



Motivation

Researchers at Monterey Bay Aquarium Research Institute (MBARI) has interest in continuously observing and filming a sea creature "larvacean". Currently, remotely controlled vehicles (ROVs) are used to track the animal, which requires long human operating time. An autonomous tracking system can significantly increase the time for observation.

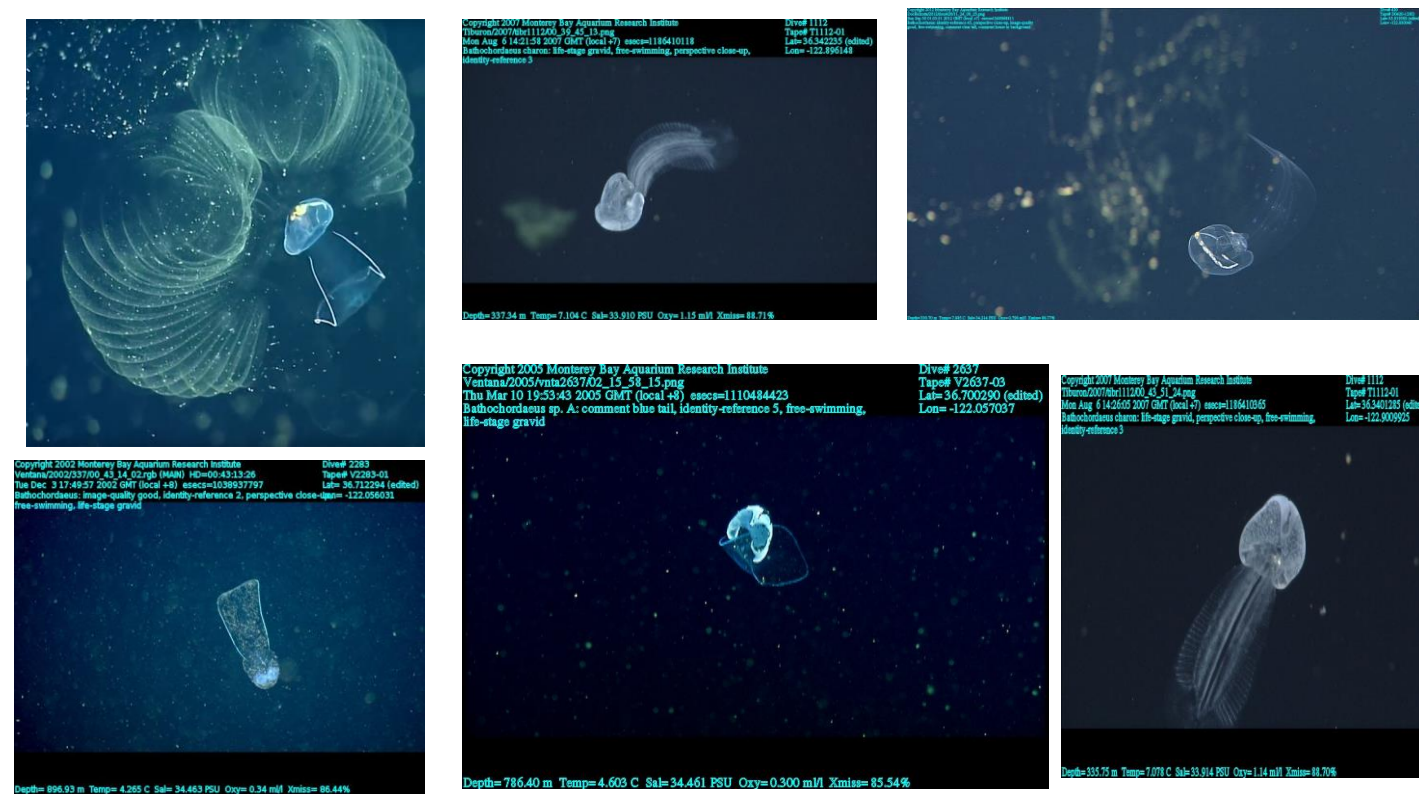
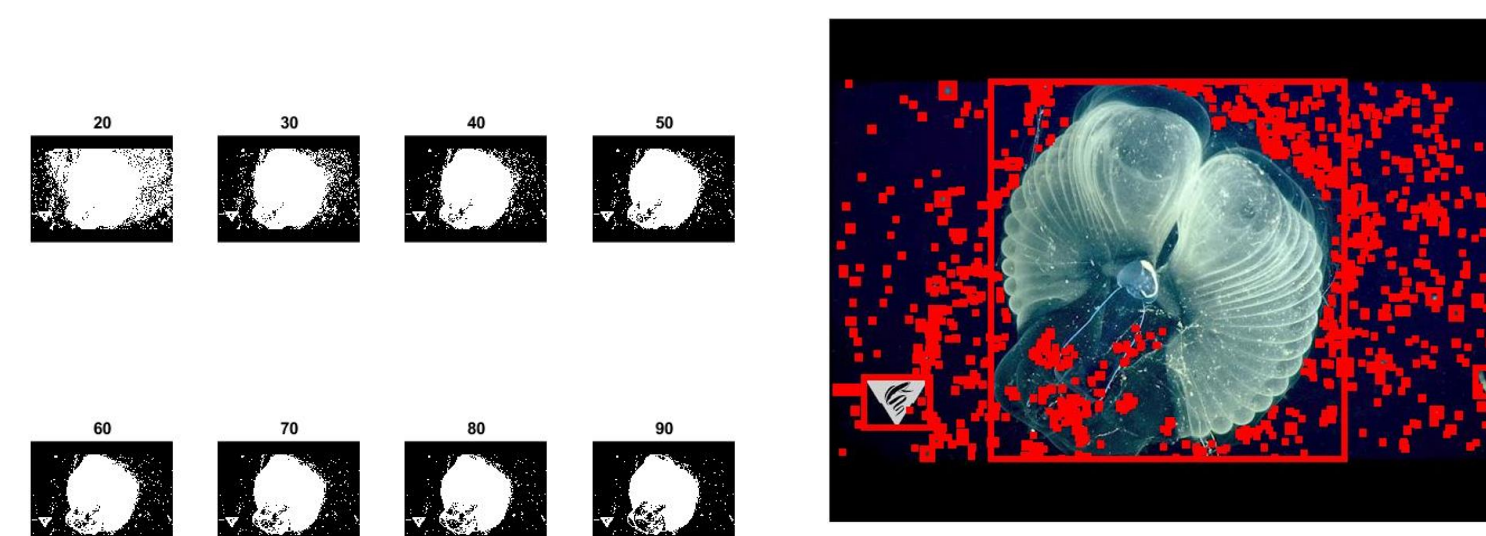


Image Preprocessing

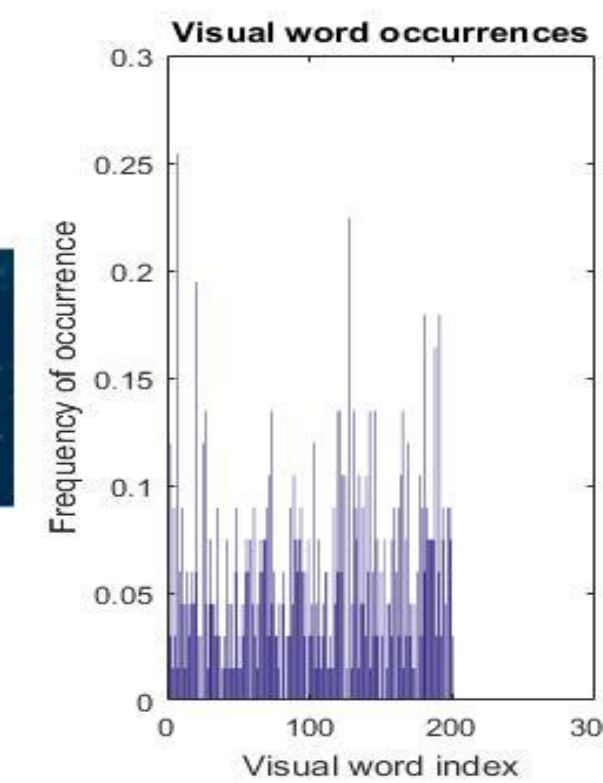
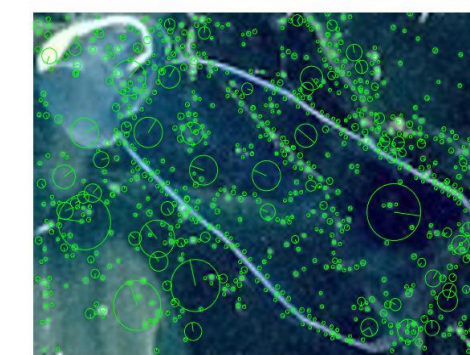
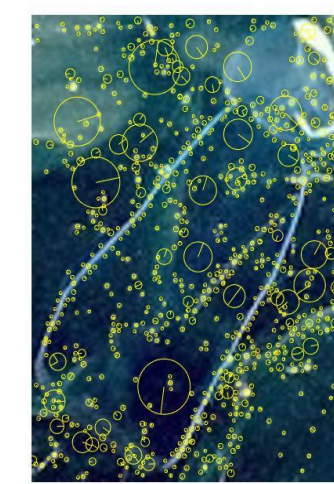
Images of larvaceans are provided by MBARI. Blobs in the images are first separated out by thresholding the image. Blobs that are above minimum threshold area are then selected as the region of interest for larvacean detection.



Feature Derivation

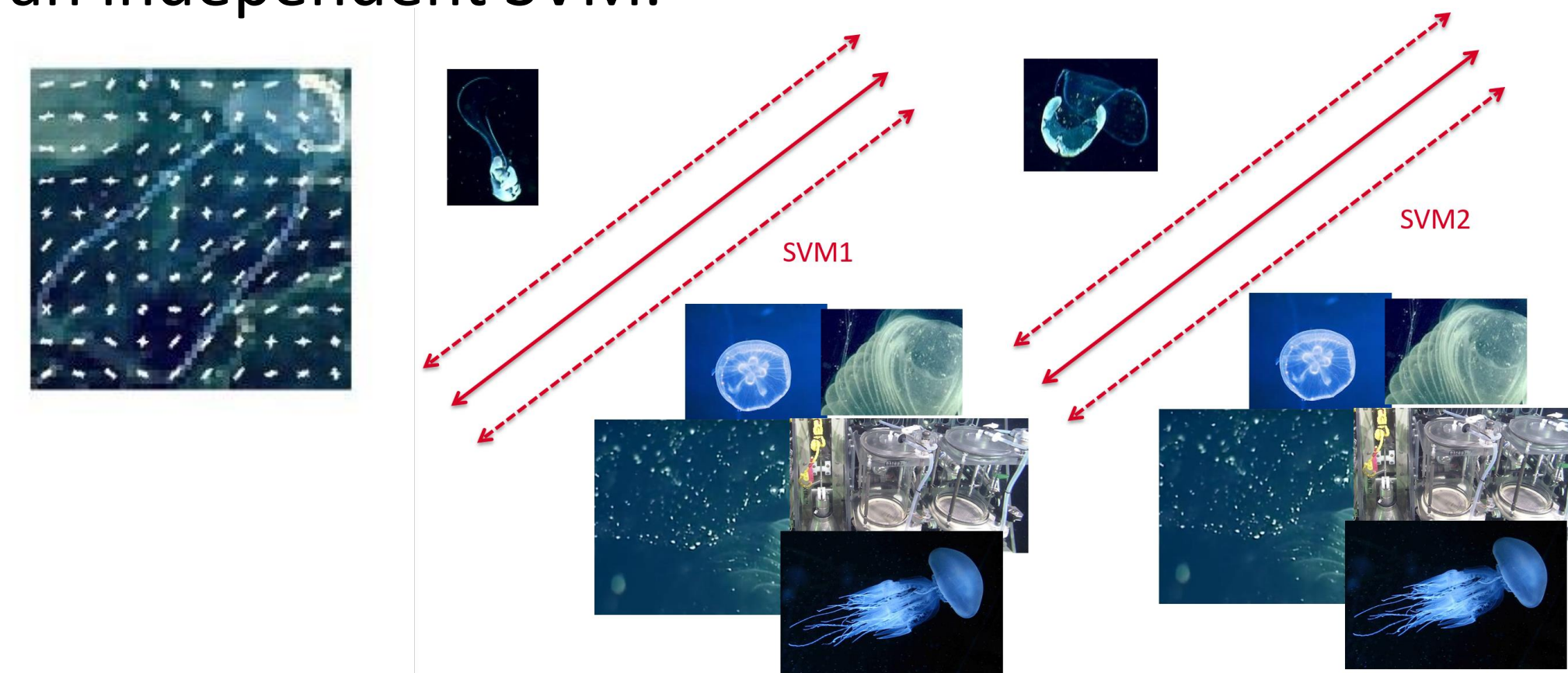
SIFT features as "Visual Words":

Scale invariant feature transform (SIFT) features are used because it is scale and rotation invariant. Compare to SURF (speeded up robust features), SIFT is more robust regarding scale, orientation, blur, affine changes [1]. K-means clustering was used to find the centroid of clusters of SIFT features (pick 'visual words'). The final feature vector of an image is a normalized histogram of occurrence of the visual words.



HOG features as exemplars:

Histogram of gradient (HOG) features are also considered, because HOG features preserves spatial information. Image input are all resized to 50x50 pixels in order to get HOG feature of same length. For orientation dependency of HOG, each orientation in the positive data pool is used as an exemplar to train an independent SVM.



Reference:

- Juan, L., Gwon, O.: A comparison of SIFT, PCA-SIFT and SURF. In: International Journal of Image Processing (IJIP), vol.3, issue 4 (2009)
- Malisiewicz, T., Gupta, A., Efros, A.A. : Ensemble of Exemplar-SVMs for Object Detection and Beyond. In: 2011 IEEE International Conference on Computer Vision, 978-1-4577-1102-2/11/ Image Courtesy of MBARI

Model Training

Two class SVM on SIFT features:

92 positive and 500 negative samples are mapped into SIFT features space and trained a 2 class SVM.

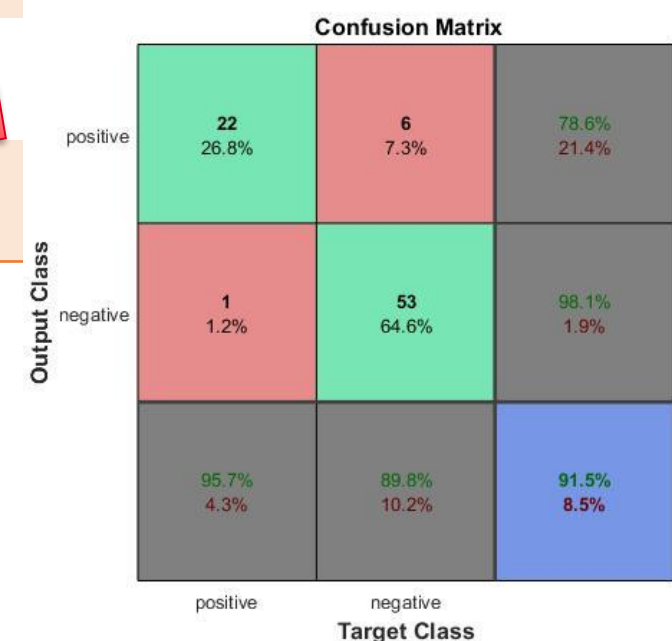
Exemplar SVM on HOG features [2]:

An independent SVM is trained on each positive example against all other negative examples. Then hold out validation is used to calibrate across exemplars using fitted sigmoid function as posterior probability transformation function. Exemplar scored higher across all other positive exemplars will have the decision boundary shifted towards it and vice versa.

Model Performance

Two class SVM on SIFT features:

Number of clusters	Training accuracy	Test accuracy
100	88.6%	82.9%
200	92.7%	91.5%
500	94.3%	90.2%



Exemplar SVM on HOG features:

Training accuracy	Test accuracy
84.4%	76.9%

Example result

