

# **SCORE++: Crowd sourced data, automatic segmentation, and ground truth for ITK4**

**Offering Institution:** Harvard Medical School

**Principal Investigator:** Sean G. Megason., Ph. D.

**Estimated Cost:** \$120,000

**Abstract:**

In a separate A2D2 proposal (SCORE: Systematic Comparison through Objective Rating and Evaluation), Marcel Prastawa (Univ. Utah) proposes the creation of an infrastructure for hosting test datasets in MIDAS, new I/O filters in ITK to access this data, and new filters for validating and comparing segmentation results. Here we propose to leverage and extend the SCORE infrastructure by crowd sourcing 1) the collection of a diverse number of image sets to place into the SCORE repository, 2) the organization of a Grand Challenge to develop new ITK filters that process datasets in the repository and will be judged using the new metrics developed in SCORE, and 3) the creation of a system to allow manual segmentation of data in the repository from users anywhere in the world to increase the availability of ground truth.

**Address of offering institution:**

Sean G. Megason, Ph.D.  
Department of Systems Biology  
Harvard Medical School  
200 Longwood Avenue  
Boston, MA 02115  
Phone: (617)432-7441  
Fax: (617) 432-5012  
Email: megason@hms.harvard.edu

**Investigators:**

Sean G. Megason - Harvard Medical School, Boston, MA  
Julien Jomier - Kitware, Chapel Hill, NC

## 1. Aims

This proposal builds upon the work proposed in “SCORE: Systematic Comparison through Objective Rating and Evaluation” by Marcel Prastawa at the University of Utah which describes the creation of an infrastructure for maintaining a repository of example image sets, the creation of ITK readers/writers for this repository, and the creation of classes for comparing segmentation results. The SCORE infrastructure proposed by Dr. Prastawa is an important and long awaited addition to ITK because it will create a system for testing and comparing segmentation algorithms. However, without data in the repository, users of the methods, or ground truth for comparison, the full potential of SCORE cannot be realized.

In this proposal, we describe an effort to leverage and extend the infrastructure created in SCORE called SCORE++. Our SCORE++ effort will capitalize on the biomedical imaging community to get data, methods, and ground truth into ITK. We feel that this “crowd-sourcing” approach is highly effective both to improve ITK itself (community→ITK) as well as in broadening the ITK community of users and developers (ITK→community). Specifically, we proposed to use crowd-sourcing in SCORE++ for the following Aims:

1. **Data:** We will collect, get copyright releases, and full annotations for a diverse collection of image sets and deposit these in the SCORE repository. To achieve this aim this we will capitalize on the connections we have by being in the Harvard Medical School community which represent users of a large diversity of imaging modalities under a single copyright holder.
2. **Grand Challenge:** To both increase the number of ITK users and the number of algorithms in ITK, we will organize a Grand Challenge. The Grand Challenge will be a competition to develop the best segmentation and tracking algorithms. Users will develop segmentation and tracking filters in ITK that process data from the SCORE repository and the results will be judged using the SCORE scoring filters. In this aim we will also extend the SCORE scoring system to *tracking*.
3. **Ground Truth:** A significant challenge in testing segmentation and tracking filters is the paucity of publically available “ground truth”. We will create a system for generating ground truth based on “artificial artificial” intelligence. This system will allow users anywhere in the world to perform manual segmentation in a quality controlled manner with an optional monetary incentive and have the results maintained in a repository. In this aim we will also extend the SCORE scoring system to consensus truth and sparse truth.

## 2. Significance to the Medical Imaging Community

A community is not a community without shared resources and interactions amongst its users. We believe that the crowd-sourcing based efforts proposed here should do a great deal to increase the shared resources and strengthen the interactions of the biomedical imaging community. Specifically, we will increase the availability of test image sets, image segmentation and tracking algorithms, and ground truth in ITK which will be invaluable to biomedical researchers

## 3. Significance to ITK


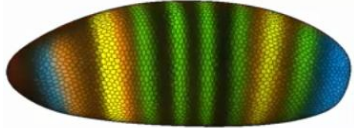
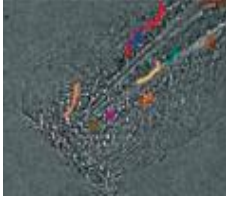


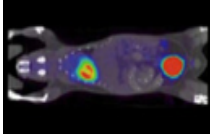
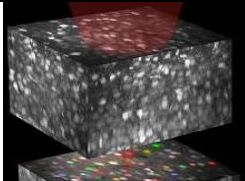
Just as this proposal will encourage ITK to benefit the biomedical imaging community, we also believe that this approach will encourage the biomedical imaging community to benefit ITK via its crowd-sourcing mechanism. This proposal will improve the number and quality of segmentation and tracking algorithms in ITK, the methods for testing and comparing these

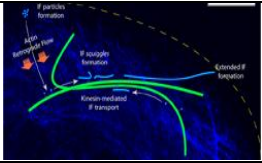
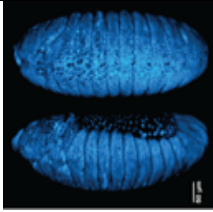
algorithms, and will increase the number of ITK users and ITK's visibility in the biomedical research community.

## 4. Implementation Details and Deliverables

### 4.1. Gathering Image Data Sets for the SCORE Repository

We will collect the following image sets:

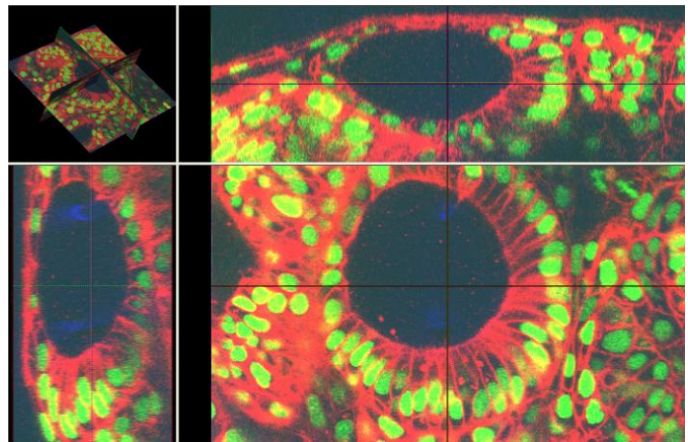
| Lab   | Test  | Imaging Modality                              | Organism   | Labels  | Dimensions  | Example image   |
|---|---|---|--|---|---|---|
| <i>Sean Megason</i> -HMS<br>Dept of Systems<br>Biology    | 3d+t cell segmentation, tracking, and lineaging. Gene expression quantitation               | Confocal fluorescent microscopy               | Live zebrafish embryos                                   | ubiquitous membrane and nuclear markers and tissue specific, time varying transgenic marker | 3 channels x 1024x1024 pixels x 100 slices x 100 time points  |    |
| <i>Angela DePace</i> -HMS<br>Dept of Systems<br>Biology   | registration of gene expression onto a reference embryo                                     | 2-photon fluorescent microscopy               | Fixed <i>Drosophila melanogaster</i> (fruit fly) embryos | in situ hybridization with registration marker and gene of interest                         | 2 channels x 1024x1024 pixels x 300 slices x 10 embryos (each stained for a different gene of interest) |    |
| <i>Tim Mitchison</i> -HMS<br>Dept of Systems<br>Biology   | 2d+t cell segmentation, tracking, and lineaging. Cell division scoring.                     | phase contrast                                | Live animal cell culture on slides                       | Phase contrast  | 1 channel x 1024x1024 pixels x 200 time points  |   |
| <i>Galit Lahav</i> -HMS<br>Dept of Systems<br>Biology     | 2d+t cell segmentation, tracking, and lineaging. Protein localization. Cell death detection | Phase contrast + wide field fluorescence      | Live animal cell culture on slides (cancer cells)        | Phase contrast and transgenic fluorescent reporters for p53                                 | 2 channels x 1024x1024 pixels x 200 time points   |  |
| <i>Roy Kishony</i> -HMS<br>Dept of Systems<br>Biology     | Bacterial colony morphology   | Flat-bed consumer scanner (brightfield)       | Live bacterial cell culture on Petri dishes              | Wild E. coli isolates treated with different drug combinations                              | 1 channel x 1024 x 1024 x 48 antibiotic combinations  |  |
| <i>Ralph Weisleder</i> -HMS<br>Dept of Systems<br>Biology | In vivo cancer diagnostics  | PET and CT                                    | Human and mouse  | Molecular markers for cancer  | 2 channel x 512x512x512 voxels  |  |
| <i>Clay Reid</i> -HMS<br>Dept of Neurobiology             | Connection and activity of neural circuits  | Transmission Electron microscopy and 2-photon | Mouse visual cortex                                      | Osmium; Fluorescent Ca++ reporters  | 4096x4096x4000 (TEM); 1 channel x 512x512 pixels x 500 timepoints (2-photon)                            |  |

|   |   |   |              |                        |   |   |
|---|---|---|--------------|------------------------|---|---|
| <i>Gaudenz Danuser</i> -<br>HMS Dept of Cell<br>Biology | Mapping of<br>molecular<br>movement in<br>cells | 2d+t Speckle<br>fluorescent<br>microscopy | Cell culture | Fluorescent<br>tubulin | 1024x1024 pixels<br>x 200 timepoints                  |  |
| <i>Scott Fraser</i> - Division<br>of Biology, Caltech   | 3d+t cell<br>tracking and<br>lineaging          | 2-photon<br>light sheet<br>microscopy     | Drosophila   | H2B-EGFP               | 1024x1024 pixels<br>x 200 slices x 1000<br>timepoints |  |

These image sets represent 10 different imaging modalities from electron microscopy to fluorescent microscopy to CT and PET, a diversity of different model organisms, and a wide-range of cutting-edge, high-impact, experimental goals and associated image analysis challenges. For each image set, we will work with the producer of the image set to get the data files, a complete annotation of the metadata, and a description of the experimental challenge and the image analysis requirements. We will also work with the producer and the copyright holder to get signed releases for placing these image sets in the public domain. It is possible for us to do this because these are all colleagues (mostly in my department) allowing me to more easily get the relevant data and annotations, and because the copyrights are all held by institutions whose intellectual property offices I've previously worked with on this issue (Harvard University and Caltech). We will convert these files to ITK readable formats and place them in the SCORE repository.

#### 4.2. Grand Challenge using SCORE Data and Methods

We will choose one data set from the SCORE repository to serve as the basis of the first ITK Segmentation and Tracking Grand Challenge. We will most likely choose a 3d+t zebrafish dataset from our lab since we are most familiar with this data, already have expert-generated ground truth, and can create a similar "hidden" dataset for use at contest time. We will provide this image set and the expert ground truth to the public to serve as the Challenge data. The goal of the Grand Challenge will be to 1) segment the nuclei and cell membranes for all the cells for all time points, 2) connect all the cells into tracks representing the movement of cells over time, and 3) connect all tracks into lineages representing the mother-daughter relationship of cells through cell division. We will provide a bare bones partial solution to the problem consisting of GoFigure for visualization of results and mediocre performing segmentation, tracking, and lineaging filters written in ITK. We provide this bare bones solution to help contestants get started and to allow them to only focus on one or two of the goals if they do not want to perform all three. Contestants can use any part of this bare bone solution in their final entry if they wish but can also develop all of their own code. We will also provide a scoring tool so contestants can evaluate the performance of their solutions along the way. All code must be developed in ITK and will become open source at the time of the contest. Contestants may develop novel



Example image set from Challenge data

algorithms or implement previously published algorithms but scoring will only be based on results. The contest will be held as part of an image analysis conference. To enter the contest, contestants will write a ~4 page paper describing their approach and submit the paper and their code to a review panel consisting of us and several other members. We will ensure that the paper adequately describes the methods, that the code builds and executes properly, and measure the performance of their method on the provided Challenge data. The top 6 scoring entrants will have their papers accepted to the conference proceedings and will be awarded \$1000 travel grants to attend the conference. We will encourage all entrants to submit Insight Journal papers. At the conference, we will unveil a new Contest image set captured under identical image settings as the Challenge set but from a different embryo. Contestants will have 24 hours to tweak any parameters and ensure that their code is working on the contest computers, and then a live 2 hour head-to-head contest. During this time, the 6 contestants will each present 10 minute talks on their approach. Winners will be announced following the talks. The winner will be awarded \$1000, second place \$500, and third place \$250 in cash or prizes.

Scoring will be based on accuracy alone with 25% for nuclear segmentation, 25% for membrane segmentation, 25% for cell tracking, and 25% for lineage. Novelty will not be a factor. We will require that the papers document the approach well to be accepted but will not further include this in the score. Speed will not be part of the scoring, but the entrants will only have 2 hours of computational time on a contest machine provided by us. Contest machines will be identical, rented, high-end multicore, multi-CPU machines running the contestants choice of operating systems. Scoring will be done using the scoring algorithms in SCORE. The SCORE proposal intends to creating several segmentation scoring algorithms but no tracking and lineage scoring methods so we will implement the method of Kao, Daggett, and Hurley (“An information theoretic approach for tracker performance evaluation”, ICCV 2009) in ITK. This same pattern can hopefully be followed by others in future years to do ITK Grand Challenges with additional datasets from the repository.

#### **4.3. System for Crowd Sourced Ground Truth using SCORE Data and Methods**

We have already developed an open-source, image analysis application geared toward microscopy called GoFigure (<http://gofigure2.sourceforge.net/>). GoFigure can currently load and display 3d+t, multichannel microscopy data; visualize this data as xy, xz, yz, and volumetrically; and allow user to manually create contours, meshes, tracks, and lineages. GoFigure is cross-platform and is built on top of Qt, VTK, and ITK. We will extend GoFigure using the newly developed IO filters from the SCORE proposal to allow GoFigure to read image sets from the SCORE repository and to write back manual segmentation results to the repository. We will also extend GoFigure to import “HIT requests”. HIT requests will contain the address of an image set in a SCORE repository, a bookmark within that image set, and instructions for the person doing manual segmentation. We will extend the SCORE repository to allow user level security through logging in, to keep track of segmentation results from different users, and to allow for the use of “sparse truth” in scoring segmentation results for quality control purposes. Sparse truth refers to calculating a segmentation score when ground truth is only known for a small minority of objects such as cells. We will also develop “consensus truth” scoring approaches which compares the results from multiple manual segmentations.

With these extensions in place, we will crowd source ground truth as follows: Requesters will post image sets and their own expertly created sparse ground truth into a SCORE repository. In



practice this could be the Kitware maintained repository or additional repositories could be set up if needed. Requesters will post requests for manual segmentations on websites such as Amazon's Mechanical Turk as HITs (Human Intelligence Tasks). These sites provide a mechanism for paying workers to manually perform HITs. Workers that choose the requester's HIT will download a HIT request from the Mechanical Turk website and GoFigure from SourceForge. Requesters will open the HIT request in GoFigure which will cause GoFigure to download the image set from the SCORE repository. The worker will manually perform the segmentation in GoFigure and submit the results to the SCORE repository via GoFigure. The requester will then login to the SCORE website and check the score for the worker's segmentation results to see if they meet quality control standard and if so will award the worker through the Mechanical Turk site.

## 5. Commitment to quality

The Megason Lab has contributed more than 60 core classes to ITK including classes for a new mesh processing framework (QEMesh) and variational level sets for segmentation, and microscopy image processing filters. These have all carried the standard ITK BSD-style license. Their group has published 15 papers in the Insight Journal from 2008-present. The papers include contribution of the variational level set framework, multiphase framework and cell tracking methods. The Megason Lab members worked with the community during every major ITK release in clearing bugs and moving clean code into the ITK repository. They are active members on the ITK-developers and users lists. Further details of their ITK involvement are included as supplementary material. The primary goal of the software engineering efforts of the Megason Lab is the creation of a software application called GoFigure for quantitative analysis of image data. It is open-source with a standard BSD license. The code is freely available on SourceForge and precompiled binaries are available on our website.

## 6. Resources and Personnel

### 6.1 Harvard University Commitment and Facilities

Harvard University has a strong commitment to the distribution of scientific and technical work for the betterment of society. As an academic institution, Harvard's intellectual property policy explicitly states that creative works produced at the University should be "disseminated for the greatest public benefit". For software, open-source is almost always the best way to disseminate intellectual property for the greatest public benefit. In the case of ITK this clearly the case and is strongly supported by Harvard.

### 6.2 Personnel

The collaborators at Harvard University are members of the Megason Lab, namely, Dr. Sean Megason, Dr. Arnaud Gelas. Mr. Nicolas Rannou, and Mr. Evan Schwab. Collaborators at Kitware are Dr. Julien Jomier who is leading their SCORE repository efforts.

## 7. Budget

Total funding requested is \$120,000. We believe this project will yield a lot more "bang for the buck" compared to many A2D2s by leveraging the community through crowd-sourcing. Details are given in the Business Proposal.