

CellProfiler: Open-Source Software to Automatically Quantify Images

Martha S. Vokes and Anne E. Carpenter
The Broad Institute of MIT and Harvard
Cambridge, MA
anne@broad.mit.edu

Researchers often examine samples by eye on the microscope — qualitatively scoring each sample for a particular feature of interest. This approach, while suitable for many experiments, sacrifices quantitative results and a permanent record of the experiment. By contrast, if digital images are collected of each sample, software can be used to quantify features of interest. For small experiments, quantitative analysis is often done manually using interactive programs like Adobe Photoshop[®]. For the large number of images that

can be easily collected with automated microscopes, this approach is tedious and time-consuming. NIH Image/ImageJ (<http://rsb.info.nih.gov/ij/>) allows users comfortable writing in a macro language to automate quantitative image analysis. We have developed CellProfiler, a free, open-source software package, designed to enable scientists without prior programming experience to quantify relevant features of samples in large numbers of images automatically, in a modular system suitable for processing hundreds of thousands of images.^{1,2}

To analyze images quantitatively, researchers build a pipeline of modules in CellProfiler (Figure 1), where each module performs a function on the images — for example, illumination correction, object identification, and measurement. The modules contain settings that allow the module to be adjusted for a particular project. The flexibility of these modules enables identifying a variety of different types of objects in a variety of types of microscopy images (Figure 2 and 3). Identification modules also can identify sub-compartments of objects (e.g. nucleus vs. cytoplasm, Figure 3). Once objects and sub-compartments are identified, measure modules count objects and calculate measurements for each object, each sub-compartment, or each overall image. These measurements can include, for example, the area, location, and shape of each object, as well as the intensity and texture (smoothness) of the staining in each channel within each object or sub-compartment. To improve the accuracy of intensity measurements, a variety of illumination correction methods are available because even when images have been white-referenced, residual patterns often remain in the images.

After a pipeline is configured, tested on a few images, and run on the full set of images, the measurements are ready for downstream analysis. The measurements are stored in an output file that also contains the pipeline with its settings; this output file can later be used to load the pipe-

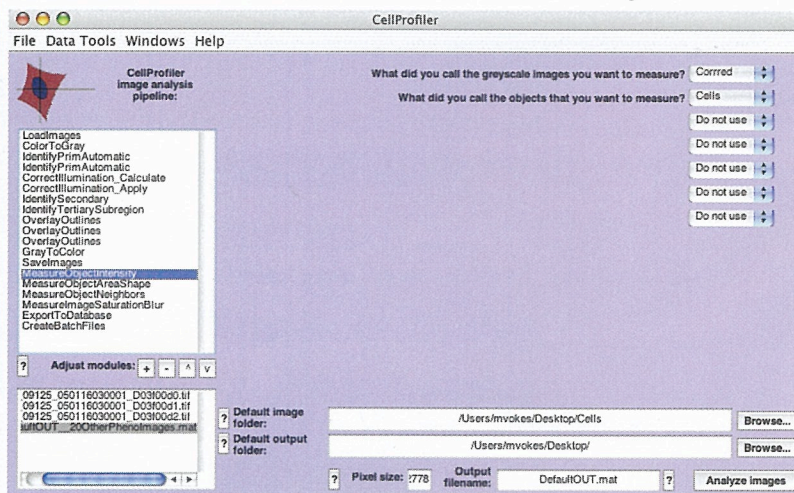


Figure 1: CellProfiler software with a pipeline loaded, showing the series of modules whose functions will be performed on every image to be analyzed. Users specify the location of their images and where to save output files containing the measurements from the analysis.

Figure 2

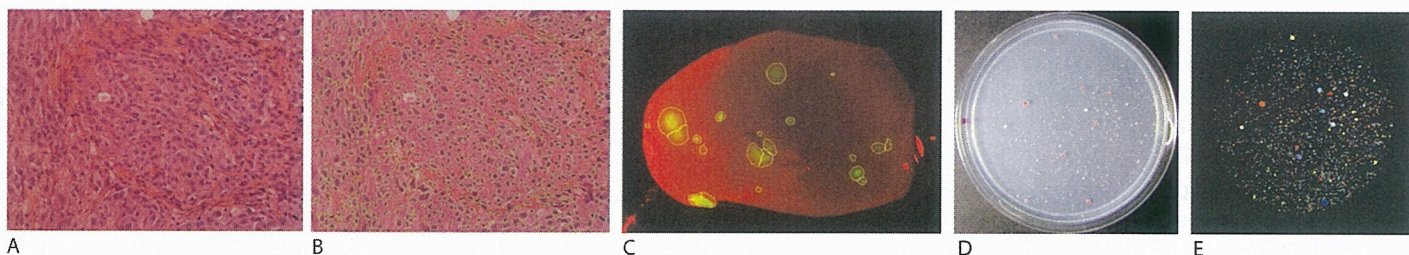


Figure 3

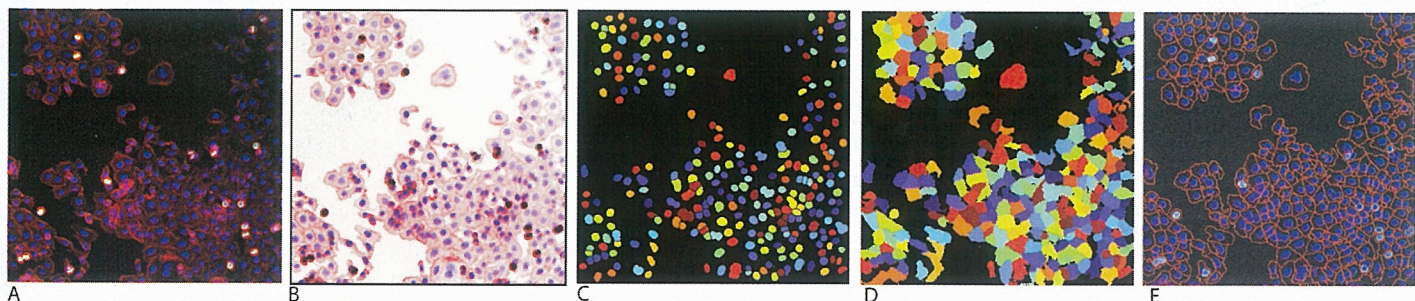


Figure 2: CellProfiler has successfully quantified features of a wide range of objects in images, including: the original image (A) and outlines (B) of cells in histological sections (20x, 330microns by 461 microns), (C) tumors in mouse lungs, and the original (D) and identified (E) yeast colonies grown on a 10 cm wide agar plate displayed in arbitrary colors.

Figure 3 (A) Original pseudo-colored image of nuclei (blue), cells (red), and phospho-histone H3 (green) with the contrast stretched. (B) An inverted pseudo-color image. Identified (C) nuclei and (D) cells displayed with arbitrary colors. (E) Image showing the outlines of each object from C and D.

line and its settings exactly as used in the analysis. The file serves as a permanent record of the analysis and can be shared with colleagues or provided as supplementary data for a published paper.

CellProfiler has tools for exporting measurement data to spreadsheet programs (e.g., Microsoft Excel®) or databases (e.g., MySQL or Oracle), creating histograms and scatter plots, and performing calculations on the data. A separate open-source software package, CellProfiler Analyst, is available for data exploration and visualization of large-scale sets of CellProfiler measurements that are stored in a database (Jones TR, Kang IH, Wheeler DB, Lindquist RA, Papallo A, Sabatini DM, Golland P, Carpenter AE, submitted). Researchers can also use the Classifier function of CellProfiler Analyst to classify objects based on complex visual features. Briefly, if researchers want to count the number of "positive" and "negative" objects in each image but cannot determine which measurement(s) are able to distinguish positives from negatives (e.g., in those cases where the distinction is complex or subtle), they provide examples of positive and negative objects and the software learns to tell the difference between them (Jones TR, Carpenter AE, Lamprecht MR, Moffat J, Silver S, Grenier J, Root D, Golland P, Sabatini DM, submitted).

CellProfiler and CellProfiler Analyst software, plus example pipelines and an online forum for questions, can be found at the CellProfiler website (<http://www.cellprofiler.org>). CellProfiler can be used to analyze hundreds of thousands of images using a computing cluster, or a smaller number of images (up to ~100) on an individual computer. It is compatible with images and movies (grayscale or color) from most microscopes (Table 1) and is

Table 1: Image and movie file formats that can be analyzed by CellProfiler.

	File formats
Images	BMP, CUR, FIG, GIF, HDF, ICO JPEG, JPG, MAT, PNG, PBM, PGM, PNM, RAS, TIF, TIFF, XWF
Movies	AVI, FLEX, STK

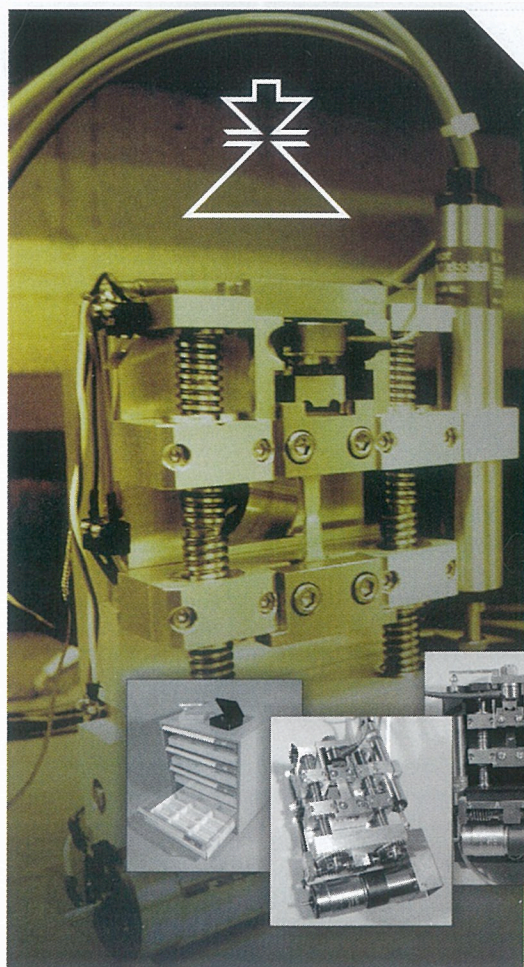
available for Mac, PC, and Unix operating systems. Users with programming experience can modify the source code using Matlab (www.mathworks.com), but compiled versions of CellProfiler that do not require Matlab are also available for download. In addition to the quantitative image analysis described here, CellProfiler also contains modules for image file format conversion, generation of montages and html-based web pages of images, and other useful image utilities. ■

Acknowledgements

We gratefully acknowledge those researchers who have contributed images from Whitehead Institute for Biomedical Research and the Massachusetts Institute of Technology: Nada Kalaany (figure 2 A and B), Kimberly A Hartwell (figure 2 C), and Leah Cowen (Figure 2 E and F). We also thank the members of the Broad Institute Imaging Platform for software development.

References

1. Lamprecht, M.R., Sabatini, D.M. & Carpenter, A.E. CellProfiler: free, versatile software for automated biological image analysis. *Biotechniques* 42, 71-75 (2007).
2. Carpenter, A.E. et al. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol* 7, R100 (2006).



CUSTOMIZING TO YOUR SPECIFIC NEEDS

Micromanipulators, preparation materials, darkroom and general lab supplies, books, grids and apertures. Many items are manufactured in our machine shop, so customizing to your specific need is not a problem.

Some of the accessories and laboratory supplies we can supply are tweezers, tools, TEM CCD imaging systems, tensile testers, turbo evaporators, sputter coaters, substages, specimen holders, standards, carbon coaters,

CONTACT US

750 Pierce Road, Suite 2
Clifton Park, New York 12065-1303
Phone: 518-877-0820, 800-833-4024
Fax: 518-877-3885

ERNEST F. FULLAM, INC.
Microscopy & Laboratory Supplies

sales@fullam.com \ www.fullam.com

