

Detection of Anomalous Particles from the Deepwater Horizon Oil Spill Using the SIPPER3 Underwater Imaging Platform

Sergiy Fefilatyev¹, Kurt Kramer¹, Lawrence Hall¹, Dmitry Goldgof¹, Rangachar Kasturi¹,
Andrew Remsen², Kendra Daly²

II. DESCRIPTION OF PLATFORM AND

The SIPPER [4] was developed by the Technology at the University of South purpose of monitoring the composition, structure of plankton and other suspended p environments. The SIPPER uses c illumination and a high speed line continuously image particles and plankt through a 10cm × 10cm sampling ap continuously scanning line scan camera ca are 10 cm in width and continuous in leng particles that enter the sampling tube are i as a single large SIPPER file with conc environmental data, such as tempera embedded within the SIPPER file. A deployment can result in hundreds of thou of individual extracted particle images (lar equivalent spherical diameter (ESD)).

Custom designed software, the i PI Classification Extraction System (PICES), quickly extract, classify, manage and anal plankton images. A database management PICES allows management of the large

For all datasets, the number of classes was set to 36,

Similarly, the *Oil Large Test Set* is a category of 30 droplet class, particles of particular interest for this research. Each dataset in the category was obtained by the category other represents the classification of all other including all predicted and validated data that passed through particles compared against oil droplets. Thus, every the following filters. Instances of the data that are a part of prediction in favor of one of the other 35 classes of the the *Oil Original Set* were removed. 5,000 images of oil datasets is summarized into the other category. We do not droplets, selected randomly, were removed for future use for validation. Another 1,072 oil droplets used for building a particular dataset, within the category of the *Original Replaced Set*, were removed as well. They are the 1,072 used to build the classifier being tested on this data set. Each dataset within the *Oil Large Test Set* category had 36 classes, 43,816 images total, of which 13,858 were oil droplets. The results of experiments where the category *Oil Large Test Set* was used are reported by averaging the performance from 30 individual classifiers within the category. For a particular experiment the classifier is trained on the *Original Set Replaced* and tested on the *Oil Large Test Set*, such that instances of the oil droplet class in training and test datasets don't intersect.

The datasets *Oil Set Original*, *Oil Large Test Set*, and *Oil Random Set* did not intersect. Datasets *Oil Set Original* and *Oil Set Original Replaced* intersected for instances of all classes except oil droplets.

VI. EXPERIMENTS

In our experiments, we report the accuracy of classification in the form of a 2x2 confusion matrix, as if we were doing binary classification, although the setup of the experiment itself was not binary. One class was the oil

discriminate the oil droplet class from others were related to the circularity of the shape, and the texture of the particle. However, it turned out it is not quite enough for completely

The false positive rate was less than 3.4% in all experiments with *Classifier II*, which was trained on a random selection of oil examples. We did an experiment with a randomly chosen test set, whose distribution mimicked what would be expected during the cruise (about 0.5% oil). For that dataset, oil droplet detection was just 63%. It is also the case that in water where there was no oil, our classifiers predicted that a small amount of oil was present.