall, precision and average precision. Technical report, Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Canada.

SUPPORTING INFORMATION
Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Ruiz-Frau, A., Martin-Abadal, M., Jennings, C. L., Gonzalez-Cid, Y., & Hinz, H. (2022). The potential of Jellytoring 2.0 smart tool as a global jellyfish monitoring platform. Ecology and Evolution, 12, e9472. https://doi.org/10.1002/ece3.9472