OCEAN OPTICS XXIV
Valamar Lacroma Dubrovnik Hotel | Dubrovnik, Croatia | October 7–12, 2018
https://oceanopticsconference.org

Monday, October 8
Oral Session 2
14:00–16:00

15:20–15:40

ADVANCING MACHINE LEARNING FOR AUTOMATED TAXONOMIC CLASSIFICATION OF PLANKTON IMAGES FROM IMAGING FLOWCYTOBOT

Recent advances in machine learning have shown promise in improving accuracy of automated detection and classification of marine organisms from high-volume data such as images and video. Newer techniques from deep learning such as convolutional neural networks (CNNs) are outperforming virtually all other image classification approaches, and offer benefits such as eliminating the need for feature engineering. We compare several machine learning approaches on a large dataset of phytoplankton imagery produced by Imaging FlowCytobot at the Martha's Vineyard Coastal Observatory: 1) support vector machines (SVM), 2) random forest classifiers (RFC), 3) multilayer perceptrons (MLP), and 4) CNNs. All four approaches are supervised and take expert human taxonomic classifications as training inputs. The SVM, RFC, and MLP approaches classify images based on a large ensemble of hand-engineered features computed for each image with a variety of image processing algorithms. In contrast, the CNN approach operates directly on images and learns image processing features as part of the training process. In our experiments with large multi-class problems, we find that CNNs fully trained on plankton images outperform other approaches, not only those based on hand-engineered features, but also CNNs that are pre-trained for generic object recognition tasks and then partially retrained to classify phytoplankton. This work is situated in an operational workflow that supports near-realtime processing and high throughput retrospective processing to handle the dataset of ~880m images. Accurate performance for such large datasets represents a Big Data challenge that requires high performance computing resources in addition to state-of-the-art algorithms.

Heidi Sosik, Woods Hole Oceanographic Institution, hsosik@whoi.edu, https://orcid.org/0000-0002-4591-2842
Joe Futrelle, Woods Hole Oceanographic Institution, jfutrelle@whoi.edu
Emily Peacock, Woods Hole Oceanographic Institution, epeacock@whoi.edu
Tyrone Lee, Woods Hole Oceanographic Institution, tlee@whoi.edu