Detecting oil spills from space

Linda Corucci, Fabio Nardelli, and Marco Cococcioni

The potential of oil-spill classification from multispectral data is enhanced by a new technique based on using support-vector machines.

The presence of oil spills at sea is of increasing public concern, since they cause serious damage to the marine ecosystem. Detection of oil spills using satellite images is an economical and straightforward way to monitor large areas simultaneously. Many government institutions already use synthetic-aperture-radar (SAR) technology for oil-spill detection. However, the high false-alarm rate renders these systems insufficiently reliable, so that many detected spills are followed by no action because of the potential risk of expensive in situ missions that could be false alarms.

On the other hand, detection of oil spills using optical satellite images allows for large-area monitoring and remote-zone control, providing more frequent information compared to SAR imagery. A new approach based on optical data could be used either on its own or in support of SAR-based solutions to meet the need of environmental-protection authorities for efficient and cost-effective monitoring tools. Moreover, the possibility of detecting oil spills by optical satellite sensors has already been demonstrated. Our approach to oil-spill classification from multispectral satellite data consists of constructing an online classifier based on an ensemble of cost-oriented support-vector machines (CO-SVMs). We applied our classifier to a data set of more than 300 oil spills and natural phenomena in the Mediterranean Sea detected by the Moderate Resolution Imaging Spectroradiometer (MODIS) on NASA’s Terra and Aqua satellites during 2008 and 2009 (see Figure 1).

Online learning allows improvements of the system’s classification capability when an existing data set is updated by oil-spill candidates that have been detected in new images. Once a data set of oil spills and false alarms (so-called lookalikes) has been collected, every time a new satellite image is downloaded and analyzed, the newly detected candidates are included. The classification system will quickly learn to discern real from false events based on analysis of these new candidates.

Moreover, since false alarms represent a significant problem, we have exploited cost-oriented classification (optimal classification with respect to a chosen cost index for misclassification), assigning different costs to misclassification errors for the oil-spill and lookalike classes. Combining cost-oriented classification with the online learning approach, we can use time varying-costs and thus change the desired target conditions (in particular the maximum false-alarm rate) according to the latest classification results. In an operational scenario, this means that, for instance, if the coast guard verifies that too many lookalikes have been labeled real oil spills, the misclassification cost for the lookalikes can be increased, thus improving the classification of future events.

More precisely, we have integrated CO-SVMs and the incremental/decremental formulation of SVMs (ID-SVMs) into a unique framework, COID-SVMs. We used an ensemble of five COID-SVMs, each with its own static misclassification-error costs. We handled the online-learning approach by introducing a sliding window for data acquisition, which happens incrementally. The SVM structures can be dynamically modified (i.e., they are not fixed by the dimension of the data acquired in batch mode). This intrinsic dynamicity improves the system’s adaptability to time-varying conditions.

We adopted the receiver operating characteristic (ROC) convex hull as method for evaluation of classifier performance and exploited the technique of Flach and Wu to repair concavities in ROC curves to improve the ROC convex hull. Figure 2 shows the results obtained by applying the classification system to the complete data set.

Continued on next page
Figure 2. Receiver operating characteristic curves for the ensemble of five online cost-oriented and incremental/decremental support-vector machines (COID-SVMs). The violet and black curves represent the convex hull and the iso-performance line used to find the optimal threshold for the classifier, respectively. The figure shows the convex hull for the ensemble (violet line) and the improved solution (purple line) obtained by repairing concavities. The red square represents the optimal threshold obtained after repairing concavities. FPR, TPR: False-, true-positive rate.

In summary, we have developed an online classifier based on an ensemble of CO-SVMs and applied it to a data set of more than 300 oil spills and natural phenomena detected in the Mediterranean Sea. The promising results that have been achieved highlight the potential of using optical satellite data for oil-spill classification. We will next seek further performance improvements of our online classifier.

We acknowledge funding from Flyby S. r. l.

Author Information

Linda Corucci
Dipartimento di Ingegneria dell’Informazione: Informatica, Elettronica, Telecomunicazioni
University of Pisa
Pisa, Italy

Linda Corucci is a particle physicist. She graduated with honors from the University of Pisa in 2007. She is currently a PhD candidate in information engineering. Her research interests include oil-pollution detection from multispectral remotely sensed images and pattern recognition.

Fabio Nardelli
Flyby S. r. l.
Livorno, Italy

Fabio Nardelli is a computer-science engineer. He graduated with honors from the University of Pisa in 2007. Since 2007, he has been working as a software engineer. His interests include intelligent and embedded systems engineering.

Marco Cococcioni
Dipartimento di Ingegneria dell’Informazione
University of Pisa
Pisa, Italy

Marco Cococcioni received his PhD in computer engineering from the University of Pisa in 2004. He is currently a visiting scientist at the North Atlantic Treaty Organization’s Undersea Research Center (La Spezia, Italy). His research interests are in computational intelligence and decision-support systems.

References