ID05 MESOPELAGIC CRUSTACEAN HABITAT IDENTIFICATION AND ANALYSIS USING DEEP LEARNING

ANTONI BURGUERA[8], FRANCISCO BONIN-FONT[7], DAMIANOS CHATZIEVANGELOU[9]

ABSTRACT

This paper presents a software infrastructure based on Deep Learning aimed at identifying habitats of certain species of crustacean that colonize areas of the marine Mesopelagic zone. Determining their presense is done, in this case, from the detection of holes in the sand that form sets of burrow structures. Preliminar inferencing models are obtained from images captured in the North Sea by trawled UWTV (undewater TV) stations, offering quite significant detection success ratios.

Keywords – Deep Learning, Species Identification, Image Proocessing.

çThe growing interest in capturing species that inhabit the Mesopelagic zone of the sea has forced accurate studies and anlysis of their habitats and habits in order to design sustainable plans of exploitation that include adequate fishing gears [1] that do not destroy the entire environment and benthos, and, in parallel, efficient preservation actions for the species with highest commercial interest. Nephrops norvegicus is a crustacean very valued in the fish markets, and a trade off between the desire of stake-holders to satisfy the continuous demand and the protection of this particular species is mandatory. Nowadays, traditional studies of Nephrops norvegicus habitats and rhythms based on trawling [2] are highly invasive. Some alternatives include the video recording, day and night, from a grid of UWTV stations mounted on sledges towed by boats, and focus their estimations in counting for burrows, either manually [3] or with Deep Leaning [4]. Burrows save individuals from the trawl tow capture, and, well identified, are a clear sign of a colony. The ongoing spanish project PLOME [5] goes one step forward in the study of the Nephrops norvegicus, grabbing video sequences in situ with underwater vehicles, and applying CNN to automatically detect burrows and animals to estimate their density.

Object Detection (OD) is a computer vision technique aimed at identifying and locating objects in images. In the recent years, Deep Learning (DL) has shown outstanding capabilities to perform OD, clearly surpassing tradidional methods. One of the most prominent DL-OD approaches is You Only Look Once (YOLO). The aim of this work is to advance in the automatic detection of Nephrops norvegicus burrows in the context of the PLOME project, experimentally evaluating the ability of YOLOv5 [6] to perform it in underwater imagery grabbed in marine environments densely colonized with this species. This work focuses on YOLOv5, though other YOLO versions could be tested. Training DL-OD systems requires large amounts of labeled data. In general this can be problematic because data labeling is a tedious and time consuming task. When it comes to underwater imagery, specially in deep sea, the problem is magnified since the data itself is scarce and difficult to obtain. Data augmentation alleviates this problem but it can lead the DL-OD to overfit. Our novel proposal is to generate DL-OD training data in the form of image sub-samples extracted arbitrarily from photo-mosaics which were built from actual underwater images, instead of using the individual images to perform the YOLOv5 training. This approach has several advantages. On the one hand, it is less tedious and error prone for a human to tag a single large image rather than hundreds or thousands of smaller images. On the other hand, each object in the mosaic is labeled only once whereas, if the individual images are used, it has to be labeled at every single image where it appears. Our approach, thus, avoids inconsistent labels among images and generates more training data than the actual input images since viewpoints that did not exist in the images can be realistically extracted from the mosaic. This approach has also two main drawbacks. First, the mosaic has to be constructed and, second, some underwater artifacts, such as vigneting or changes in illumination, are removed by the mosaic building tools. Our proposal to solve this latter problem is to artifically add these artifacts when creating the labeled data. The fully documented source code of the proposed dataset generator is publicly available at https://github.com/aburguera/MOSAICDATASET.

EXPERIMENTAL RESULTS

A total of 1810 burrows were hand labeled in five different underwater mosaics of sizes ranging from 1900x39603 pixels to 1900x50615 pixels. The photo-mosaics were supplied by the Functioning and Vulnerability of Marine Resources research group of the ICM (Institut de Ciences del Mar-Barcelona) and images used to form them were originally grabbed from a UWTV system formed by sledges equipped with cameras, and towed at a constant speed from a vessel in the confluence of the Baltic and the Scandinavian north sea. All mosaics represent rectilinear transects approximately 20 meters long. Four of them have been used to build the train (4000 images) and validation (500 images) datasets using our proposal whilst the fifth mosaic was solely used to construct the test dataset (500 images). All the generated images have a resolution of 640x480 pixels and were grayscaled.

YOLOv5 MODEL	INF. TIME	mAP@0.5	mAP@0.5:0.95
YOLOv5s	11.571 ms	0.797	0.422
YOLOv5m	28.181 ms	0.755	0.401
YOLOv5l	51.253 ms	0.792	0.427

Tab 1. Quality metrics

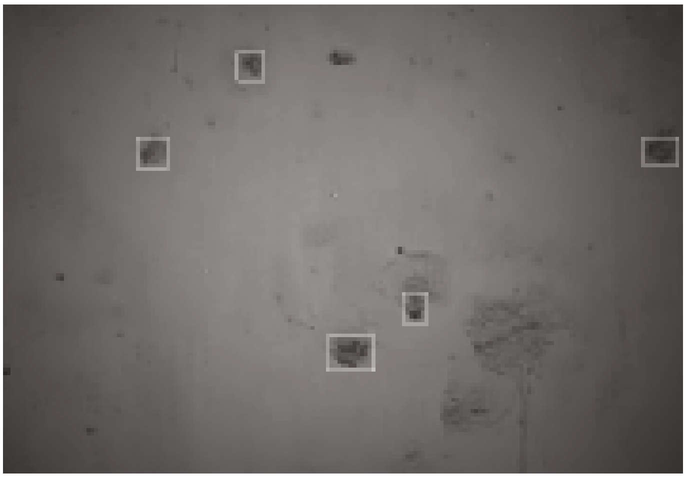

Fig 1. Image showing the infered burrows.

Afterwards, small, medium and large YOLOv5 architectures (YOLOv5s, YOLOv5m and YOLOv5l) were trained with the train dataset fine tuning the hyper-parameters with the validation dataset. Then, the quality of each model was assessed on the test dataset. The evaluation was performed on a standard laptop computer (i7 CPU at 2.9 GHz) equipped with a NVIDIA GeForce GTX 1650 and using torch-1.11.0+cu113 over Ubuntu 20.03. Table 1 summarizes the results. Even though the larger the model the slower the inference, the resulting quality is almost identical, being close to a mean Average Precisions (mAP)@0.5 of 0.8 and a mAP@0.5:0.95 of 0.4 in all cases. Figure 1 shows a sample image where inferred burrows appear in squared bounding boxes.

Since results suggest no major differences in mAP between models, we decided to focus on YOLOv5s because it shows a reasonably stable detection rate of 86.423 fps. Figure 2 shows, in the left, the Recall-Precision curves for different level combinations of Intersection over Union (IoU) and Average Precisions (AP), and, on the right, the F1-Score changes depending on the IoU and the confidence score thresholds used. This allowed us to determine the optimal thresholds, which are 0.1 for the IoU and 0.5 for the confidence score. Using this optimal configuration we reached not only a precision close to 0.8 but also a recall of 0.74 and an F1-Score of 0.77.

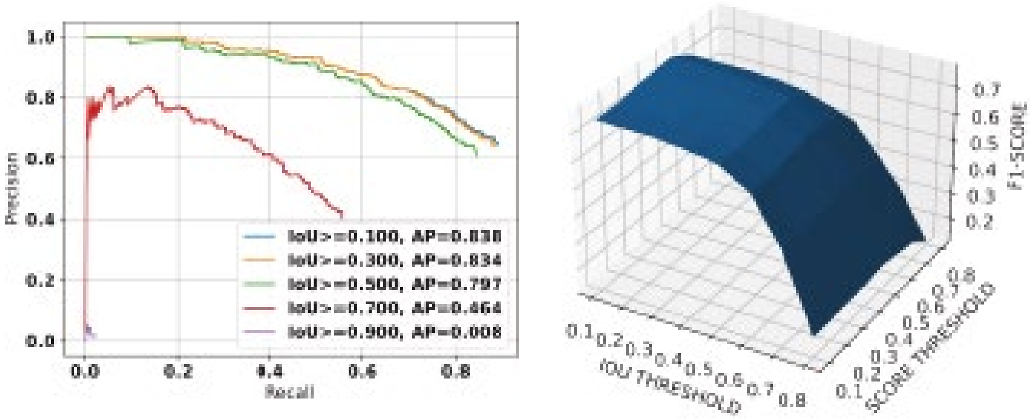

Fig 2. Recall-Precision (left) and F1-Score (right) curves for the trained YOLOv5s.

ACKNOWLEDGEMENTS

This work is partially supported by Grant PLEC2021-007525/AEI/10.13039/501100011033 funded by the Agencia Estatal de Investigación, under Next Generation EU/PRTR.

REFERENCES

[1] J. Brčić, B. Herrmann, M. Mašanović, S.K. Šifner, and F. Škeljo, "CREELSELECT---A Method for Determining the Optimal Creel Mesh: Case Study on Norway Lobster (Nephrops norvegicus) Fishery in the Mediterranean Sea", Fisheries Research, vol. 204 , pp. 433-440, 2018.

[2] J. Aguzzi and F. SardÃ . A History of Recent Advancements on Nephrops norvegicus Behavioral and Physiological Rhythms. Reviews in Fish Biology and Fisheries 18, 235â€“248 (2008).

[3] C. Lordan, J. Doyle, R. Bunn, D. Fee and C. Allsop. Aran, Galway Bay and Slyne Head Nephrops Grounds (FU17) 2011 UWTV Survey Report (2023).

[4] A. Naseer, E. N. Baro, S. D. Khan and Y. V. Gordillo, "Automatic Detection of Nephrops norvegicus Burrows in Underwater Images Using Deep Learning," Global Conference on Wireless and Optical Technologies (GCWOT), Malaga, Spain, 2020, pp. 1-6.

[5] PLOME Project, https://plomeproject.es/

[6] Ultralytics, YOLOv5 https://pytorch.org/hub/ultralytics_yolov5/, https://github.com/ultralytics/yolov5