August 2014

Proposal for Team Project ($5K)

Analysis of the information content of biological images

Participants:

1. Student Name: **Swetha Ramadesikan** E-mail: **sramades@purdue.edu**

Faculty Advisor Name: **R. Claudio Aguilar** E-mail: **claudio@purdue.edu**

Department: **Biological Sciences** Institution: **Purdue University**

2. Student Name: **Kayalvizhi Madhivanan** E-mail: **kmadhiva@purdue.edu**

Faculty Advisor Name: **R. Claudio Aguilar** E-mail: **claudio@purdue.edu**

Department: **Biological Sciences** Institution: **Purdue University**

3. Student Name: **Yucong Zhang** E-mail: **yczhang@purdue.edu**

Faculty Advisor Name: **Mark Ward** E-mail: **mdw@purdue.edu**

Department: **Physics** Institution: **Purdue University**

4. Student Name: **Priyadharshini S Venkatasubramani** E-mail: **priya26@neo.tamu.edu**

Faculty Advisor Name: **Aniruddha Dutta** E-mail: **data@ece.tamu.edu**

Department: **Electrical & Computer Engineering** Institution: **Texas A&M University**

5. Student Name: **Shanmuga Sundar Velkur** E-mail: **svelkur2@jhu.edu**

Faculty Advisor Name: **Michael Shields** E-mail: **michael.shields@jhu.edu**

Department: **Civil Engineering** Institution: **Johns Hopkins**

Faculty Advisor Name	**Faculty Advisor Signature**
1. R. Claudio Aguilar	
2. Mark Ward	
3. Aniruddha Datta	
4. Michael Shields	

August 2014

Proposal for Team Project ($5K)

Analysis of the information content of biological images

Participants:

1. Student Name: **Swetha Ramadesikan** E-mail: **sramades@purdue.edu**

Faculty Advisor Name: **R. Claudio Aguilar** E-mail: **claudio@purdue.edu**

Department: **Biological Sciences** Institution: **Purdue University**

2. Student Name: **Kayalvizhi Madhivanan** E-mail: **kmadhiva@purdue.edu**

Faculty Advisor Name: **R. Claudio Aguilar** E-mail: **claudio@purdue.edu**

Department: **Biological Sciences** Institution: **Purdue University**

3. Student Name: **Yucong Zhang** E-mail: **yczhang@purdue.edu**

Faculty Advisor Name: **Mark Ward** E-mail: **mdw@purdue.edu**

Department: **Physics** Institution: **Purdue University**

4. Student Name: **Priyadharshini S Venkatasubramani** E-mail: **priya26@neo.tamu.edu**

Faculty Advisor Name: **Aniruddha Datta** E-mail: **datta@ece.tamu.edu**

Department: **Electrical & Computer Engineering** Institution: **Texas A&M University**

5. Student Name: **Shanmuga Sundar Velkur** E-mail: **svelkur2@jhu.edu**

Faculty Advisor Name: **Michael Shields** E-mail: **michael.shields@jhu.edu**

Department: **Civil Engineering** Institution: **Johns Hopkins**

Faculty Advisor Name **Faculty Advisor Signature**

1. R. Claudio Aguilar

2. Mark Ward

3. Aniruddha Datta

4. Michael Shields

In order to develop novel and more efficient therapeutic approaches we need to understand disease mechanisms, which in turn requires the collection and analysis of biologically-relevant information. Unfortunately, the complex nature and high-information content of biological data, together with the lack of proper analysis methods make the information gathering and interpretation process very slow and imprecise. For example, following microscopy, image analysis is typically performed by visual examination of hundreds of cells or structures per experiment. This method, in addition to limiting the amount of samples that are to be analyzed, is not accurate and potentially poorly reproducible as it introduces human error and bias. Although some algorithms are currently available, they rely on subjective assumptions and are not suitable for widespread applications.

The *objective* of this student-initiated proposal is **to develop tools to extract and analyze quantitative biological information related to the process of cell migration.**

Cell migration is one of the major cellular process life scientists are focused on. On the one hand, this process is essential for normal embryo development, and indeed we discovered that *defective cell migration is observed in a developmental disease called Lowe Syndrome* (LS)(1). On the other hand, *enhanced cell migration of cancer cells is the cause of metastasis* which currently accounts for 90% of cancer-related deaths (2, 3).

Cell migration requires the coordination of multiple elements of the cellular machinery. This is particularly true for the Golgi apparatus which establishes the direction of migration, and for carriers of proteins and lipids necessary for the migration-required remodeling of the cell membrane (4). Therefore, the availability of information such as area of the Golgi apparatus, number, density and lifetime of protein-lipid carriers in patients' versus normal cells would be crucial for understanding and fighting diseases.

During the recent CSoI Student Workshop "Science of Information for Biological Data" (July 7th-11th 2014; Purdue University, West Lafayette Campus), we assembled two multidisciplinary (computer science/engineering/biology students) and multi-institution (Purdue/Texas A&M/Johns Hopkins) teams that developed the foundations for automated analysis of biological information content in fluorescent microscopy images. Here we propose to articulate both teams in a single effort to produce a unique application that will also incorporate biologically-relevant constraints into the existing module created during the workshop.

Specifically, during the workshop we devised a basic code using the "R"-software package to determine the total area of the Golgi apparatus and developed a *Matlab* routine to address the problem of particle (e.g. protein/lipid carrier) recognition. However, due to the limited time of the workshop, some important issues could not be addressed; such as the automatic detection of biologically relevant boundaries from multiple cells or structures in a single image. These features are crucial to be able to reliably extract data per cell and therefore, to obtain accurate statistical information.

Indeed, the routines that we had devised could not extract data for individual cells, but only for the entire image. In addition, the algorithms' ability to identify individual objects was largely dependent on factors such as noise cancellation procedures and arbitrary thresholds applied by

the user per image. We propose to develop a method that relies on scanning the image to detect maximum and minimum intensities and automatically determine thresholds which will eliminate the introduction of arbitrary values by the user and consequent variation.

We will analyze microscopy images, in which each color represents a different cell component (this is achieved using different cell biology techniques such as immunofluorescence). It should be noted that the color information is stored into different matrices that can be worked on independently or as a group. Some of these components will be used to define a cell boundary; for example, the signal from a membrane marker whose intensity is analyzed pixel by pixel using a raster scan can be used to detect the extent of the cell and mark boundaries once it encounters background (no signal). Other components (e.g., a transport carrier- or a Golgi-marker) would be used as markers to identify organelles and intracellular structures. For this latter purpose, we plan to use additional statistical criteria and already existing routines and algorithms such as Object Based Colocalisation and the Mander's Correlation Coefficient (5, 6).

We anticipate that these improved image analysis algorithms (combined in a single application) will expedite accurate image analysis and data interpretation. We also envision that these routines will be used to analyze time-lapse images to study cell migration, dynamics of transport carriers and organelles or structures in patient vs normal cells. Further, these capabilities will provide a valuable platform to perform drug screens and evaluate the efficacy of candidate compounds for correcting cellular defects. Finally, we hope that these quantitative measurements can become the basis for the development of diagnostic methods based on numerical and biological correlations.

In order to achieve our goal, the team members will be mentored by their PIs who will be providing valuable suggestions and feedback to shape the proposed algorithms. Additionally, the team would regularly update their mentors on progress, potential challenges and use the expertise of their advisors in implementing solutions. All students will be involved in algorithm design and implementation. The biologists in the team will be providing key biological constraints to be incorporated in the code and also help in image acquisition and validating the code in terms of its biological relevance and accuracy, while the rest of the team will be providing Information Theory expertise and implementing algorithms. There will be constant communication between the members for algorithm design and evaluation in terms of its efficiency and accuracy in addressing the problem statement. We plan to meet at least once a month via video conferencing and once in three months in person.

In order for the image analysis tool to be available and accessible by the scientific community, we are seeking to participate in national conferences where we plan to present our final version of the project. We plan on attending a conference that would be held during or after our final meeting, in order to ensure that the project has attained completion. Importantly, we anticipate a collaborative publication describing this development. We are confident that this collaboration would help in creating a versatile biological image analysis tool freely available to the scientific community and can be applied in many research problems involving fluorescence microscopy analysis and not just described area of focus.

References:

1. Coon BG, Mukherjee D, Hanna CB, Riese DJ, II, Lowe M, Aguilar RC. Lowe syndrome patient fibroblasts display Ocrl1-specific cell migration defects that cannot be rescued by the homologous Inpp5b phosphatase. Human Molecular Genetics 2009;18(23):4478-4491.
2. Taketo MM. Reflections on the Spread of Metastasis to Cancer Prevention. Cancer Prevention Research 2011;4(3):324-328.
3. Coon BG, DiRenzo DM, Konieczny SF, Aguilar RC. Epsins' novel role in cancer cell invasion. Communicative & Integrative Biology 2011;4(1):95-97.
4. Friedl P, Gilmour D. Collective cell migration in morphogenesis, regeneration and cancer. Nature Reviews Molecular Cell Biology 2009;10(7):445-457.
5. Lachmanovich E, Shvartsman DE, Malka Y, Botvin C, Henis YI, Weiss AM. Co-localization analysis of complex formation among membrane proteins by computerized fluorescence microscopy: application to immunofluorescence co-patching studies. Journal of Microscopy-Oxford 2003;212:122-131.
6. Bolte S, Cordelieres FP. A guided tour into subcellular colocalization analysis in light microscopy. Journal of Microscopy-Oxford 2006;224:213-232.