Experiments for Gland Segmentation

Please download the new dataset from the gland segmentation contest webpage.

Metrics to be computed per experiment (three metrics):

1. Pixel level accuracy [1]
   1. Jaccard index
   2. Dice index
   3. F1 score
2. Object level accuracy [1]
   1. Object level Jaccard index
   2. Object level Dice index
   3. F1 score

Experiments to perform [1] (**We have to implement other methods for comparison**):

1. Test on following three sets (We have to generate these test sets)
   1. whole test data (benign and malignant together) – (test\_A and test\_B combined)
   2. Benign only (select benign images from whole test data)
   3. Malignant only (select malignant images from whole test data).
2. Healthy and adenomatous samples: The experiment should be tested on 37 test from the Warwick-QU dataset (list of images - exp1.txt).
3. Generalization on Bilkent dataset: Training on the Bilkent dataset and testing on the Warwick-QU dataset.
4. Moderately differentiated samples: The experiment should be conducted on 47 moderately differentiated test images from the Warwick-QU dataset (list of images - exp3.txt).
5. Moderately-to-Poorly and Poorly Differentiated Samples: We should conduct the experiment on 20 moderately-to-poorly differentiated and 24 poorly differentiated test images (list of images -exp4.txt).
6. Healthy samples: All 72 images in the Bilkent dataset are assessed as healthy by the experts. The 24 images are used for training and the rest for testing.
7. Execution times: Average test time on the test dataset of step 2.

Note: A metric per experiment is computed as average (mean +- standard deviation) across the test images in that step.

References:

[1]. Sirinukunwattana, Korsuk, David RJ Snead, and Nasir M. Rajpoot. "A stochastic polygons model for glandular structures in colon histology images." *Medical Imaging, IEEE Transactions on* 34.11 (2015): 2366-2378. <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7109172>