*Available online at* www.scholarsresearchlibrary.com

Extended Abstract

**Journal of Computational Methods in Molecular Design, 2021, 11(2)**
https://www.scholarsresearchlibrary.com/journals/journal-of-computational-methods-in-molecular-design/

**ISSN 2231-3176**

# The cell image library: An open source central repository and research platform for microscopy images

**David Orloff**

*University of California, California*

*E-mail:* dorloff@ncmir.ucsd.edu

The Cell Image Library (CIL) is a public and easily accessible resource database of images, videos, and animations of cells, capturing a wide diversity of organisms, cell types, and cellular processes. The purpose of this database is to advance research on cellular activity, with the ultimate goal of improving human health. This comprehensive and easily accessible Library is designed as a public resource first and foremost for research, and secondarily as a tool for education. The long-term goal is the construction of a library of images that will serve as primary data for research. A discussion of the development of the open source library and a demonstration of the features focusing on large data sets will be presented. Lastly, there will be a discussion of new tools that have been incorporated to interact with the data in the CIL to aid in original research will be presented. Key to the success of CIL-CCDB is its ability to bridge the abstract concepts of cell biology and cell structure with the reality of that seen in the laboratory. This is most notably demonstrated by the interactive cell illustration on the homepage that invites users to 'Explore the Cell'. This feature, unavailable with general search engines such as Google Images, provides easy access to narrative and to images and videos for both non-scientists and experts. By hovering over a cellular feature on the illustration or hovering over the word on the right, users see both the word and all those features highlighted in the illustration. Simply clicking starts an active search for microscopic images of that feature or organelle. Recently added for a number of these features is a second level of interaction. For example, selecting the mitochondrion takes one to a Mitochondrion page with a new interactive illustration of the mitochondrion, a brief explanation of its function, the five most commonly used annotation terms for the Molecular Functions and the five most commonly used annotation terms for the Biological Processes. These annotation terms are from the Gene Ontology. The annotation terms in both cases are hyperlinks to searches of all images with that term in CIL-CCDB. Microscopic images of mitochondria then follow this new information. Thus, the user goes from the illustration to the raw data, emphasizing the types of cell images that give rise to the illustrations one sees in textbooks. While there is a difference of opinion about using terms from Gene Ontology, the advantage is that these position CIL-CCDB to create links directly from an image or video into the nucleic acid and protein databases, Genbank and PDB. The intention is to allow users to move from the cell to the molecules and from the cell to the organism via federation with other databases as has been modeled by the Neuroscience Information Framework (NIF). The more general search engines do not provide these types of connections. In addition to browsing CIL-CCDB with the graphical interface, there is a browse bar located at the top of the page for users to explore CIL-CCDB by different categories, including Cell Process, Cell Component, [(both from the Gene Ontology], Cell Type, [from the Cell Type ontology], Organism [from the NCBI organismal classification ontology] or Recent images. The results pages for this sorting can be customized to sort by a variety of different parameters, including thumbnails of images, alpha order or even the number of images populating a category in the library. These customized sortings are unavailable with the general search engines. The advanced search offers a great deal of power to the user. Images can be searched based on multiple specific attributes such as whether it is a still image, a video or an animation, whether it is 3D with a Z dimension or whether it is a time series. These can be combined for even more selective searching. Selecting both the Z-stack and the Time Series filters will return those images that have data in 'both' the Z dimension of space and the time dimension. These images can be explored further in a multitude of ways. For example, one can move to a certain plane in the Z dimension and watch what happens through time or one can move to a certain time point and move through the different spatial planes. At the top of every page is a simple search box. This search engine is another way to begin accessing the data in CIL-CCDB. After three characters are entered in the search box, suggested terms are present for selection. Only terms that will actually return a result are suggested to prevent the selection of a term with a null result. Molecular cell biology and microscopy have undergone a revolution that led to an explosion in complex, dynamic, high-dimensional im-aging data (Reynaud etal., 2015; Ouyang and Zimmer, 2017). The lack of computational methods to extract information from such rich and high-content data is now becoming a critical bottleneck, and thus the field of cell imaging is in great need of computational sci-entists. However, there is a huge gap between biologists who pro-duce, analyze, and hold the data, and computational scientists whose technical and analytical skills can enable extraction of more information from it. This gap is caused by differences in culture, communication, academic motivation, and reward.One key step toward filling this gap is making cell image data publicly available. Data availability will attract computational scien-tists by exposing them to fresh and challenging problems at the in-terface of computer vision, data science, and cell biology.

***Bottom Note:*** *This work is partly presented at* 6th World Congress on NATURAL PRODUCT & SYNTHETIC CHEMISTRY June 24-25, 2019 | New York, USA.