ISBI 2019 Challenge Proposal

https://biomedicalimaging.org/2019/call-for-challenges

Title 2
Organizing Team 2
  Contact Persons 2
Space and Hardware Requirements 2

Challenge Description 3
  Problem description 3
  Data Quality and Availability Post-Challenge 3
  Training & Testing Data 6
  Reference Standard 6
  Evaluation Metrics 7
  Website and computing facilities 7
  Publications 7
  Other 7
  References 8

Figure 1. A sample set of sequence of stain histological slices with marked corresponding landmarks.
Title

Automatic Non-rigid Histological Image Registration (ANHIR)

Organizing Team

- **Jiri Borovec**, Department of Cybernetics, Faculty of Electrical Engineering, CTU in Prague, Czech Republic
- **Arrate Muñoz-Barrutia**, Bioengineering and Aerospace Engineering Department, Universidad Carlos III de Madrid, Leganés, Spain
- **Jan Kybic**, Department of Cybernetics, Faculty of Electrical Engineering, CTU in Prague, Czech Republic
- **Gloria Bueno**, VISILAB Group, Universidad de Castilla-La Mancha, Spain
- **Michael Gadermayr**, Institute of Imaging & Computer Vision, RWTH Aachen University, Aachen, Germany
- **Ignacio Arganda-Carreras**, Department of Computer Science and Artificial Intelligence, University of the Basque Country, Spain
- **Oscar Deniz**, VISILAB Group, Universidad de Castilla-La Mancha, Spain
- **Dmitry Sorokin**, Laboratory of Mathematical Methods of Image Processing, Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University, Russia

Contact Persons

- Jiri Borovec, jiri.borovec@fel.cvut.cz
- Jan Kybic, kybic@fel.cvut.cz

Space and Hardware Requirements

One of the previous image registration challenges/benchmarks hosted about 10 teams. For instance, RIRE [1] evaluated 12 methods, EMPIRE10 focusing on CT images [2] evaluated 20 methods, and in CRC [3] participated 3 teams with 9 methods (participants were required to reimplement their methods). The number of image registration methods has increased since then and for many, the implementation is freely available. We will also try to make the entry barrier low. For these reasons, we expect between 10-20 teams to take part.

For the public (training) part of the dataset, the participants will submit registration result to an automatic evaluation system that we will prepare. For the private (test) part, we require to be able to run the image registration method on a standard Linux machine (our server). Regarding conference participation, we would need a standard presentation equipment.
Challenge Description

Problem description

In Digital Pathology, one of the most simple and yet most useful features is the ability to view serial sections of tissue simultaneously on a computer monitor [4–6]. This enables the pathologist to evaluate the histology and expression of multiple markers for a patient in a single review [7]. However, the rate-limiting step in this process is the time taken for the pathologist to open each individual image, align the sections within the viewer, and then manually move around the section. In addition, due to tissue processing and pre-analytical steps, sections with different stains have non-linear variations between the two acquisitions. That is, they will stretch and change shape from section to section. To date, there is not a workable solution to automatically align the serial sections. [8–12]

The proposed challenge focuses on comparing the accuracy and possibly the speed of automatic non-linear registration methods on a set of large images from the same tissue samples but stained with different biomarkers. The registration of these images is a challenging task due to the differences in the appearance model, the repetitive texture of the details and the large image size, between other issues. We plan to evaluate the registrations accuracy on manually annotated landmarks, the method robustness with respect to successful registrations and, as an auxiliary criterion, its processing computation time. We require that all methods run fully automatically, with no image specific parameters.

Data Quality and Availability Post-Challenge

High-resolution (40x magnification) whole-slide images of different types of tissue (lesions, lung-lobes, mammary-gland) were acquired - the original size of our images varies and goes from 15k x 20k up to 45k x 45k pixels. The acquired images are organized in sets of consecutive tissue slices where each slice was stained by a different dye (Cc10, CD31, He, Ki67, proSPC, CNEU, ER, PR). We form pairs of all images within each set, which can be meaningfully registered. Our benchmarking data is composed of 50 tissue cases with different stains (21 train and 29 test) which yields in 722 image pairs (214 train and 508 test). We are currently negotiating with two more institutions in order to provide some more data. For convenience, we also provide downscaled versions of the images in 100%, 50%, 25%, 10% and 5% of the original size.

The task is to register the images within sets. We have marked significant structures in the tissue with landmarks which are spread approximately uniformly over the tissue. Landmarks were manually identified in each image, with correspondences within each set, which allows us to validate the geometric registration accuracy between any two images in each set.

The images will be divided into training and test sets (as described in the following sections). Training images will be freely available from the beginning of the challenge, or are available already, including their landmarks. Test images and their landmarks will be used for evaluation only, on a server node. The evaluation framework is freely available.
Figure 2. Preview of lesions tissue. In rows, we show slices belonging to the same tissue sample and different stains in columns.
Figure 3. Preview of lung lobes tissue. In columns, we show slices belonging to the same tissue sample and different stains in rows.
Figure 4. Preview of mammary glands. In columns, we show slices belonging to the same tissue sample and different stains in rows.

Training & Testing Data

We have split the dataset into two parts - public (214 image pairs) and test (508 image pairs). The public part is available for user experiments. The test dataset will be kept private and used for the final evaluation of the image registration methods.

Reference Standard

In this task, the gold standard is expert annotation - placing landmarks at key points consistently in the whole sequence of stained slices. For each image pair, there are 87 landmarks on average.
Evaluation Metrics

We evaluate the registration performance with respect to geometric accuracy on the manual landmarks. We compute the mean Euclidean distance (displacement error) between reference landmarks and transformed landmarks by each image registration method.

Besides the accuracy, we want to measure execution time. To be able to evaluate the execution time of different methods run on a different computer we set a reference method where we know the execution time on our machine.

Website and computing facilities

We shall create a web-page (probably hosted on https://grand-challenge.org) with a basic automatic registration and evaluation system based on submitted data from the participants.

We would prefer all participant methods to be publicly available so the medical experts can benefit from this wide comparison.

For running on our infrastructure we have available a computing node with AMD K8 2000 processor (16-cores) and 128GB RAM running on Linux.

Publications

We plan to write a summary journal publication with the challenge description and overall evaluations of all participating methods. We may include other methods in the evaluation for which a publicly available implementation is found [13–18]. We will include an extensive discussion on the suitability of different methods and the experimental validation results on the public dataset.

A reduced version of the dataset and a preliminary evaluation framework has already been described in a conference article [19].

Other

- The public part of the used dataset: http://cmp.felk.cvut.cz/~borovji3/?page=dataset
- Landmark annotations for the public dataset: https://borda.github.io/dataset-histology-landmarks
- Proposed image registration framework: https://borda.github.io/BIRL
- Template for the web page of the proposed challenge: https://anhir.grand-challenge.org
References


