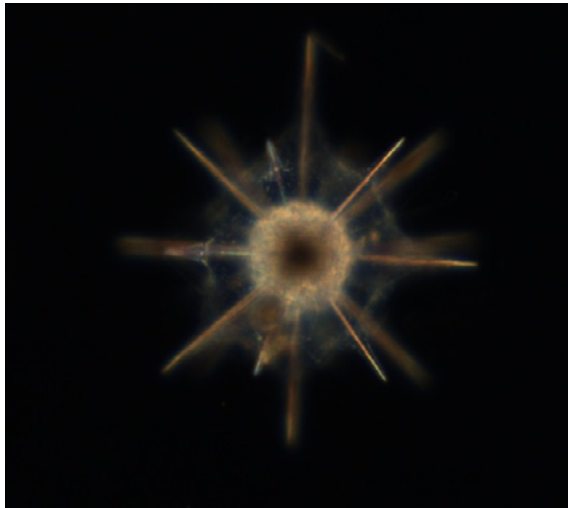# *The freest of lunches:*
Using out-of-domain data to boost oceanographic image classification

Eric Orenstein

University of California, San Diego

May 23, 2018
11:00 a.m. – Pacific Forum

Over the past decade, the biological oceanographic community has increasingly relied on in situ digital imaging to sample the denizens of the sea. These data sets have grown intractably large, requiring countless hours of human labor for analysis. Oceanographers have begun to leverage advances in machine learning to automate the process. In this talk, I will outline ongoing efforts in the Jaffe Laboratory for Underwater Imaging to speed classification of data from the Scripps Plankton Camera. I will focus on experiments using out-of-domain data to boost the performance of machine classifiers and present two early time series analyses—one examining a parasite-host relationship, the other tracking the occurrence of chain-forming diatoms. Data from other sources do indeed improve accuracy, but more work remains to judge the quality of the labels and develop consistent annotation pipelines. There is, after all, no such thing as a free lunch.

MBARI Monterey Bay Aquarium Research Institute

Monterey Bay Aquarium Research Institute, 7700 Sandholdt Road, Moss Landing, Ca 831-775-1700