

New technique accelerates biological image analysis

Researchers in Carnegie Mellon University's Lane Center for Computational Biology have discovered how to significantly speed up critical steps in an automated method for analyzing cell cultures and other biological specimens.

The new technique, published online in the *Journal of Machine Learning Research*, <http://jmlr.csail.mit.edu/>, promises to enable higher accuracy analysis of the microscopic images produced by today's high-throughput biological screening methods, such as the ones used in drug discovery, and to help decipher the complex structure of human tissues.

Improved accuracy could reduce the cost and the time necessary for these screening methods, make possible new types of experiments that previously would have required an infeasible amount of resources, and perhaps uncover interesting but subtle anomalies that otherwise would go undetected, the researchers said.

The technique also will be applicable in fields beyond biology because it improves the efficiency of the belief propagation algorithm, a widely used method for drawing conclusions about interconnected networks.

"Current automated screening systems for examining cell cultures look at individual cells and do not fully consider the relationships between neighboring cells," said Geoffrey Gordon, associate research professor in the School of Computer Science's Machine Learning Department. "This is in large part because simultaneously examining many cells with existing methods requires impractical amounts of computational time."

In many cases, computer vision systems have been shown to distinguish patterns that are difficult for humans to detect, he added. However, even automated systems may confuse two similar patterns, and the confusion may be resolvable by considering neighboring cells.

Gordon and his fellow authors, biomedical engineering student Shann-Ching "Sam" Chen and computational biologist Robert F. Murphy, were able to expand their focus from single to multiple cells by increasing the efficiency of the belief propagation algorithm. The algorithm has become a workhorse for researchers because it enables a computer to make inferences about a set of data by drawing on multiple sources of information. In the case of biological specimens, for instance, it can be used to infer which parts of the image are individual cells or to determine whether the distributions of particular proteins within each cell are abnormal.

But as the number of variables increase, the belief propagation algorithm can grow unwieldy and require an impractical amount of computing time to solve these problems.

The belief propagation algorithm assumes that neighbors — whether they are cells, or bits of text — have effects on each other. So the algorithm represents each piece of evidence used to make inferences as a node in an interconnected network, and exchanges messages between nodes. The Carnegie Mellon researchers found shortcuts for generating these messages, which significantly improved the speed of the entire network.

Murphy, director of the Lane Center for Computational Biology, said this technique could improve the

performance of belief propagation algorithms in many applications, including text analysis, Web analysis and medical diagnosis. For this paper, the researchers applied their techniques to analysis of protein patterns within HeLa cells. They found the technique speeded analysis by several orders of magnitude.

In high-throughput screening processes used for drug discovery and other research, tens of thousands of wells — each containing tens or hundreds of cells — need to be analyzed each day, Murphy said. Automated analysis of the cellular relationships within so many wells would be impossible without the sort of speedups achieved in the new study, he added.

Source: Carnegie Mellon University

This document is subject to copyright. Apart from any fair dealing for the purpose of private study, research, no part may be reproduced without the written permission. The content is provided for information purposes only.