

Real-Time Protein Image Acquisition and Classification System

Madhav Sigdel¹, Salma Begum¹, Marc L. Pusey², Ramazan Aygun¹

¹ Dept. of Computer Science, University of Alabama in Huntsville, Huntsville, AL, USA

² Dept. of Biological Sciences, University of Alabama in Huntsville, Huntsville, AL, USA

Abstract

In this paper, we describe the design and implementation of a stand-alone real-time system for protein image acquisition and crystallization classification. The goal is to assist crystallographers in scoring crystallization trials. In-house assembled fluorescence microscopy system is used for the image acquisition. The images are classified into three categories (non-crystals, likely crystals and clear crystals). The proposed image classification can be executed in parallel to acquisition. Image classification consists of two main steps – image feature extraction and application of data mining algorithms. Our feature extraction involves applying different thresholding techniques, noise filtering, extracting high intensity regions, applying connected component labeling to separate the objects (blobs) and extracting features related to shape and size of the objects. Classification is done applying a multilayer perceptron (MLP) neural network. We performed our experiments on 2250 images consisting 67% non-crystal, 18% likely crystals and 15% clear crystal images and tested our results using 10-fold cross validation. Our results demonstrate that the method is very efficient (< 3 seconds to process and classify an image) and has comparatively high accuracy. Our system detected 97% of the crystals with overall accuracy of 89.7% accuracy and with 7.5% unnecessary checks for images that did not have crystals. We also compare our results with exhaustive set of binary classifiers of MLP neural networks that include both one-versus-one and one-versus-all classifiers.

Email addresses:

ms0023@cs.uah.edu (M. Sigdel), sb0034@cs.uah.edu (S. Begum), puseym@uah.edu (M. L. Pusey), raygun@cs.uah.edu (R. Aygun)