2016 13th Conference on Computer and Robot Vision

Subsea fauna enumeration using vision-based marine robots

Karim Koreitem*, Yogesh Girdhar†, Walter Cho‡, Hanumant Singh†, Jesus Pineda†
and Gregory Dudek*
*School of Computer Science, McGill University, Montreal, Canada.
Email: karimkor@cim.mcgill.ca
†Woods Hole Oceanographic Institution, Woods Hole, Massachusetts, USA.
Email: yogi@whoi.edu
‡Point Loma Nazarene University, San Diego, CA, USA

***Abstract*—This paper describes a robotics system for population density estimation of marine organisms and vision-based algorithm for computing the associated population estimates. We focus on benthic fauna, through the use of Seabed AUV to collect benthic imagery, and then employ a support vector machine (SVM) for automated analysis of these images to estimate the population of the fauna of interest. The proposed approach is a significant improvement over existing techniques such as trawling, or manual inspection of images collected by a towed vehicle.**

We tested our proposed technique by first collecting benthic image data using the Seabed AUV at Hannibal seamount in Panama, and then predicting the counts of the crabs and squat lobsters in the data. We compare our predictions with ground-truth data from thousands of sample locations containing manual counts estimated by a team of experts, and found that our estimates have 94% precision and recall on held out test data.

***Keywords*-Marine Robotics; Visual Learning;**

I. Introduction

This paper considers estimation of the density of bottom (benthic) subsea fauna, specifically crabs and squat lobsters, using fully automated robotic sampling and vision-based detection and density estimation. Our approach is based on the analysis of image data from a robotic data collection system.

Knowledge of the distribution and abundance of bottom organisms is important for understanding their ecology, and for protecting these organisms, e.g., for the design of marine protected areas.

To better understand the ecology of bottom organisms, we first need to have accurate population estimates for various species. At the very minimum, we require estimates of distribution and abundance. Current standard approaches for computing these estimates includes catching the organisms, or visually analyzing images. These approaches are labor intensive and can damage the habitat. Moreover, the cost of performing such a survey can be substantial.

Through the use of robotic systems, we believe that an approach with lower environmental impact and perhaps even improved accuracy can be developed. This entails two critical functions: deployment of robotic systems to cover regions of interest on the sea floor, and the collection and assessment of data regarding the population density and prevalence of species of interest. Our long-term objective is to develop autonomous robotic systems that can operate in the ocean and collect and analyze such data on an ongoing basis. In the interim, we are describing work in which an autonomous underwater vehicle robot collects underwater image data and the collected data is analyzed offline after the end of the expedition. An image of the robot and a successful set of target detections on the data collected by the robot can be seen in Fig. 1.

Our work is based on the collection of video imagery along a predetermined trajectory selected prior to the deployment of the robot [1]. While some researchers, including teams we have been associated with, have examined autonomous real-time navigation, the depth and expense of this deployment preclude that approach and the emphasis of this paper is on the data analysis. We consider an automated image processing system that computes population densities of two species of interest (Galatheid crabs and squat lobsters) using batch computation after the robot returns to the surface.

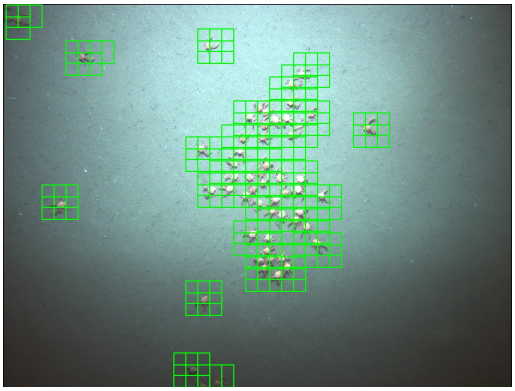
In this paper, after a brief discussion of related research, we consider the overall scope and nature of the data analysis system being developed at Woods Hole, and how this component relates to it. We then discuss our robotic platform and the specifics of the data collection missions. Then we describe our approach to vision-based population assessment using filtering and machine learning. Finally, we present results on data-driven parameter selection and an analysis of experimental results and error rates.

II. Related Work

A number of groups have considered the use of robotic systems for shallow water as well as benthic exploration and data collection. In prior work, we have deployed a small vehicle for coral reef observation and data collection. Several groups have developed and deployed vehicles that can make close approaches to the ocean floor, coral reefs, or aquatic structures [2], [3]. Close range observation of undersea environments is challenging due not only to the logistic

978-1-5090-2491-9/16 $31.00 © 2016 IEEE
DOI 10.1109/CRV.2016.53

(a)



(b)

Figure 1. In this work we use the seabed robot (a) to collect benthic imagery, and present an automated technique to enumerate fauna of interest for the purpose of quantifying the health of a subsea ecosystem. Figure (b) show an example of fauna detection, in this case of Galatheid crabs. The green boxes indicate a patch that is detected as containing a crab.

challenges of marine deployment, but also because (a) the propulsion systems for ocean-hardened vehicles may be unsafe to operate close to fragile underwater environments; (b) some devices have limited maneuverability; (c) it is difficult for humans to produce pre-planned trajectories since sensor feedback underwater is often poor, communications are difficult and terrain models are rarely complete; (d) inappropriate thrust can disturb sediments or marine species thus interfering with the measurement process.

Beijbom et al [4] looked at coral image classification using multi-scale color and texture filter banks at multiple scales. They found that a careful selection of multi-scale filter banks can outperform standard texture-only methods when estimating coral coverage in the open ocean. While related in spirit, that work focused on a rather different task where the classification of high-quality images largely filled by a single coral species was the task. In contrast, we consider the entire robotics pipeline and focus on density estimates in images where the target class(es) may only occupy a small fraction of the overall image. Similarly, Bewley et al [5] tackled detection of seaweed species in sea floor images and relied on local image features combined with a supervised learning model.

Many authors have considered the issues of developing stable AUVs for underwater data collection [6], [7], [8]. In prior work, we have considered the use of a small portable underwater vehicle for image processing tasks with high levels of maneuverability [9], [10], [11]. This has included automated coral identification from relatively low-resolution imagery collected as the vehicle moves over shallow water reefs. In contrast, the present paper addresses density estimation on a different class of vehicles in much deeper water and uses a richer representation of image content to allow a higher level of performance.

Several authors have examined the assessment of environmental parameters in marine environments using robotics systems. In addition to free swimming autonomous underwater vehicles, towed surface sensors provide an effective mechanism to collect some types of data such as shallow-water coral images [12]. Of course, due to light absorption and scattering, surface vehicles are suitable only for shallow water. Measurements in deeper water present serious logistical challenges either in terms of deployment of fixed immobile sensors, or in the deployment and operation of mobile systems, yet progress is being made. For example, in [13] 3D reconstructions of the sea floor are recovered. Several groups have used such vehicles typically with highly specialized operational parameters [14]. In exciting work that includes tracking of motile species, Dunbabin et al. consider the tracking of invasive crown of thorns on coral reefs using vision based imaging [15].

Large-scale surveys of marine environments are only recently becoming efficient due to the combination of robust autonomy, vehicles with sufficient endurance, and suitable data processing infrastructures. One notable landmark is the large scale survey of kelp forests near Australia using an autonomous underwater vehicle whereby thousands of square meters of shallow-water sea floor were surveyed [16]. That work did not apply image-based classification methods to the data set, but it did illustrate the potential to cover large regions of the ocean. In the current work we examine much smaller regions, but at rather greater depths and with image-based methods to identify specific marine taxa.

III. Approach

A. Problem Formulation

Our objective is to obtain population estimates (i.e counts) of organisms at or near the ocean bottom to estimate their distribution and abundance. Doing this can impact the habitat when direct sampling methods are used (e.g trawling), and this process can also be very laborious. For example, when images are visually analyzed by human beings, substantial effort is typically expended to obtain consistent meaningful data. Automatic methods to analyze images and extract

counts from photographs are expected to make a significant improvement in the process of obtaining counts from images, and thus of obtaining low-impact population estimates. Here we present a method of enumerating bottom-dwelling organisms using vision-enabled marine robots.

Our approach to the problem can be split into two parts. First we deploy an underwater robot equipped with a high resolution, high dynamic range camera, to collect images of the seafloor, over a statistically representative path. Then, given the collected images, we train a classifier on a small subset of the data to recognize the fauna of interest, and then apply it to the entire dataset. As we are aiming to tackle a density estimation problem on top of the detection problem, we decide to follow a supervised approach. We validate the approach by comparing the predicted counts with held out ground truth data.

B. Robot Platform

We used the Jaguar autonomous underwater vehicle (AUV), a member of the Seabed family of robots from Woods Hole Oceanographic Institution, to collect the data presented in this paper. Jaguar, shown in Fig.1(a), is an ideal platform for visual seafloor inspection tasks. Its dual hull design puts most of the payload weight in the lower hull, while keeping the upper hull relatively more buoyant. This sets up a naturally stable vehicle configuration that allows for collecting image data in low light conditions.

As mentioned above, Jaguar is equipped with a high resolution camera with a high dynamic range sensor size and a powerful strobe light to capture high quality images of the seafloor. Images collected from the Jaguar camera are color-corrected using the approach developed by Singh et al. [17].

Jaguar AUV has an endurance of 24 hours, and can go to depths of up to 5000m. The piezoelectric pressure sensor onboard the robot can measure depths with 1cm of precision. The AUV navigates underwater using an optical north-seeking gyro for heading, a doppler velocity log (DVL) for measuring ground speed and altitude, and Long Baseline (LBL) acoustic beacons for absolute localization. The vehicle can keep itself localized with position uncertainty of less than a meter.

Our image analysis process was conducted offline using data recovered in the mission(s) as described above. We developed a classifier pipeline that is able to process incoming data streams from the robot and estimate the density of species of interest.

Our approach to species sampling, classification and density estimation is a combination of manually selected filtering, machine learning, and data driven calibration. Our image processing pipeline and the learning procedure is described in III-D, but a critical precursor is the acquisition not only of sample data, but also known ground-truth data for use in training the learning-based components of the classifier.

C. Data Annotation

A selection of images and image patches were extracted from our data set to constitute the training data for learning and tuning the classifier. The resolution of the collected images is 1360x1024 while the image patches are of size 64x60. This data was selected based on randomly sampling the data set and then manually selecting additional images to assure coverage of the diverse kinds of terrain and illumination experienced by the robot.

Each of the randomly sampled images was divided into 357 non-overlapping patches of size 64x60 according to the same sampling parameters used by the automated classifier (described below). These subwindows were then presented to a "domain expert" to classify them into three simple classes for each species of interest:

- **Yes**: The species of interest occupies at least 50 *per cent* of the image patch.
- **No**: The species of interest occupies under 30 *per cent* of the image.
- **Reject**: The image is indeterminate with respect to the above criteria or it is not representative of the target data set (i.e. due to an imaging failure) or occlusion.

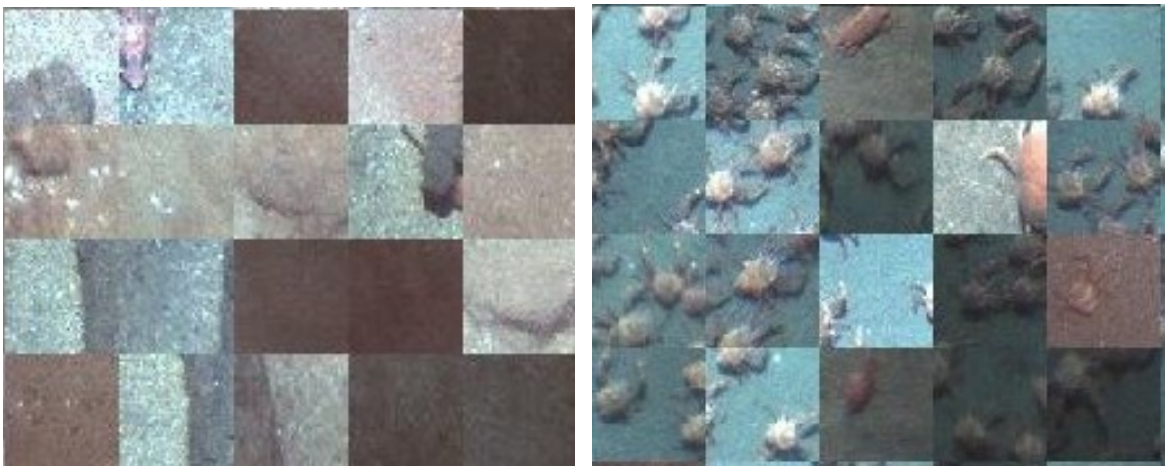
The data set of images with positive ("yes") and negative ("no") image patches were then subdivided into mutually exclusive subsets used for training and testing as described below.

1) Training set: The training dataset consists of 5145 manually annotated image patches balanced in a 56/44% negative/positive split. Compilation images of subsets of these negative and positive patches are presented in Fig.2).

2) Test set: The test dataset consists of 921 image patches balanced in a 55/45% negative/positive split. These patches were arbitrarily selected from the dataset and the test and training sets are mutually exclusive. The patches that constitute the test set are also not subwindows of images that were used in the training set.

D. Method

In this section, we describe the proposed method for crustacean detection and density estimation. The objects we are trying to detect are generally small and occupy small fractions of each image individually. However, in some cases, they are sufficiently numerous that they either occlude one another or occupy a substantial fraction of the overall scene. As a result, we employ a sliding window scheme that allows us to examine small fractions of an image in isolation. It is also important to discriminate several different types of fauna from one another as well as from the background. In particular, the species that we are most interested in consist of a combination of Galatheids and Brachyuran crabs.



(a) (b)

Figure 2. **Compilation of annotated positive and negative patches:** The two images presented above are compilations of (a) 5x4 negative patches not containing any fauna of interest and (b) 5x4 positive patches containing examples of crabs and squat lobsters that the system is designed to detect. It is worth noting that the negative patches can contain any of the following and more: clear sand floor, pebbles, boulders, cobbles, sediments, sea pens, sea urchins, sponges, corals and other miscellaneous fish. On the other hand, positive patches contain a variety of galatheids and brachyuran crabs.

These crustaceans of interest have a conspicuous shape that makes them fairly easy for a biologist to recognize, when seen in isolation (at least at a coarse taxonomic level). While shape cues alone may be sufficient for isolated organisms, the fact that they are often occluded, particularly by one another, led us to the observation that chromatic cues are also very valuable, even if they can be distorted by wavelength-dependent absorption. In addition, the non-homogeneity of the lighting conditions and the fact that our images were captured from a top view meant that brightness and orientation invariance were necessary. Therefore, we opted to use a texture descriptors based approach which was ideal in representing the low frequency intensity variations of crustaceans as opposed to the dense high frequency variations of the sandy ocean floor, overwhelming present in the background of the dataset.

1) Image Features: Our image classifier exploits shape-related filters to extract texture information and color-space filters to extract co-occurring color or hue features. Our shape features are based on Gabor functions which constitute a local representation for frequency-domain information.

The Gabor function is a well-established filter family for measuring energy in various frequency bands of the image while maintaining spatial localization [18]. They are described by sinusoids with Gaussian envelopes and have proven effective in texture classification [19] and when used in appropriate combinations, as applied here, they can be described as a wavelet. A Gabor wavelet has the form

$$g_{mn} = \frac{\|k_{mn}\|^2}{\sigma^2}\exp\left(-\|k_{mn}\|^2\frac{\|(x,y)\|}{2\sigma^2}\right)\left[e^{ik_{mn}(x,y)} - e^{-\sigma^2/2}\right]. \quad (1)$$

The wave function is given by

$$k_{mn} = k_n e^{i\phi_m}, \quad (2)$$

and has components $k_n = k_{max}/f^n$ and $\phi_m = \pi m/8$. The maximum frequency is denoted k_{max} and f is the frequency spacing. An image can be transformed using a wavelet through the convolution operation:

$$G_{mn}(I) = \int I(x_1, y_1) g^*_{mn}(x - x_1, y - y_1) dx_1 dy_1. \quad (3)$$

The m parameter determines the orientation of the wavelet, n selects the frequency and σ determines the spatial support. By filtering with multiple wavelets, it is possible to extract a variety of content from an image. Specifically, in order to achieve some level of rotational invariance, we apply eight filter orientations and the signatures for each filter is represented by its amplitude histogram across the image. Our signal is further reduced by representing the amplitude histogram by its mean-variance pair, an approach that allows robust comparisons between histograms to be computed [20], [21].

In order to represent the color information of each patch, a color histogram is extracted from the patch in the HSV color space. For the three different HSV color channels, we use 24, 3 and 3 bins respectively to quantify the distribution of each channel separately and concatenate each computed 1D histogram of each channel into one final 1x30 vector containing the three histograms.

Both the Gabor features and the color histogram vector are then concatenated to form one final 1x62 length feature vector. The Gabor vector is 1x32 including the 8 orientations. With each patch described by a specific vector, we now learn

a support vector machine (SVM) from the indexed training images. With the learned model, we can then proceed to make crab and lobster detection predictions on new test images.

E. Implementation

1) SVM parameters selection and cross-validation: The scikit-learn implementation [22] of the libsvm support vector machine (SVM) were used throughout the work. In order to optimize the selection of the SVM parameters, we use k-fold cross-validation with k = 5 by partitioning our original previously described training set into five equal sized subsets. Of these new subsets, one subset at a time is set aside as a validation set for testing our model. We then proceed to repeat the process k-times (k=5) for each subset using all five as validation test sets separately, averaging the performance over the five folds to provide a single estimation. This cross-validation technique is then repeated for a number of SVM parameter and kernel combinations in order to identify the best kernel and respective parameters to use in our system. Because the SVM parameter selection process is cross-validated, the risks of over-fitting our datasets are minimized.

The SVM parameters are as follows (taken from scikit-learn):

1) kernel: Specifies the kernel type to be used in the algorithm. Examples: linear, polynomial, and radial basis function (rbf).
2) C: penalty parameter C of the error term.
3) gamma: Kernel coefficient for 'rbf' and 'poly' kernels.

As shown in Fig. 6, the optimized parameters for the SVM are in our case: **polynomial** kernel, C = 10, gamma = 10.

Before we can proceed to train the system, we must first manually annotate the training image patches. This is done by a "domain expert". Once the training data is annotated, we can proceed with our training procedure.

2) Algorithm: Training Step: In order to train the system, we first index the training dataset by describing each patch in the dataset and and then train the support vector machine using the indexed dataset. To do this, the following steps are taken:

1) Convert the images into the HSV color space.
2) Apply adaptive thresholding on the V channel of the image patch in order to simply segment more obvious foreground objects in the images, specially with a more uniform sandy background.
3) Describe each patch with our designed feature descriptors as described above.
4) Train the support vector machine using the newly described training set.

After we have learned our model, we can now use the classifier to evaluate new test images.

3) Testing and Evaluation: The same procedure that is used to describe the training patches is used to extract the feature descriptors of test patches and the SVM is then used to classify the test patch as either containing crab or not. False positives are also taken and dynamically added to the negative training set in order to continuously improve the system.

IV. Data collection and assessment

A. Data collection

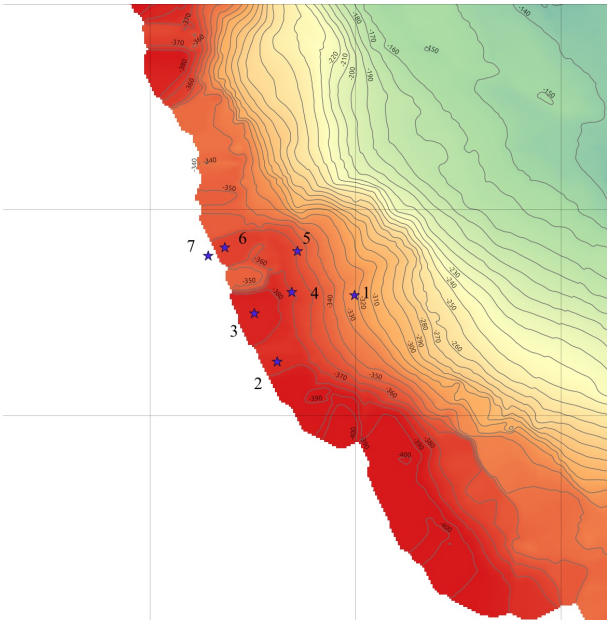
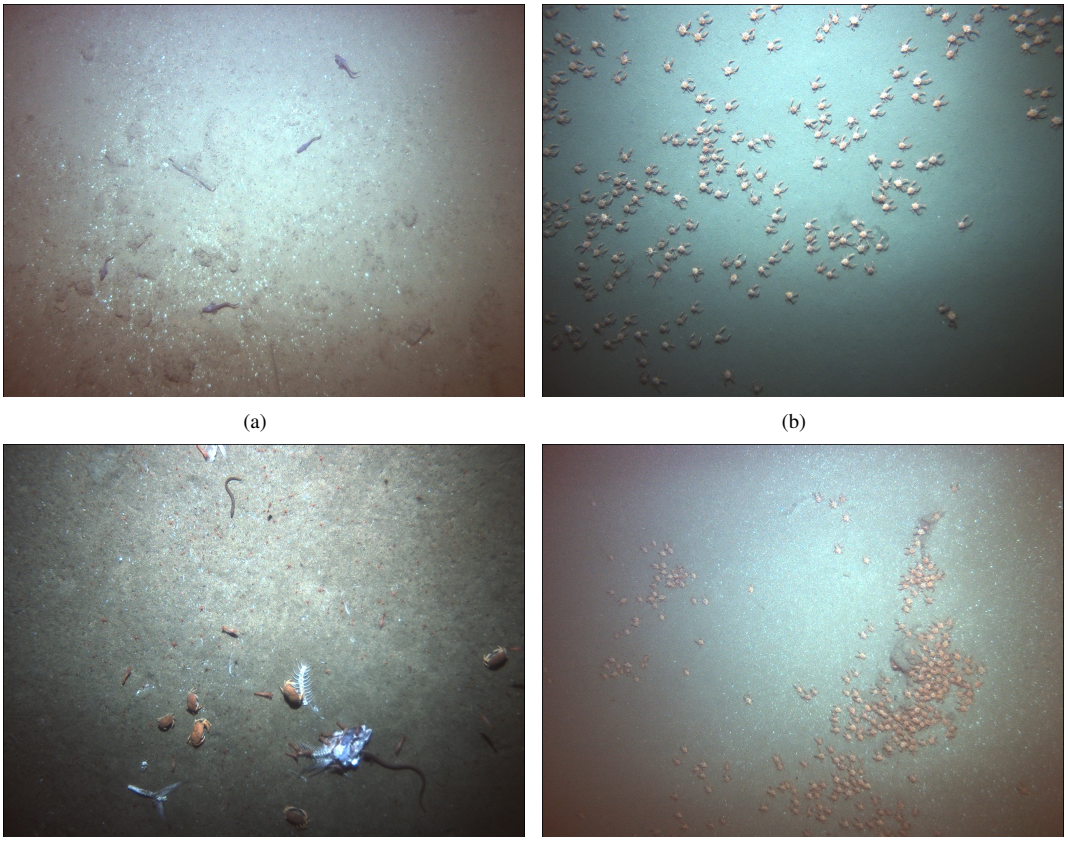


Figure 3. **AUV Waypoints:** The seafloor map showing waypoints used for planning the AUV mission to collect the dataset.

The data set was collected using the Jaguar AUV at Hannibal seamount, off the Pacific coast of Panama, at a height of 4 meters above the seafloor, with depth varying between 300-400 meters [1]. Figure 3 shows the waypoints used by the AUV for the data collection task. A total of 2296 continuous images were taken over the span of roughly 3 hours and 40 minutes at a constant frequency. These images are taken in sequence as the robot moves autonomously through waypoints that were pre-defined by the expedition crew. Therefore, chronological and spatial continuity can be assumed to be features of the dataset although our system does not make use of these features.

The goal of the dataset was to capture various species of Galatheids and Brachyuran crabs. Many of these species travel in aggregates while some are found in more isolated situations. However, the dataset also comes across other types of fish and organisms that are not of interest to this paper.

Sample images of the dataset are included in Fig. 4.



(a) (b) (c) (d)

Figure 4. **Dataset sample images:** Image (a) shows an example image that does not contain species of interest, but other fish while images (b), (c) and (d) illustrate the diversity in the target species that our system can handle. Note that (c) in particular also contains several non-target objects.

B. Experimental Results

At this stage, we evaluate our detection system both (a) qualitively by running our classifier over sliding windows with new test images and (b) quantitavely by running the classifier over our carefully annotated test set.

1) Qualitative Analysis: By running our sliding window evaluation, we can visually review the performance of our classifier by reviewing clear false positive and false negative detections. While this method is purely visual and not quantifiable, it is still valuable to confirm the general performance of the system. An example test image in which detected fauna are highlighted is presented in Fig.5.

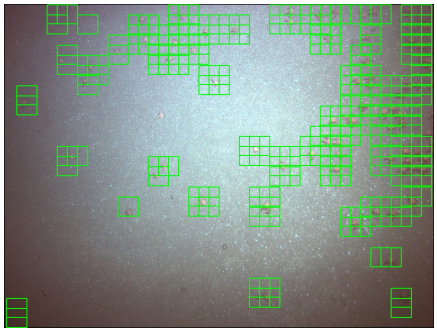


Figure 5. **Sliding Window Detector: After running the classifier in a sliding window fashion over the above test image, we can qualitatively evaluate the performance of the system by reviewing false positive and negative detections.**

2) Quantitative Analysis: To formally evaluate our system, we proceed in two steps. The first evaluation tests the system's learned model and its classification performance. To do this, we evaluate the test set that is independent from our training set as described in section III-C2. We evaluated our classifier on this test set and the results are presented in Fig.6. On our test set, our system achieves 94% for both recall and precision.

In order to evaluate our population density estimates, we simply tally the counts over each image of the entire dataset. We then compared our estimates to the manually annotated images which had ground truth counts of the various species of fauna present. The results are presented in the form of a bar graph pair which overlays the ground

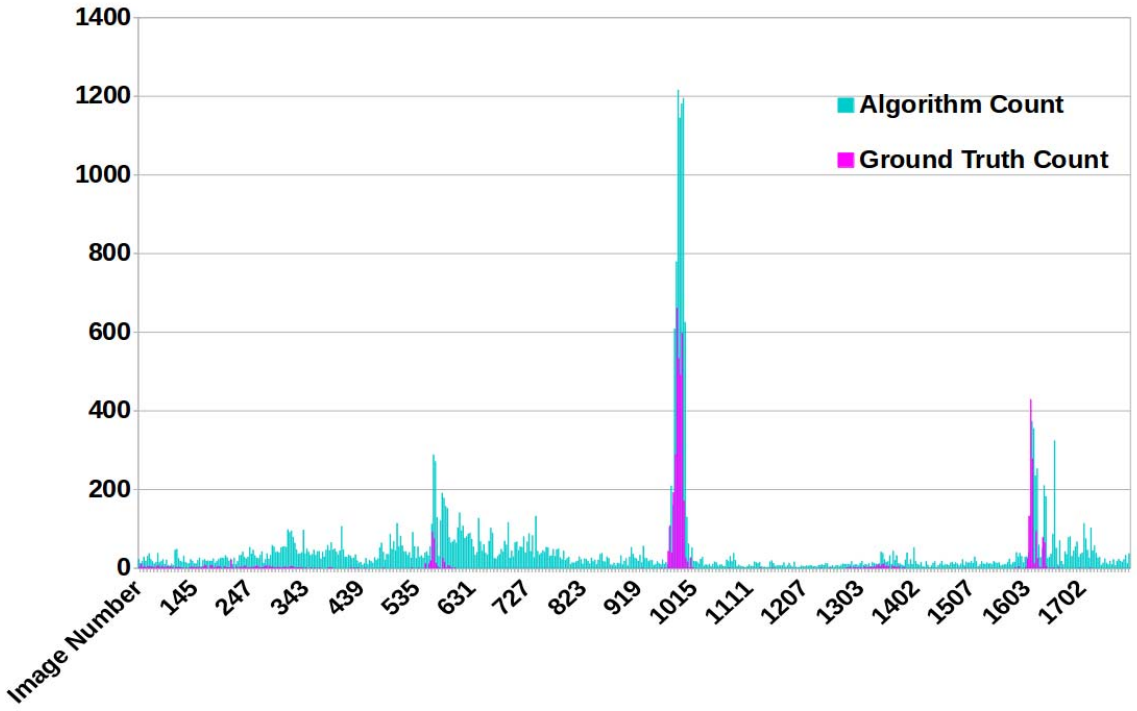


Figure 7. **Population density estimates:** Overlapping bar graphs of galatheids squat lobsters and brachyuran crab counts over time. The cyan graph showcases the number of windows detected as containing crab by the classifier over time while the magenta graph represents the ground truth counts carefully determined through manual annotation by domain experts. The time axis is represented by the image numbers of the images collected chronologically in the Panama dataset. The matching peaks of the graph clearly demonstrate the tight correlation between estimated and manually tallied crab population densities throughout the dataset. It is important to note that because we are using overlapping sliding windows, the system counts are systematically (and correctably) higher than the actual counts as can be seen in Fig.1(b) and Fig.5.

Training	precision	recall	f1-score	support
negative examples	0.97	0.97	0.97	2906
positive examples	0.96	0.95	0.96	2239
avg/total	**0.96**	**0.96**	**0.96**	5145
Testing	**precision**	**recall**	**f1-score**	**support**
negative examples	0.95	0.95	0.95	509
positive examples	0.94	0.93	0.94	412
avg/total	**0.94**	**0.94**	**0.94**	921

Figure 6. **Classifier performance:** The above table summarizes the performance of our classifier pipeline on a separate test set. The test set contains a balanced 921 image patches. The svm parameters are the following: **polynomial** kernel, C = 10, gamma = 10.

truth counts per image over time over the system's counts. The results are presented in Fig.7. It is important to note that because we are using overlapping sliding windows, the system counts are higher than the ground truth counts as can be seen in Fig.1(b) and Fig.5. However, the graph clearly showcases how our system is able to provide biologists with coarse estimates of the highest fauna density regions across the dataset over time. These regions are illustrated by the apparent overlapping peaks in the graphs such as at images 550, 1000 and 1650. As mentioned, the systems count are however not perfect as a measure because of their inability to handle clutter in sequences which contain heavy aggregates of crabs for example.

V. Conclusion

We have described a marine robot system that uses image data collected on the sea floor to non-destructively estimate the populations of crustaceans. In contrast to existing traditional methods that harvest and potentially deplete the stock to compute the population, our approach is non-invasive. To do this, we developed and proposed a vision-based technique for quantitatively evaluating the population of a target marine species in the wild using color image data. We first collected data using a Seabed autonomous robot in a previously unexplored part of the world, identified target species of interest, and then estimated the number of individuals of the target species in each image in the dataset. We compared our estimated counts with data that was manually annotated by a team of expert biologists, and found that our technique gives 94% precision and 94% recall rate. Major limitations of the work are found in situations of clutter where aggregates of fauna are overlapping both in depth and spatially. Another limiatation of the current method is the

inability to identify overlapping regions captured across the dataset. As a result of the previously mentioned limitations, in future work we will be focusing on extending this work to tackle more complex cluttered habitats as well as novel and more mobile species.

ACKNOWLEDGMENT

This work was supported by the Natural Sciences and Engineering Research Council (NSERC) through the NSERC Canadian Field Robotics Network (NCFRN).

REFERENCES

[1] J. Pineda, W. Cho, V. Starczak, A. F. Govindarajan, H. M. Guzman, Y. Girdhar, R. C. Holleman, J. Churchill, H. Singh, and D. K. Ralston, "A crab swarm at an ecological hotspot: patchiness and population density from AUV observations at a coastal, tropical seamount," *PeerJ (in-press)*, 2016.

[2] J. Eskesen, D. Owens, M. Soroka, J. Morash, F. Hover, C. Chryssostomidis, J. Morash, and F. Hover, "Design and performance of odyssey iv: A deep ocean hover-capable auv," MIT, Tech. Rep. MITSG 09-08, 2009.

[3] M. Woolsey, V. Asper, A. Diercks, and K. McLetchie, "Enhancing niust's seabed class auv, mola mola," in *Proceedings of Autonomous Underwater Vehicles*, 2010, pp. 1–5.

[4] O. Beijbom, P. Edmunds, D. Kline, B. Mitchell, and D. Kriegman, "Automated annotation of coral reef survey images," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, June 2012, pp. 1170–1177.

[5] M. S. Bewley, B. Douillard, N. Nourani-Vatani, A. Friedman, O. Pizarro, S. B. Williams, "Automated species detection: An experimental approach to kelp detection from sea-floor AUV images," *Australasian Conference on Robotics and Automation (ACRA)*, 2012.

[6] S. Mohan and A. Thondiyath, "A non-linear tracking control scheme for an under-actuated autonomous underwater robotic vehicle," *International Journal of Ocean System Engineering*, vol. 1, no. 3, pp. 120 – 135, 2011.

[7] J. Vaganay, L. Gurfinkel, M. Elkins, D. Jankins, and K. Shurn, "Hovering autonomous underwater vehicle - system design improvements and performance evaluation results," in *Proceedings of Unmanned Untethered Submersible Technology*, 2009.

[8] D. Font, M. Tresanchez, C. Siegentahler, T. Pallej, M. Teixid, C. Pradalier, and J. Palacin, "Design and implementation of a biomimetic turtle hydrofoil for an autonomous underwater vehicle," *Sensors*, vol. 11, no. 12, pp. 11 168–11 187, 2011.

[9] D. Meger, F. Shkurti, D. Cortes Poza, P. Giguere, and G. Dudek, "3d trajectory synthesis and control for a legged swimming robot," in *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*. IEEE, 2014, pp. 2257–2264.

[10] D. Meger, J. C. G. Higuera, A. Xu, and G. Dudek, "Learning legged swimming gaits from experience," in *International Conference on Robotics and Autonomous Systems (ICRA)*, 2015.

[11] T. Manderson, D. Meger, J. Li, D. Corts Poza, N. Dudek, and G. Dudek, "Towards autonomous robotic coral reef health assessment," *Proc Conference on Field and Service Robotics 2015*, vol. 10, no. 1, 2015.

[12] S. Williams, O. Pizarro, M. Johnson-Roberson, I. Mahon, J. Webster, R. Beaman, and T. Bridge, "Auv-assisted surveying of relic reef sites," in *OCEANS 2008*, Sept 2008, pp. 1–7.

[13] O. Pizarro, R. M. Eustice, and H. Singh, "Large area 3-D reconstructions from underwater optical surveys," *IEEE Journal of Oceanic Engineering*, vol. 34, no. 2, pp. 150–169, 2009.

[14] T. Maki, A. Kume, T. Ura, T. Sakamaki, and H. Suzuki, "Autonomous detection and volume determination of tubeworm colonies from underwater robotic surveys," in *OCEANS 2010 IEEE - Sydney*, May 2010, pp. 1–8.

[15] F. Dayoub, M. Dunbabin, and P. Corke, "Robotic detection and tracking of crown-of-thorns starfish," in *Proc. IEEE/RSJ Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2015.

[16] E. M. Marzinelli, S. B. Williams, R. C. Babcock, N. S. Barrett, C. R. Johnson, A. Jordan, G. A. Kendrick, O. R. Pizarro, D. A. Smale, and P. D. Steinberg, "Large-scale geographic variation in distribution and abundance of australian deep-water kelp forests," *PloS one*, vol. 10, no. 2, pp. 1–21, 2015.

[17] H. Singh, C. Roman, O. Pizarro, R. Eustice, and A. Can, "Towards high-resolution imaging from underwater vehicles," *The International journal of robotics research*, vol. 26, no. 1, pp. 55–74, 2007.

[18] D. Gabor, "Theory of communication," *Journal of the IEE*, vol. 93, pp. 429 – 457, 1946.

[19] I. Fogel and D. Sagi, "Gabor filters as texture discriminator," *Biological Cybernetics*, vol. 61, no. 2, pp. 103–113, 1989. [Online]. Available: http://dx.doi.org/10.1007/BF00204594

[20] J. Pickands III, "Statistical inference using extreme order statistics," *the Annals of Statistics*, pp. 119–131, 1975.

[21] P. Meer, J. Jolion, and A. Rosenfeld, "A fast parallel algorithm for blind estimation of noise variance," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 12, no. 2, pp. 216–223, 1990.

[22] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.