Challenges in Bio-Molecular Imaging and Information Discovery: Developing a Searchable, Distributed Retinal Image Database

J. Byun[1], L. Wang[1], G. P. Lewis[2], S. K. Fisher[2,3] and B.S. Manjunath[1]

PURPOSE. To develop a digital library of immunofluorescence images of vertebrate retina. The long-term objective is to provide new imaging and information processing technologies allowing the creation of large, distributed image databases that are searchable based on metadata.

METHODS. Digital images of feline retina were generated from tissue sections and wholemounts labeled with different combinations of antibodies using a BioRad 1024 confocal microscope. The database was used to study the response of each cell type in the retina before and after retinal detachment. Machine learning and pattern recognition techniques were used for the automatic annotation of over 1000 of these immunofluorescence images. As a first step, we extracted the MPEG-7 homogeneous texture feature vector from every image block (64x64 pixels) in each image. The support vector machine (SVM) learning method was used for learning the distribution of the feature vectors corresponding to layers of cell bodies and synapses in the retina.

RESULTS. To allow queries to the cat retinal image database, we developed metadata for each image. The metadata consists of file name, status, view, antibody used, the cell of interest and magnification information. The initial result of the SVM trial based on the texture vectors for identifying the outer nuclear layer in the normal cat retina achieved about 90% accuracy.

CONCLUSIONS. This preliminary work demonstrates the possibility of applying information processing techniques to large-scale databases of biological images that currently must be analyzed by visual inspection. Further challenges include: applying these methods to a much larger set of images, improving them, and extending them to time series, multispectral and multimodal images. Such searchable databases will also make it possible to share very large numbers of images that otherwise would remain inaccessible and for an individual to search for and analyze similar data distributed among many different shared databases. This will allow for a broader understanding and interpretation of the data, leading to a more complete and integrated understanding of cellular structure, molecular organization, and function in normal retina and retina altered by injury or disease.

[1]Dept of Electrical and Computer Engineering, [2]Neuroscience Research Institute, [3]Dept of MCD Biology , University of California, Santa Barbara, CA