Studying the Sky/Planets Can Drown You in Images:
Machine Learning Solutions at JPL/Caltech

Usama M. Fayyad
Machine Learning Systems Group
Jet Propulsion Laboratory
California Institute of Technology

Today's science instruments can collect huge amounts of image data, making traditional human-based comprehensive analysis an infeasible endeavor. Our long-term goal is to develop a domain-independent system capable of small-scale object recognition in large image databases for science analysis. We view this effort as a step towards a new generation of analysis tools where users specify what to search for simply by providing the system with training examples, and letting the system automatically learn what to do. The system would then automatically search all images and produce a catalog of objects of interest for analysis purposes. Two applications at JPL will be covered. The first targets automating the cataloging of sky objects in digitized sky survey consisting of three terabytes of image data and containing on the order of three billion sky objects. The Sky image ('cataloging and Analysis Tool (SKICAT) developed at JPL achieves a sky object classification accuracy of about 94%, allowing accurate scientific analysis and theory verification/refutation. The learning algorithms played an integral and enabling role by automating the classification and cataloging of hundreds of millions of objects, the majority of which being too faint for visual recognition by astronomers, and by recognizing objects that are at least one isophotal magnitude fainter than any objects cataloged in large-scale photographic surveys to date. The second part of the talk will cover JARtool (JPL Adaptive Recognition Tool) targeting the detection and cataloging of over an estimated 1 million small volcanoes visible in the Magellan SAR database of over 30,000 images of Venus. The techniques described are applicable to a wide range of domains since the problem of visual pattern recognition appears in many domains, including medical and biological imaging, industrial product inspection, surveillance as well as astronomy and planetary science.

NOTES:
SKICAT work is in collaboration with N. Weir and S. Djorgovski of the Caltech Astronomy Dept.

The work described in this paper/presentation was carried out in part by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.