Multi-organ Nuclei Segmentation and Classification Challenge 2020

Ruchika Verma*, Neeraj Kumar*, Abhijeet Patil†, Nikhil Cherian Kurian†, Swapnil Rane‡, and Amit Sethi†

*Department of Biomedical Engineering, Case Western Reserve University, Cleveland, OH, USA
†Department of Electrical Engineering, Indian Institute of Technology Bombay, Mumbai, Maharashtra, India
‡Department of Pathology, Tata Memorial Cancer Centre, Mumbai, Maharashtra, India
Correspondence: monusac2020@gmail.com

I. Introduction

In order to assess the variations in the tumors and their microenvironments across organs and patients, identification of nuclei morphologies and classification of their types are essential. In multi-organ nuclei segmentation and classification (MoNuSAC) challenge, the organizers will provide carefully annotated dataset of H&E stained whole slide images of four organs (breast, kidney, lung and prostate) with hand-annotated nuclei boundaries and cell-types. The participants will use the training data of the challenge to build machine learning models to segment and identify the type of cells present in a given whole slide image. The participants are also welcome to use machine-learning-free techniques for their model development provided they solve both segmentation and classification tasks. Subsequently, the participants will be provided with a testing dataset of unseen patients to report the results of their models to the organizers for evaluation. The test annotations will be withheld from the participants and will be used to rank the entries based on the performance of their models. MoNuSAC is an official satellite event of the IEEE International Symposium on Biomedical Imaging (ISBI) 2020.

This document provides information about the challenge timeline, rules and regulations for participating in this challenge, registration procedure, details of training and testing data, evaluation metric, and submission format. A post-challenge journal paper summarizing the challenge outcomes will be prepared after formally concluding the challenge at ISBI 2020 workshop.

II. Challenge Timeline

November 15, 2019: Challenge open for registration (Please see Section III)
December 20, 2019: Training data release (Images + Ground Truth)
February 01, 2020: Testing data release (Images only)
February 25, 2020: Submission of testing results along with a manuscript describing the algorithm and the testing code (Please check the submission instructions in Section VI)

March 10, 2020: Preliminary leaderboard will be released online
April 3, 2020: Declaration of challenge winners at ISBI 2020 challenge workshop

III. Registration Procedure

Prospective participants can register for this challenge by filling in their details in a Google form available here. Challenge participants should adhere to the following rules and regulations, otherwise their participation may be cancelled (without notice) by the organizers at any time during the course of the challenge.

- Anonymous registrations are strictly prohibited.
- Each team is allowed to have at most 5 members but the team registration should be done ONLY by ONE of them. S/he will be the point of contact of the team with the organizers.
- For successful registration each team must provide complete and correct information about the Name of the contact person, Affiliation (including department, university/institute/company, country) and valid E-mail address.
- Participants should carefully select their team names as they will not be changed during the course of the challenge.
- Redundant and incomplete registrations will be removed without any notification.
- Participating teams maintain copyright to the associated intellectual property and software they develop in course of participating in MoNuSAC 2020[1]. The testing code and model parameters submitted during the challenge will be used by the organizers for the sole purpose of model evaluation and will not be released publicly in any form (even with the post-challenge journal paper).

IV. Challenge Data

H&E staining of human tissue sections is a routine and most common protocol used by pathologists to enhance the contrast of tissue sections for tumor assessment (grading, staging, etc.) at multiple microscopic resolutions. Hence, we will provide the annotated dataset of H&E stained digitized tissue images of several patients acquired at multiple hospitals using one of

---

[1] https://monusac-2020.grand-challenge.org/

the most common 40x scanner magnification. The annotations will be done with the help of expert pathologists.

A. Training Data

Training data spanning four organs (breast, kidney, lung and prostate) with cell-boundary and cell-type annotations for epithelial cells, lymphocytes, macrophages and neutrophils was prepared from the whole slide images of 45 patients (scanned at 31 hospitals) downloaded from The Cancer Genome Atlas (TCGA) data portal[2]. Plurality of patients, organs, hospitals and disease states (benign and malignant tumors) will help in learning the morphological variations of diverse cell-types included in this challenge. The sources of such variations may be the differences in the slide preparation protocols adopted at multiple hospitals, patient-specific tumor biology or different developmental stages of the annotated cell-types.

Training set annotations were done using Aperio ImageScope® and were saved in .xml files. Different cell-types were annotated using a unique marker color: epithelial cell were annotated in red, lymphocytes in yellow, macrophages in green and neutrophils in blue. Figure 2 shows an example annotated image. In total, the training data contains 31,411 hand-annotated nuclei instances including 14,539 epithelial cells, 15,654 lymphocytes, 587 macrophages and 631 neutrophils.

The training data, comprising H&E stained images and .xml annotation files, was released to the registered participants on December 20, 2019. If you want to get access to the training data, please register for the challenge (see Section III). The code for reading .xml annotation files can be downloaded by clicking here. Additionally, participants are free to use our previous datasets released as part of MoNuSeg 2018[3] to train a generalized nuclei segmentation module [1], [2].

B. Testing Data

The testing data will be prepared using the similar protocol as adopted for creating the challenge training data. However, the testing set will be created from the patients not included in the training set. The testing data will also contain annotations for the *ambiguous* regions with white boundaries in addition to the usual red (epithelial), yellow (lymphocytes), green (macrophages) and blue (neutrophils) boundary annotations. The ambiguous regions will be the ones which will not be used for computing the evaluation metric for ranking the participants because (1) these regions might have very faint nuclei with unclear boundaries or (2) the annotators might be unsure of their true class. Only H&E stained tissue images and the *ambiguous* region annotations of the testing set will be released to the participants on February 1, 2020. The organizers will evaluate participant's algorithms by using the withheld testing cohort annotations (for the classes of interest).

V. EVALUATION CRITERIA

The metric adopted for evaluating the participant's performance is weighted average of the class-specific panoptic quality (PQ) [3]. We will match each predicted nucleus (p) with the ground truth nucleus (g) if their intersection over union (IoU) is strictly greater than 0.5, this matching will be done separately for each cell-type. For a given class c, the unique matching splits the predicted (p_c) and ground truth (g_c) nuclei into three sets: true positives (TP_c), false positives (FP_c), and false negatives (FN_c), representing matched pairs of segments, unmatched predicted segments, and unmatched ground truth segments, respectively. Given these three sets, class-specific PQ_c (for c^{th} class) will be computed as:

$$PQ_c = \frac{\sum_{(p_c,g_c)\in\{TP_c\}} IoU(p_c,g_c)}{|TP_c| + \frac{1}{2}|FP_c| + \frac{1}{2}|FN_c|} \tag{1}$$

For each image in the test set, the weighted panoptic quality will be computed as the weighted sum of the class-specific panoptic quality, i.e. $wPQ = \sum_{c=1}^{4} w_c PQ_c$, where the weights for the four cell-types are given as follows:

$$w_c = \begin{cases} 1 & \text{for c = epithelial cells or lymphocytes} \\ 10 & \text{for c = macrophages or neutrophils} \end{cases} \tag{2}$$

Macrophages and neutrophils are given more weights in metric computation because of their under-representation in the training and testing sets. The average of the weighted PQs across all images in the testing set will be used as the ranking metric for this challenge.

VI. SUBMISSION INSTRUCTIONS

Each participant should send an email to monusac2020@gmail.com with three attachments (1) zipped file containing prediction masks, (2) A manuscript providing algorithm details and (3) Testing code for evaluating the proposed algorithm, in the format described below.

Prediction mask submission format
1) Create a folder and name it with the patient name.
2) Within the folder created in step 1, create a sub-folder for saving the results of each sub-image
3) Within the sub-folder created in step 2, create a sub-sub-folder to save .mat files for segmented instances of each cell-type. You can name the .mat file as per your convenience. However, make sure that your folder name represents patient name, sub-folder represents sub-image name, and sub-sub-folder represents cell-type.
4) The folder hierarchy mentioned above is depicted in the Figure 2.

The folders containing the results of all testing images should be saved in a single zipped file as "TeamName_MoNuSAC_test_results.zip". Click here to see an example of submission file.
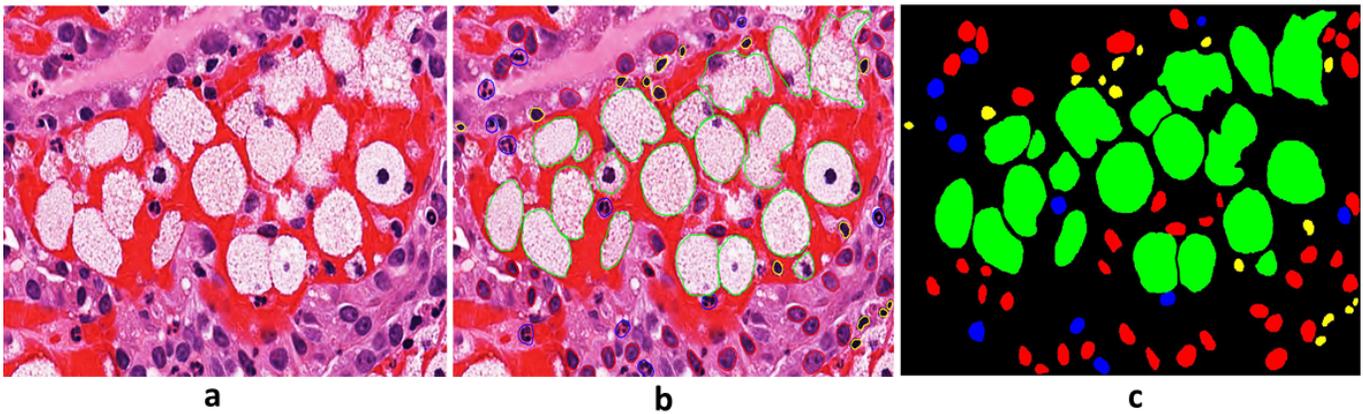
Fig. 1: (a) A sub-image cropped from whole slide image of a patient included in the training set, (b) boundary annotations of different cell-types done using unique marker colors and (c) masks generated from the annotations - epithelial cells are shown in red, lymphocytes in yellow, macrophages in green and neutrophils in blue.
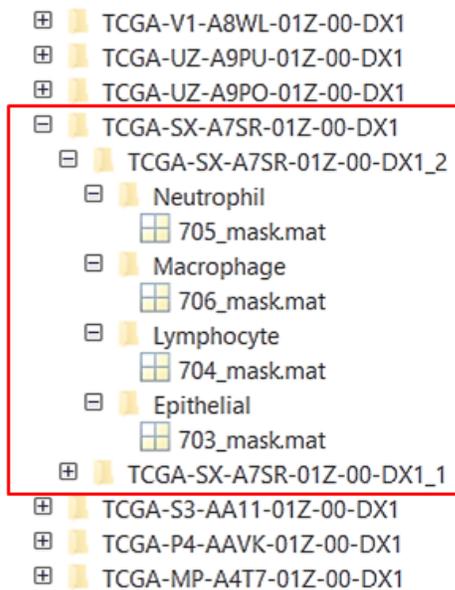

Fig. 2: Folder hierarchy to save the testing results.

Composition of each .mat file

Each mat file should contain instances of only one class. All pixels that belong to a segmented instance should be assigned the same unique positive integer (>0). Hence, the number of unique integers (excluding 0s) within a mat file will correspond to the number of instances saved in that file and 0 will represent the background.

Please click here to see the output file generated for a couple of patients to familiarize yourself with the submission format. The snapshot of the folder hierarchy is shown above.

Manuscript submission format

Each participating team must submit a manuscript describing their algorithm in detail.

- The organizers will review the paper for sufficient details required to understand and reproduce the algorithm and hold the right to exclude participants in case their method description is insufficient.
- An example of a well written manuscript (from our previous challenge - MoNuSeg) is provided on this link. There is no page limit to allow you to give detailed information about every step of your proposed algorithm. Please include a flowchart illustrating your algorithm along with the details of all parameters, hyper-parameters, data augmentation methods, loss functions, etc. used for training the models in your manuscript.

The manuscript should be submitted in a .pdf format with name "TeamName_MoNuSAC_manuscript.pdf".An example of a well written manuscript (from our previous challenge - MoNuSeg) is provided on this link.

Testing code submission format

Each team should submit a Python script and trained model weights to facilitate organizers to evaluate their algorithms. The submitted script should contain commented lines to load the trained models, testing data and saving the predictions (in the prediction mask submission format given above). The testing code should also clearly mention the libraries required to run the models.

The participants might have trained their models using other programming languages but they should make sure that their models can be readily tested in Python using popular machine/deep learning libraries.

* The testing script should be named as "Team-Name_MoNuSAC_testing_code.py".

* Participating teams maintain copyright of the associated intellectual property and software they develop in course of participating in MoNuSAC 2020. The testing code and model parameters submitted during the challenge will be used by the organizers for the sole purpose of model evaluation and will not be released publicly in any form (even with the post-challenge journal paper).

Please send one email to monusac2020@gmail.com with subject "TeamName MoNuSAC Submission" and 3 afore-mentioned attachments to complete your participation in the challenge.

Anyone who does not follow these submission instructions will be disqualified.

VII. Leaderboard Announcement

The preliminary leaderboard will be announced publicly during the challenge workshop at ISBI 2020. The reproducibility of top 3 techniques will be ensured before the ISBI workshop to declare the challenge winners. The final leaderboard will be released on the challenge webpage after verifying the results of all participants.

References

[1] N. Kumar, R. Verma, D. Anand, Y. Zhou, O. F. Onder, E. Tsougenis, H. Chen, P. A. Heng, J. Li, Z. Hu *et al.*, "A multi-organ nucleus segmentation challenge," *IEEE transactions on medical imaging*, 2019.

[2] N. Kumar, R. Verma, S. Sharma, S. Bhargava, A. Vahadane, and A. Sethi, "A dataset and a technique for generalized nuclear segmentation for computational pathology," *IEEE Transactions on Medical Imaging*, vol. 36, no. 7, pp. 1550–1560, July 2017.

[3] A. Kirillov, K. He, R. Girshick, C. Rother, and P. Dollár, "Panoptic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 9404–9413.