

Article

Automating Jellyfish Species Recognition through Faster Region-Based Convolution Neural Networks

Adam Gauci ^{1,*}, Alan Deidun ¹ and John Abela ²

¹ Department of Geosciences, Faculty of Science, University of Malta, Msida MSD 2080, Malta; alan.deidun@um.edu.mt

² Department of Computer Information Systems, Faculty of ICT, University of Malta, Msida MSD 2080, Malta; john.abela@um.edu.mt

* Correspondence: adam.gauci@um.edu.mt

Received: 15 October 2020; Accepted: 17 November 2020; Published: 21 November 2020



Abstract: In recent years, citizen science campaigns have provided a very good platform for widespread data collection. Within the marine domain, jellyfish are among the most commonly deployed species for citizen reporting purposes. The timely validation of submitted jellyfish reports remains challenging, given the sheer volume of reports being submitted and the relative paucity of trained staff familiar with the taxonomic identification of jellyfish. In this work, hundreds of photos that were submitted to the “Spot the Jellyfish” initiative are used to train a group of region-based, convolution neural networks. The main aim is to develop models that can classify, and distinguish between, the five most commonly recorded species of jellyfish within Maltese waters. In particular, images of the *Pelagia noctiluca*, *Cotylorhiza tuberculata*, *Carybdea marsupialis*, *Verella vellella* and salps were considered. The reliability of the digital architecture is quantified through the precision, recall, f_1 score, and κ score metrics. Improvements gained through the applicability of data augmentation and transfer learning techniques, are also discussed. Very promising results, that support upcoming aspirations to embed automated classification methods within online services, including smart phone apps, were obtained. These can reduce, and potentially eliminate, the need for human expert intervention in validating citizen science reports for the five jellyfish species in question, thus providing prompt feedback to the citizen scientist submitting the report.

Keywords: citizen science; convolution neural networks; image processing; jellyfish detection; machine learning

1. Introduction

Citizen science incorporates scientific research and monitoring projects for which members of the public collect, categorise, transcribe or analyse scientific data. Citizen science has definitely come of age in recent years in generating knowledge, creating new learning opportunities and enabling civic participation [1]. The marine conservation and biodiversity monitoring potential of citizen science campaigns is increasingly being recognised [2,3], being deployed, for example, in the monitoring of invasions by marine alien species [4] and in an array of other marine applications summarised in [5]. Despite an increased uptake of marine monitoring protocols based on the citizen science rationale, a number of challenges still undermine the mainstreaming of the same protocols. Chief among these challenges is the need to validate individually submitted citizen science reports in order to avoid the contamination of the main dataset with misidentification reports. In fact, for citizen science projects to fully fulfil their potential, they need to generate good-quality datasets [6].

A decade ago, smart phone apps and the virtual/gaming approach were considered as cutting-edge, emerging technologies which held great promise for the advancement of citizen

science [7]. The state-of-the-art has progressed to a state where smart phone apps are being enhanced with a new suite of sensors to enable greater citizen science functionality of the same technologies. Given the image-intensive nature of many ongoing marine citizen science campaigns, the image analyses domains of Artificial Intelligence and Machine Learning have increasingly been applied to automate the validation of the same images [8] and thus to expedite the normally lengthy process. These protocols have been incorporated within fixed-location, jellyfish bloom early-warning systems, including the JellyMonitor prototype deployed on the Scottish seabed and which makes use of both classic computer vision and deep learning neural networks to detect the occurrence of jellyfish blooms [9]. The automated processing of underwater footage, besides static imagery, has also been achieved through the same protocols. For instance, the JellyToring system automatically detects and quantifies different species of jellyfish based on a deep object detection neural network, allowing the user to automatically record jellyfish presence during long periods of time [10]. To date, our University of Malta research group has proposed an image-analysis protocol for the characterisation of Large Microplastics (LMPs) extracted from beach sediment [11].

Jellyfish swarms, or blooms, consist of large numbers of jellyfish which aggregate occasionally, mainly as a result of the direction of prevailing water currents, and the occurrence of such jellyfish aggregations has attracted increased attention in recent years (see, e.g., in [12]) by virtue of their socio-economic impact. The “Spot the Jellyfish” initiative, run by the University of Malta and the International Ocean Institute (IOI) since June 2010, follows a citizen science approach and relies on the collaboration of the general public, mariners, divers and especially the younger generations through their teachers and parents, by recruiting their assistance in recording the presence and location of different jellyfish within Maltese waters. The reporting is done by simply matching a sighted jellyfish with a simple visual identification guide, giving the date and time of the sighting, and indicating the number of jellyfish individuals observed. Sightings can also be reported online, through the campaign’s social media page, via SMS or email. Citizen scientists are also encouraged to supplement their report with a photo, which is used for verification purposes and to quality control the database. All jellyfish reports within this campaign are also shared with Maltese tourism authorities for beach management purposes. Current citizen science validation protocols are labour-intensive and subject to potential human errors, such that citizen scientists do not receive prompt feedback to their submitted reports, leading to a possible disengagement of the same contributors. On the pro side, current protocols are characterized by a high degree of taxonomic precision, such that the rate of misidentification is generally very low.

In this work, the hundreds of photos submitted to the “Spot the Jellyfish” initiative for the five most commonly recorded species within Maltese waters (*Pelagia noctiluca*, *Cotylorhiza tuberculata*, *Carybdea marsupialis*, *Veleva veleva* and salps) are used to train and evaluate the performance of a tailored-designed, region-based Convolution Neural Network (R-CNN). The main aim of the developed digital architecture is to identify correctly the jellyfish species recorded in the submitted photo, thus eliminating or reducing the need for human expert intervention and providing a prompt feedback to the citizen scientist. Semiquantitative estimates of jellyfish individual abundance will also be generated by the algorithm when presented with photos of jellyfish blooms. A region-based method is applied to identify and record the jellyfish sighting and the confidence level for the model’s taxonomic identification output. Over the last decade, R-CNNs have been researched and successfully applied to a wide-range of applications including the detection of tumours [13], the detection of ships [14], the detection of debris on the airfield pavement [15], to give an early warning of sharks in coastal areas through the use of drones [16], for traffic sign recognition [17], to detect three-dimensional objects during autonomous driving [18], as well as for the classification of cracks in concrete [19].

The most basic form of R-CNN makes use of an external algorithm such as the Edge Boxes method to come up with region proposals [20]. Pixels corresponding to the identified regions are then extracted, resized and processed via a support vector machine (SVM) that is trained using CNN features [21]. The Fast R-CNN implementation makes use of the Edge Box algorithm to identify regions

that might contain an object of interest [22]. Although the CNN features of the identified regions are extracted, the results coming from overlapping parts are shared. While this makes the process faster, the entire images still need to be processed. In the Faster R-CNN method, a Region Proposal Network (RPN) is embedded within the detector to eliminate the need of an external algorithm to identify important sub-areas [23]. Apart from making the implementation faster, this allows the detection of regions to be attuned to the dataset itself. RPNs are able to identify valid boxes of any size quickly, by making use of anchor boxes. In particular, processing is carried out within a set of predefined boxes, with dimensions that correspond to the objects that are to be identified. This eliminates the need to perform computations on every possible position of a sliding window. In this work, the anchor boxes allow the definition of the expected aspect ratios of image regions that show a jellyfish. The classification models then try to fit objects within the given ratios. This approach allows for a better performance to be obtained when multiple jellyfish cover each other, such as in the case of a bloom. Faster R-CNNs are designed to give quick and accurate results, even when a large number of regions are defined. Being a deep learning method, features are extracted from the provided data automatically. Throughout the supervised learning process, filters in the hidden layers are modified so as to extract important features and minimise the prediction error.

All development in this study was done using Matlab by Mathworks (<https://www.mathworks.com>). In particular, use of the Vision toolbox was made to label and augment the dataset, as well as to train the models. The Image Labeler App was used to define bounding boxes and the corresponding species types around every visible jellyfish in all images within the database. All of the annotated data and the developed code will be made available upon request. The public distribution of such files will allow for a common base of labelled images that can be used to benchmark and compare the performance of future jellyfish classification algorithms.

The ultimate objective is to embed the optimised algorithm within the “Spot the Jellyfish” campaign website as well as on a mobile phone app. This service will simplify and speed up the report validation process drastically, providing a prompt response to citizen scientists on the taxonomic identity of the recorded species, and this is the main motivation behind this exercise. When fully implemented, users would just need to capture and upload an image of the observed jellyfish. The time and geolocation information will be collected from the mobile phone and the server-based Artificial Intelligence software will complete the information by estimating the number of jellyfish individuals in the bloom, as well as suggesting the taxonomic identity of the sighted jellyfish species. All the submitted photos will be used for online learning and to continuously improve the classification model deployed by the algorithm in an adaptive approach. Once fully implemented, through embedding within a customised app, for instance, the optimised algorithm can serve as a blueprint for citizen science campaigns relying on image analyses. In fact, by facilitating the automated extraction and interpretation of information from submitted images, and thus the report validation process, the same algorithm can make a positive contribution by improving the efficacy and attractiveness of the same campaigns.

2. Image Dataset

The “Spot the Jellyfish” citizen science campaign provides an interactive interface that shows a map of the Maltese Islands with all the validated jellyfish sighting reports received by connecting to the central campaign database [24]. In order to train the developed classification model, all of the jellyfish photos submitted through the campaign were initially organised in different directories, according to species. Additional photos for each jellyfish species were also obtained from online sources retrieved through the Google image search feature. A mixture of different submitted jellyfish image typologies was availed of, consisting of images that show jellyfish within the water column, jellyfish beached on shores, jellyfish photographed against complex backgrounds as well as a mixture of images with and without watermarks. Figure 1 shows images of *Pelagia noctiluca*, *Cotylorhiza tuberculata*, *Carybdea marsupialis*, *Velella velella* and salps, while Table 1 highlights the number of images defined

in each class. In total, 602 images were used to train the classification model, while 200 images were used for its testing. The 40 images used to test the classification accuracy of each species were chosen randomly from the available set of photos.

The original images were of varying sizes. In the preprocessing stage, the containment area within each image that showed jellyfish individual/s, was defined by recording the lower left coordinate, the width, and the height (in terms of pixel numbers), of the bounding box. During this stage, some of the images were manually cropped to remove unnecessary image background so that the species occupied around 60 percent of the image. The process is demonstrated in Figure 2.



Figure 1. Images of *Carybdea marsupialis*, *Cotylorhiza tuberculata*, *Pelagia noctiluca*, salp and *Verella verella*.



Figure 2. Original (left) and cropped (right) versions of a *Pelagia noctiluca* training image.

Table 1. Number of training and testing images per species.

Species	Number of Training Images	Number of Testing Images
<i>Carybdea marsupialis</i>	89	40
<i>Cotylorhiza tuberculata</i>	201	40
<i>Pelagia noctiluca</i>	170	40
Salps	52	40
<i>Verella verella</i>	90	40

In a second processing stage, all images and the corresponding boundaries that specify the regions that contain the jellyfish, were resized so that the maximum width was that of 224 pixels. These dimensions were applied in order to apply transfer learning protocols and take advantage of models that were pretrained using the same image sizes. The rasters were downscaled using a bicubic interpolation scheme with antialiasing. Up to this stage, the image height was left to vary proportionally, so not to distort the same images.

3. Classification Model

Although some of the jellyfish images extracted from the “Spot the Jellyfish” campaign database were captured by SCUBA divers, and thus show the entire jellyfish against a consistently blue background, the majority of database photos were captured from a boat or from the shore. Therefore, even though the images were all of high resolution, in most cases the jellyfish individuals only occupied a small region of the frame. As a result, highly reliable detection and classification schemes were required to correctly identify the jellyfish in the images.

Initially, conventional convolution neural networks (CNNs) were trained to classify the raw and unprocessed images into five categories for the identification of *Pelagia noctiluca*, *Cotylorhiza tuberculata*, *Carybdea marsupialis*, *Verella verella* and salps jellyfish. While good results were generally obtained,

the applicability of region-based convolution neural networks (R-CNNs) was subsequently tested and found to further improve the classification method's accuracy. Such deep learning techniques improve on the CNN models by following a two-stage approach. Initially, sub-regions within the image that might contain a jellyfish are identified. Features are then extracted from the convolved areas and adopted as an input for the classifier.

In this study, five classification models were implemented to distinguish between five different jellyfish species. In all cases, transfer learning was used to improve the accuracy. Such an approach involves the use of a pre-trained CNN that can extract features from thousands of natural images coming from various categories. The developed faster R-CNN models were then optimised to extract features that are specific to jellyfish images. The five models were based on ResNet18, ResNet50 and GoogLeNet. The ResNet18 and ResNet50 networks were 18 and 50 layers deep, respectively. These have been trained on more than a million images [25]. On the other hand, GoogLeNet is 22 layers deep [26]. Since these networks are based on 224×224 images, all jellyfish frames had to be resized in order to ensure consistent dimensions. The resizing of images was carried out through a bicubic interpolation scheme. Table 2 provides a summary of the salient differences between the five deployed classification models. The time required to classify the 200 test images is also listed.

Table 2. Details of the five implemented classification models.

Model No.	Feature Extraction Network	Feature Layer	Number of Anchors	Time to Classify 200 Images (s)
Model 1	ResNet50	activation_40_relu	3	172.90
Model 2	GoogLeNet	inception_3a-output	3	312.36
Model 3	GoogLeNet	inception_5b-output	3	59.43
Model 4	ResNet18	res3b_relu	3	169.16
Model 5	GoogLeNet	inception_3a-output	10	341.56

Data augmentation was also applied to increase the number of training examples. Deep learning techniques such as the Faster R-CNN method often rely on big data. However, in this case, only a few examples were available. In such circumstances, the model might learn the function very well by overfitting. Although very accurate results will be noted during the training process, the performance on unseen images will not be equivalent. In this study, as only 602 training images were available, overfitting was avoided through augmentation. As discussed by Shorten and Khoshgoftaar [27], such a process will help to create synthetic instances to increase the size of the training set. In this case, the jellyfish images were vertically and horizontally flipped and in each case, a different feature layer was defined. For the majority of cases, the number of anchors was set to three. However, in classification Model 5, this was increased to ten. The test images were not part of the training set. As required by the models, these were resized to 224×224 pixels. When more than one jellyfish was detected within the submitted image, the one with the highest percentage probability was considered. The adopted framework for the training of each classifier is graphically represented in Figure 3.

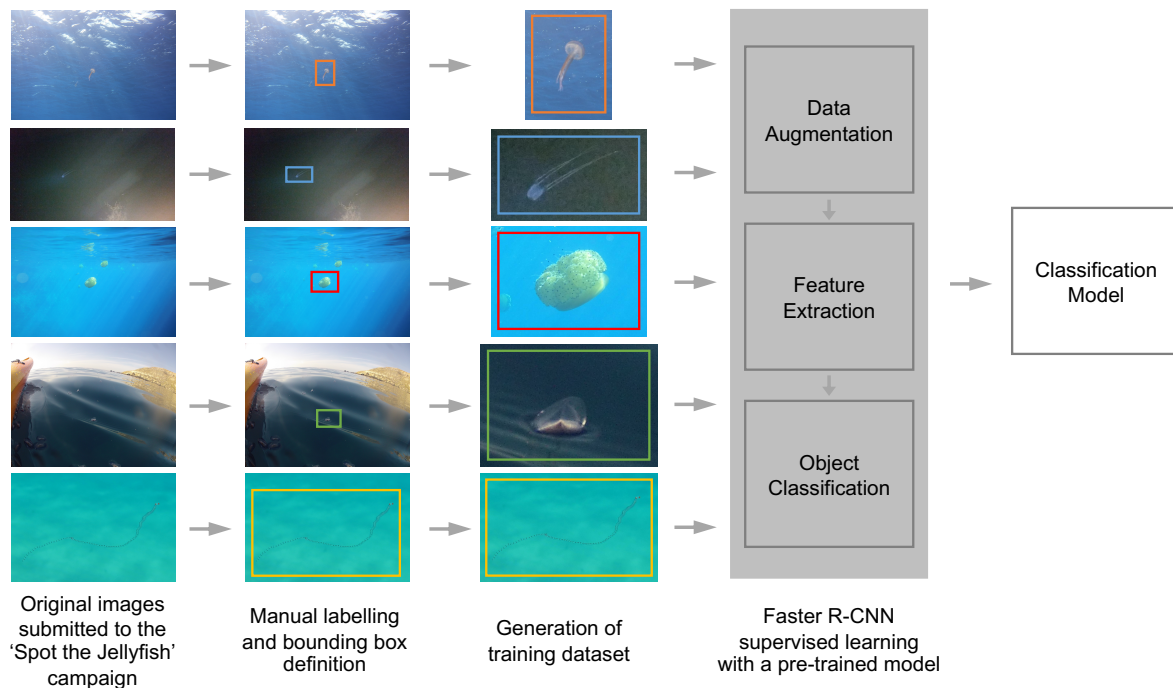


Figure 3. Adopted framework for the training of each classifier.

4. Results

The confusion matrices obtained from testing all five classification models with the same set of images are presented in Table 3. Here, the columns represent the true (actual) class of the jellyfish shown in the test images, while the rows correspond to the category predicted by the classifier. A “not classified” category was added to the list of categories within the “predicted” section, to denote the images that registered a jellyfish which was not characterised by the classifier. The test set contained 40 images for each jellyfish species. For a perfect classifier, the values within the diagonal would be 40, implying that the correct species was identified every time. As demonstrated below, this was not always the case. Confusion matrices help to identify whether a classifier is behaving randomly or whether for instance, it cannot distinguish between two classes.

The precision, recall and f_1 score for every class by every model were computed by Equations (1)–(3). Here, TP represents the number of true positive cases and corresponds to the number of instances for which the correct class was predicted. On the other hand, FP corresponds to the false positives and indicates the number of times other jellyfish were incorrectly classified with the group being tested. The number of instance that belong to the class under test but were labelled otherwise are treated as false negatives (FN). The results obtained for the trained models are presented in Table 4. The precision score gives an estimate of the fraction of the predicted jellyfish species that actually belong to the same species. The recall score indicates the percentage of correctly classified instances. Obviously, classifiers with high precision and recall scores are preferred. The f_1 score combines these two metrics by taking the harmonic mean. The Cohen’s κ score that measures the convergence between the actual classes and the predicted ones, was computed as shown in Equation (4). For this metric, p_o represents the actual observed agreement and is equivalent to the accuracy of the model while p_e is the probability expected by a random classifier. In this case, a value of 1 denotes perfect agreement, while a value of zero indicates that the classifier performance is similar to that of a random draw.

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$f_1 = \frac{2(Recall \times Precision)}{Recall + Precision} \tag{3}$$

$$\kappa = \frac{p_o - p_e}{1 - p_e} \tag{4}$$

Table 3. Confusion matrices for classification.

		Actual				
		<i>Carybdea marsupialis</i>	<i>Cotylorhiza tuberculata</i>	<i>Pelagia tuberculata</i>	Salps	<i>Velella velella</i>
Model 1	<i>Carybdea marsupialis</i>	36	0	2	1	0
	<i>Cotylorhiza tuberculata</i>	0	38	0	0	0
	<i>Pelagia noctiuca</i>	2	1	35	0	1
	Salps	2	0	2	39	0
	<i>Velella velella</i>	0	0	0	0	39
	Not classified	0	1	1	0	0
Model 2	<i>Carybdea marsupialis</i>	37	0	1	0	0
	<i>Cotylorhiza tuberculata</i>	0	38	0	0	0
	<i>Pelagia noctiuca</i>	1	2	36	0	0
	Salps	1	0	2	40	0
	<i>Velella velella</i>	1	0	0	0	40
	Not classified	0	0	1	0	0
Model 3	<i>Carybdea marsupialis</i>	40	0	0	0	1
	<i>Cotylorhiza tuberculata</i>	0	39	1	1	0
	<i>Pelagia noctiuca</i>	0	1	39	0	3
	Salps	0	0	0	39	0
	<i>Velella velella</i>	0	0	0	0	36
	Not classified	0	0	0	0	0
Model 4	<i>Carybdea marsupialis</i>	38	0	2	0	0
	<i>Cotylorhiza tuberculata</i>	0	34	0	0	0
	<i>Pelagia noctiuca</i>	2	5	35	1	0
	Salps	0	0	3	39	0
	<i>Velella velella</i>	0	1	0	0	40
	Not classified	0	0	0	0	0
Model 5	<i>Carybdea marsupialis</i>	40	0	1	0	0
	<i>Cotylorhiza tuberculata</i>	0	40	0	0	0
	<i>Pelagia noctiuca</i>	0	0	39	0	0
	Salps	0	0	0	40	0
	<i>Velella velella</i>	0	0	0	0	40
	Not classified	0	0	0	0	0

Clearly, the implementation of transfer learning has allowed very accurate jellyfish species classifications/assignments to be made. The models that performed best were those that used the GoogLeNet feature extraction network. While an increase in the number of anchors resulted in a slower classification process, a higher accuracy was also concurrently obtained through this measure. As expected, ResNet50 gave better results than ResNet18 and took less than 4 more seconds to classify the 200 test images. Figures 4–8 present a sample of the results obtained through the implementation of classification Model 5.

Table 4. Precision, recall, f_1 score and κ metrics for predictions by the five classification models.

Model	Species	Precision	Recall	f_1 Score	κ
1	<i>Carybdea marsupialis</i>	0.92	0.90	0.91	0.92
	<i>Cotylorhiza tuberculata</i>	1.00	0.95	0.97	
	<i>Pelagia noctiluca</i>	0.90	0.87	0.88	
	Salps	0.91	0.98	0.94	
	<i>Veleva veleva</i>	1.00	0.98	0.99	
2	<i>Carybdea marsupialis</i>	0.97	0.93	0.95	0.95
	<i>Cotylorhiza tuberculata</i>	1.00	0.95	0.97	
	<i>Pelagia noctiluca</i>	0.92	0.90	0.91	
	Salps	0.93	1.00	0.96	
	<i>Veleva veleva</i>	0.98	1.00	0.99	
3	<i>Carybdea marsupialis</i>	0.98	1.00	0.99	0.96
	<i>Cotylorhiza tuberculata</i>	0.95	0.98	0.96	
	<i>Pelagia noctiluca</i>	0.91	0.98	0.94	
	Salps	1.00	0.98	0.99	
	<i>Veleva veleva</i>	1.00	0.90	0.95	
4	<i>Carybdea marsupialis</i>	0.95	0.95	0.95	0.92
	<i>Cotylorhiza tuberculata</i>	1.00	0.85	0.92	
	<i>Pelagia noctiluca</i>	0.81	0.88	0.84	
	Salps	0.93	0.98	0.95	
	<i>Veleva veleva</i>	0.98	1.00	0.99	
5	<i>Carybdea marsupialis</i>	0.98	1.00	0.99	0.99
	<i>Cotylorhiza tuberculata</i>	1.00	1.00	1.00	
	<i>Pelagia noctiluca</i>	1.00	0.98	0.99	
	Salps	1.00	1.00	1.00	
	<i>Veleva veleva</i>	1.00	1.00	1.00	

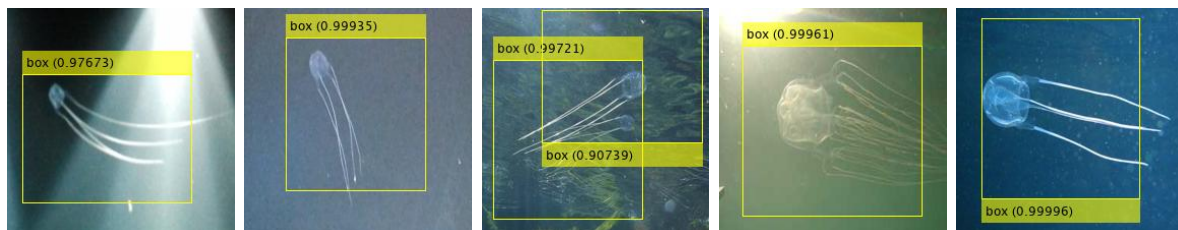


Figure 4. Output by classification model 5 for *Carybdea marsupialis* (box jellyfish).



Figure 5. Output by classification model 5 for *Cotylorhiza tuberculata* (friedegg jellyfish).

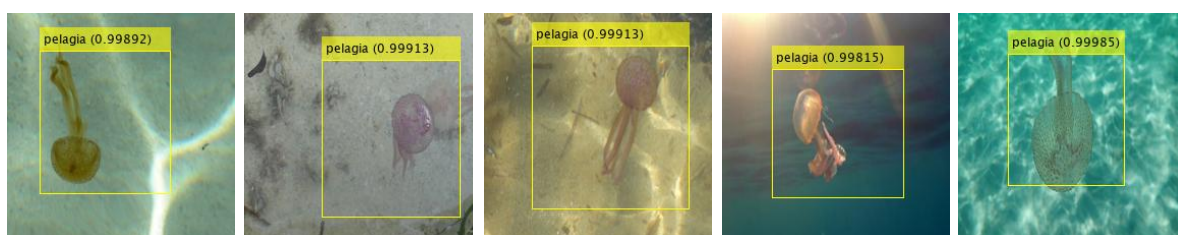


Figure 6. Output by classification model 5 for *Pelagia noctiluca*.

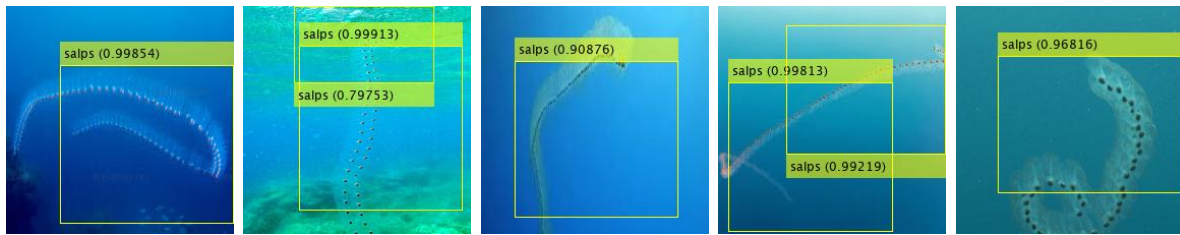


Figure 7. Output by classification model 5 for salps.

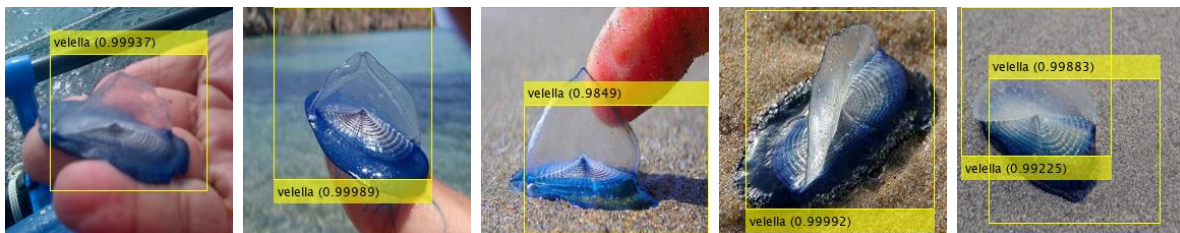


Figure 8. Sample of images that show *Velevella velevella* individuals which were used for training.

5. Conclusions

The Cohen's κ output values were high, ranging between 0.917 (Model 4) and 0.994 (Model 5), suggesting that the automated identification of jellyfish through images submitted to these classification models is generally robust and reliable. In this work, the default model parameters were used. The use of hyperparameter optimisation methods that will further fine-tune the selected model is recommended for future work. This exercise is expected to further improve on the results presented here.

In terms of precision values, the five jellyfish species adopted for the current study were classified correctly in the following decreasing order; *Cotylorhiza tuberculata* (4/5 models reported a value of 1 for precision), *Velevella velevella* (3/5 models reported a value of 1 for precision), salps (2/5 models reported a value of 1 for precision), *Pelagia noctiluca* (1/5 models reported a value of 1 for precision) and *Carybdea marsupialis* (precision values of 1 were not reported by any model).

Corresponding results obtained by Martin-Abadal et al. [10] similarly rank *Cotylorhiza tuberculata* (along with *Rhizostoma pulmo*, which was not considered in the current study) higher than *Pelagia noctiluca* in terms of correct identification success. These authors attribute this result to the fact that *Cotylorhiza tuberculata* is a larger jellyfish whose body remains relatively unchanged whilst swimming, thus rendering them more amenable for correct identification, while in *Pelagia noctiluca*, the relative position of the tentacles in relation to the main body (umbrella) changes to a greater extent with the movement of the animal, adopting a multitude of shapes, making it more difficult to identify. The fact that *Carybdea marsupialis* was the least correctly identified jellyfish species might be tentatively attributed to the highly transparent nature of the species and to the high incidence of night-time (and, consequently, low-lit) images of the species, due to its nocturnal and photo-tactic habits.

The lowest precision score obtained from the five classification models was that of 0.814 (*Pelagia noctiluca*, Model 4), with this being one value, out of just two, which were inferior to a threshold of 0.9. Similarly, high recall and f_1 score metric values were recorded through the implementation of the five classification models. This further substantiates the robust performance of the same classification models in correctly identifying jellyfish species from the images provided. The f_1 score metric values obtained in the current study, ranging between 0.843 (*Pelagia noctiluca* within classification Model 4) to 1 (*Velevella velevella*, salps and *Cotylorhiza tuberculata* within classification Model 5) are comparable to the same metric values reported within [10], which range between 0.936 and 0.952.

It is essential for every citizen science campaign to continuously monitor and validate data quality [28]. We are confident that the results presented in this study support our upcoming aspirations to embed the image analysis protocols described therein within a smart phone app and within our campaign website to enable its broad application. This would in turn spearhead its further performance testing and subsequent optimisation through the submission of an increasing number of

jellyfish images and output monitoring. Additionally, the formulation of guidelines on the taking of high-quality jellyfish photos by citizen scientists in order to ensure their applicability within image analysis protocols is recommended.

Author Contributions: All three authors contributed significantly to the publishing of this work. A.G. drafted the manuscript, planned, and implemented the experiments. J.A. assisted in the methodology and model development, validation, as well as in the review and editing of the paper. A.D. carried out the analysis of results and helped in the preparation, writing, review, and editing of the manuscript. A.D. manages the “Spot the Jellyfish” database from which most of the images used in this work were obtained. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: Most of the images used to train and test the models were obtained from the “Spot the Jellyfish” database. The authors would like to thank all those who contributed to this campaign.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Turrini, T.; Dörlner, D.; Richter, A.; Heigl, F.; Bonn, A. The threefold potential of environmental citizen science-Generating knowledge, creating learning opportunities and enabling civic participation. *Biol. Conserv.* **2018**, *225*, 176–186. [[CrossRef](#)]
2. Dickinson, J.L.; Zuckerberg, B.; Bonter, D.N. Citizen science as an ecological research tool: Challenges and benefits. *Annu. Rev. Ecol. Evol. Syst.* **2010**, *41*, 149–172. [[CrossRef](#)]
3. Chandler, M.; See, L.; Copas, K.; Bonde, A.M.; López, B.C.; Danielsen, F.; Legind, J.K.; Masinde, S.; Miller-Rushing, A.J.; Newman, G.; et al. Contribution of citizen science towards international biodiversity monitoring. *Biol. Conserv.* **2017**, *213*, 280–294. [[CrossRef](#)]
4. Giovos, I.; Kleitou, P.; Poursanidis, D.; Batjakas, I.; Bernardi, G.; Crocetta, F.; Doumpas, N.; Kalogirou, S.; Kampouris, T.E.; Keramidas, I.; et al. Citizen-science for monitoring marine invasions and stimulating public engagement: A case project from the eastern Mediterranean. *Biol. Invasions* **2019**, *21*, 3707–3721. [[CrossRef](#)]
5. Cigliano, J.A.; Meyer, R.; Ballard, H.L.; Freitag, A.; Phillips, T.B.; Wasser, A. Making marine and coastal citizen science matter. *Ocean Coast. Manag.* **2015**, *115*, 77–87. [[CrossRef](#)]
6. Kosmala, M.; Wiggins, A.; Swanson, A.; Simmons, B. Assessing data quality in citizen science. *Front. Ecol. Environ.* **2016**, *14*, 551–560. [[CrossRef](#)]
7. Newman, G.; Wiggins, A.; Crall, A.; Graham, E.; Newman, S.; Crowston, K. The future of citizen science: Emerging technologies and shifting paradigms. *Front. Ecol. Environ.* **2012**, *10*, 298–304. [[CrossRef](#)]
8. Rife, J.; Rock, S.M. Segmentation methods for visual tracking of deep-ocean jellyfish using a conventional camera. *IEEE J. Ocean. Eng.* **2003**, *28*, 595–608. [[CrossRef](#)]
9. French, G.; Mackiewicz, M.; Fisher, M.; Challiss, M.; Knight, P.; Robinson, B.; Bloomfield, A. JellyMonitor: Automated detection of jellyfish in sonar images using neural networks. In Proceedings of the 2018 14th IEEE International Conference on Signal Processing (ICSP), Beijing, China, 12–16 August 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 406–412.
10. Martin-Abadal, M.; Ruiz-Frau, A.; Hinz, H.; Gonzalez-Cid, Y. Jellytoring: Real-Time Jellyfish Monitoring Based on Deep Learning Object Detection. *Sensors* **2020**, *20*, 1708. [[CrossRef](#)] [[PubMed](#)]
11. Gauci, A.; Deidun, A.; Montebello, J.; Abela, J.; Galgani, F. Automating the characterisation of beach microplastics through the application of image analyses. *Ocean Coast. Manag.* **2019**, *182*, 104950. [[CrossRef](#)]
12. Pitt, K.A.; Lucas, C.H.; Condon, R.H.; Duarte, C.M.; Stewart-Koster, B. Claims that anthropogenic stressors facilitate jellyfish blooms have been amplified beyond the available evidence: A systematic review. *Front. Mar. Sci.* **2018**, *5*, 451. [[CrossRef](#)]
13. Akselrod-Ballin, A.; Karlinsky, L.; Alpert, S.; Hasoul, S.; Ben-Ari, R.; Barkan, E. A region based convolutional network for tumor detection and classification in breast mammography. In *Deep Learning and Data Labeling for Medical Applications*; Springer: Cham, Switzerland, 2016; pp. 197–205.
14. Liu, Z.; Hu, J.; Weng, L.; Yang, Y. Rotated region based CNN for ship detection. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 900–904.

15. Cao, X.; Wang, P.; Meng, C.; Bai, X.; Gong, G.; Liu, M.; Qi, J. Region based cnn for foreign object debris detection on airfield pavement. *Sensors* **2018**, *18*, 737. [CrossRef] [PubMed]
16. Sharma, N.; Scully-Power, P.; Blumenstein, M. Shark detection from aerial imagery using region-based CNN, a study. In Proceedings of the Australasian Joint Conference on Artificial Intelligence, Wellington, New Zealand, 11–14 December 2018; Springer: Berlin/Heidelberg, Germany, 2018; pp. 224–236.
17. Liu, Z.; Du, J.; Tian, F.; Wen, J. MR-CNN: A multi-scale region-based convolutional neural network for small traffic sign recognition. *IEEE Access* **2019**, *7*, 57120–57128. [CrossRef]
18. Li, P.; Chen, X.; Shen, S. Stereo r-cnn based 3d object detection for autonomous driving. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 16–20 June 2019; pp. 7644–7652.
19. Deng, L.; Chu, H.H.; Shi, P.; Wang, W.; Kong, X. Region-Based CNN Method with Deformable Modules for Visually Classifying Concrete Cracks. *Appl. Sci.* **2020**, *10*, 2528. [CrossRef]
20. Zitnick, C.L.; Dollár, P. Edge boxes: Locating object proposals from edges. In Proceedings of the European Conference on Computer Vision, Zurich, Switzerland, 6–12 September 2014; Springer: Berlin/Heidelberg, Germany, 2014; pp. 391–405.
21. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Region-based convolutional networks for accurate object detection and segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *38*, 142–158. [CrossRef] [PubMed]
22. Girshick, R. Fast r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 1440–1448.
23. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. In Proceedings of the Advances in Neural Information Processing Systems, Montreal, QC, Canada, 7–10 December 2015; pp. 91–99.
24. Spot the Jellyfish. 2020. Available online: <http://oceania.research.um.edu.mt/jellyfish/stats.html> (accessed on 15 November 2020).
25. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
26. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.
27. Shorten, C.; Khoshgoftaar, T.M. A survey on image data augmentation for deep learning. *J. Big Data* **2019**, *6*, 60. [CrossRef]
28. Danielsen, F.; Burgess, N.D.; Balmford, A. Monitoring matters: Examining the potential of locally-based approaches. *Biodivers. Conserv.* **2005**, *14*, 2507–2542. [CrossRef]

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).