Toward Consistent Cell Segmentation: Quality Assessment of Cell Segments via Appearance and Geometry Features

Andrew Brinker[1], Annika Fredrikson[2], Xiaofan Zhang[3], Richard Souvenir[3], Shaoting Zhang[3]

[1]California State University, CA, US
[2]Principia College, Elsah, Illinois, US
[3]University of North Carolina, Charlotte, NC, US

ABSTRACT

Computer-Aided Diagnosis (CAD) systems based on histopathological images rely on quality low-level image processing, including cell segmentation. Many methods for cell segmentation lack in generality and struggle with the wide variety of cell appearance and inter-cell structure present in histopathological images. We present a computationally efficient system to classify segmentation results as the first step toward automatic segment correction. This general method can applied to existing or future cell segmentation methods to provide corrections for low-quality results. Specifically, with a small collection of easy-to-compute features, we can identify incorrect segments with a high degree of accuracy, which then can be used to determine the needed corrections based on the type of segmentation failure present.

Keywords: Cell segmentation, breast cancer, histopathological images

1. INTRODUCTION

Breast cancer is the second most common form on cancer in the United States today, afflicting 220,097 women and 2,078 men in 2011.[1] Fortunately, early diagnosis and treatment can significantly increase the survival rate among patients. In this context, Computer-Aided Diagnosis (CAD) systems, which have already shown promise for the diagnosis of osteoporosis and intracranial aneurysms,[2] have risen in an effort to improve diagnostic consistency relative to double-reading practices and reduce the time necessary for accurate diagnosis.[3–9] For CAD systems targeting breast cancer, cell segmentation of histopathological images is an important component.[10–17]

Previous work on CAD systems for breast cancer diagnosis have taken a variety of approaches. Using high-resolution images taken from breast cancer biopsy slides, these methods attempt to construct actionable models of the diseased regions in order to differentiate types of breast cancer: usual ductal hyperplasia (UDH), atypical ductal hyperplasia (ADH), or ductal carcinoma in situ (DCIS).[18,19] Petushi et al.[20] developed a system for identifying cell nuclei in these images via morphology-based supervised classification. Dundar et al. developed a binary classifier using shape, size, and intensity based features taken from individually segmented cells.[19] Doyle et al. used an SVM classifier with textural and architectural features to classify cases as cancerous or non-cancerous, and to differentiate the type of cancer if present.[9] In each of these cases, information about individual cells is critical to the overall system.[21] In some systems, segmentation is even combined with cell detection (and is thus even more critical), as in the work of Arteta et al. on cell detection and segmentation via non-overlapping extremal regions.[22]

Reliable segmentation results for all cells in the images is crucial to the final determinations of these algorithms. However, while many cell segmentation methods exist, they often struggle with the wide variety of cell appearances specific to breast cancer images, even in images collected from a single patient.[21,23] One such example is the pylon model for segmentation.[22,24] Applying this algorithm to our data set (Figure 1) results in inconsistent segmentation.* While, the algorithm correctly segments the majority of cells, it fails at a high rate in

This work was conducted when Andrew Brinker and Annika Fredrikson were NSF REU students at the University of North Carolina at Charlotte (Research Experience for Undergraduates Summer Programs).

*Data provided by Clarion Pathology Lab, Indianapolis and the Computer Science Department at IUPUI, Indiana. Images taken with ScanScope digitizer at 40× magnification.

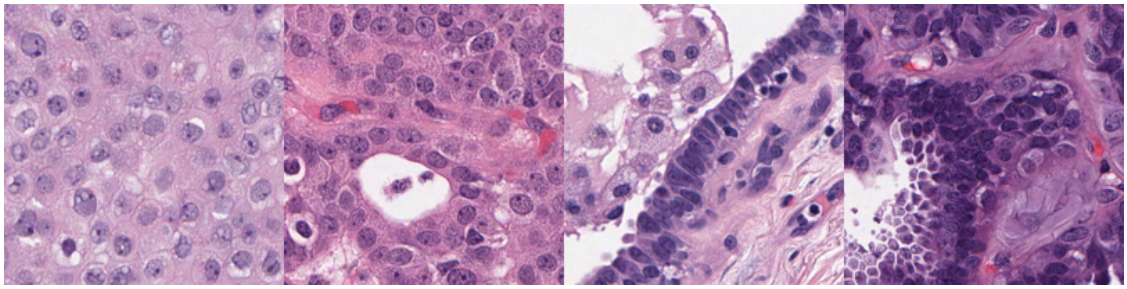

Figure 1. Examples histopathological images, illustrating the wide variety of cell appearance.

the extreme cases, such as conjoined cell clumps or low-contrast cell boundaries. In the first case, the algorithm incorrectly combines cell segments, thus under-represents the cells in the image, and incorrectly represents many others. In the second case, the method misses large portions of the cells that are visually similar to surrounding tissue, or experiences a curling effect as it traces a rough boundary rather than the cell itself. Examples of these failures are shown in Figure 2.

In each case, the varying types of tissue and the variance of the inter-cell structures erode the quality of cell segments and increase the difficulty of cell segmentation in histology slide images. To preserve the correct segments in the average case while improving segments at the extrema, we propose a two-phase system of cell segment correction. This system can be applied to any existing or future segmentation method and can efficiently identify and correct segmentation failures. Through the use of a post-processing step in the cell segmentation phase of a larger CAD system, this method can improve segmentation results and thus improve overall diagnostic quality. We present the first of the two phases: a classifier that can differentiate correct and incorrect cell segments with a high degree of accuracy in advance of the second phase of automatic correction.

2. METHODS

Our data-driven framework, shown in Figure 3, takes cell segmentation results as input and determines whether or not the segmentation correctly delineates the cell in the images in order to correct low-quality segmentation results. Specifically, our method relies on a small collection of computationally-efficient features, shown in Figure 4, split into two categories: appearance-based metrics and shape-based metrics. The appearance-based

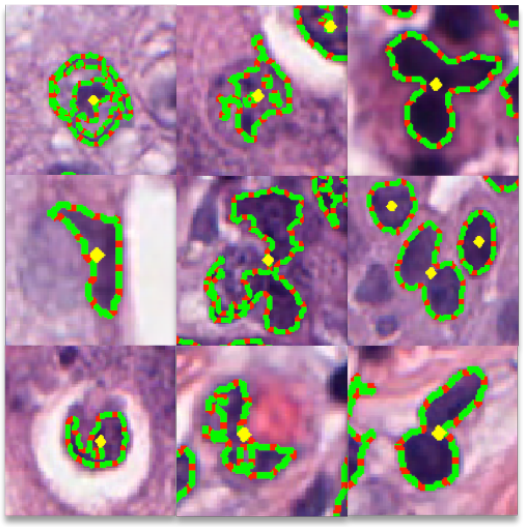

Figure 2. Example failures from the of pylon model[22, 24] on histopathological images.

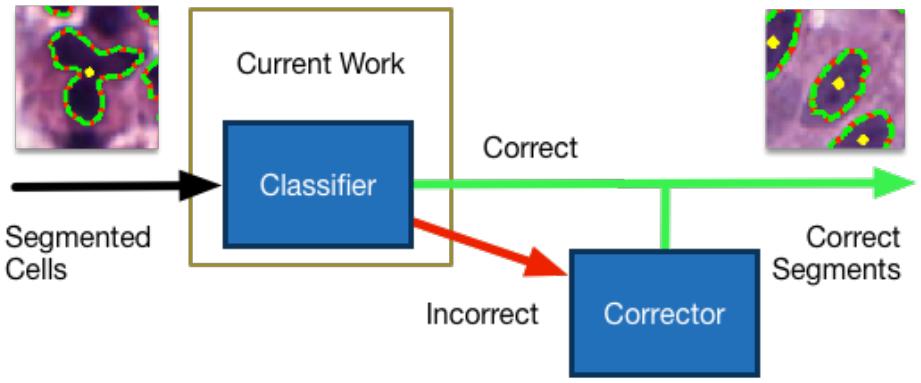

Figure 3. The general process for our proposed method.

metrics (area, circularity, and solidity) represent the relatively consistent shape of cells in histopathological images. The shape-based metrics (intensity variance and entropy) represent the degree of internal consistency seen in these cells.

While there is some degree of variation, cells in a given histopathological image are roughly consistent in area, which makes this measure useful for differentiating correct and incorrect segments. The *area* metric is most useful in cases without clear differentiation between the cells and the surrounding tissue, where some segmentation methods return multiple abnormally small segments. Similarly, the ratio of area to perimeter, which we refer to as *circularity*, is relatively consistent in cells in histopathological images and can be used to identify segments that are overly thin relative to their area. Finally, the *solidity*, defined as the ratio of the area of a segment to the area of the enclosing convex hull, can help to identify segments with a large degree of concavity. This metric helps to identify cases where two cells have been improperly joined together. Collectively, these three shape-based metrics serve to assist in the proper identification of a wide degree of segmentation failures common to cell segmentation methods applied to histopathological images.

Cells in histopathological images are generally consistent in coloration. *Intensity variance* assists in identifying segments where the cell has been improperly separated from the surrounding tissue. If the area enclosed by a segment has a large degree of intensity variance, this is typically because the segment encloses some portion of surrounding tissue. The second appearance-based metric, *entropy*, serves a similar purpose. The Shannon Entropy of the segmented region characterizes the distribution of the intensity values in a given region. Segments which include both a cell and its surrounding tissue generally have a wider distribution of intensity values compared to correct segments. Together, the two appearance based metrics complement the shape-based metrics to, collectively, identify incorrect segments.

Because of given histopathological image can contain hundreds or thousands of cells, we selected both the shape- and appearance-based metrics to be efficient to compute. Area can be computed as the sum of pixels enclosed in the region, circularity only requires the calculation of the perimeter, and solidity requires the solving for the convex hull of the input shape, which is a basic 2D image processing step. The appearance-based metrics, intensity variance and entropy are straightforward, closed-form computations. So, for a large number of cells, our metrics can identify failure cases computationally efficiently. Theses features are normalized and used to train a Support Vector Machine (SVM) classifier.

3. RESULTS

To test the ability of our system to correctly classify segmentation results as low-quality, we applied pylon segmentation[24] with Maximally Stable Extremal Regions to a collection of images with a total of 687 individual segmentations, and input the results into our system. Ground truth was supplied via manual classification of the test data and compared against the results of the system.

Figure 5 shows the results of a 10-fold cross-validation experiment. For each instances of the experiment, we measured the accuracy, $\frac{P_T+N_T}{P+N}$, where P_T is the number of true positives, N_T is the number of true negatives,

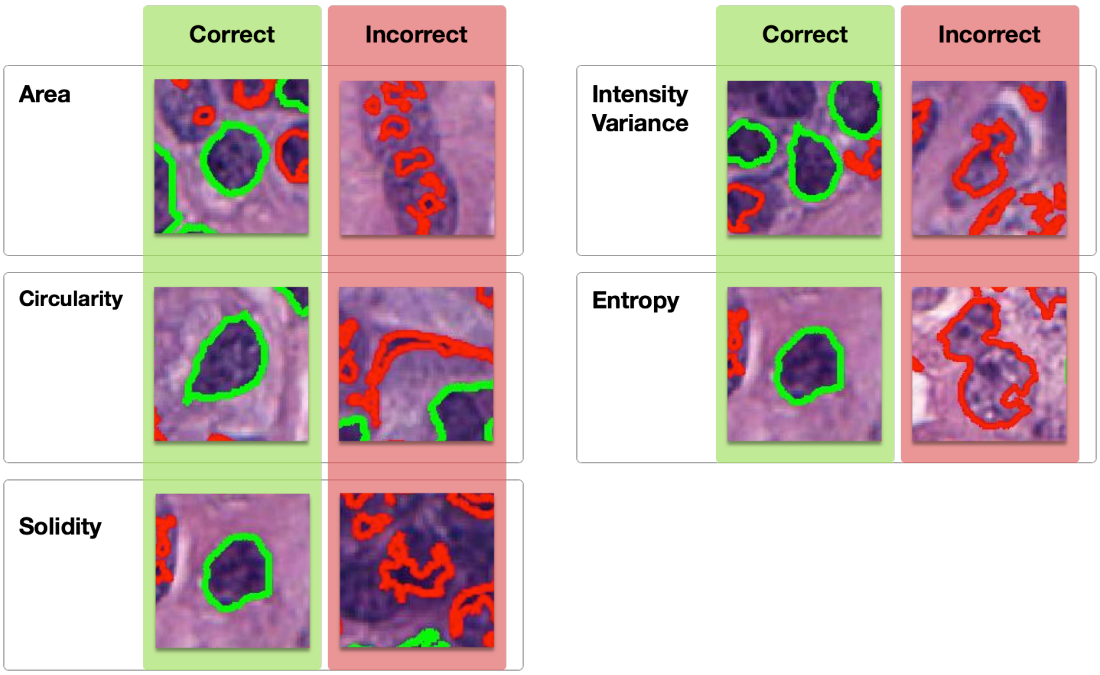

Figure 4. Correct and incorrect segments organized by feature. On the left are the shape-based metrics, on the right are the appearance-based metrics.

P is the number of correctly classified segments, and N is the number of incorrectly classified segments. Across the 10 folds, the average accuracy is 80.38%, which measures the correspondence of our predictions with the human-provided ground truth classifications (both positive and negative).

Further improvements could be achieved using additional, computationally expensive features. Initially, our motivation was to to provide a fast method that could be applied to large-scale data sets. While the current results are promising and can be used to improve the segmentation results from the best-performing algorithms, our framework can be naturally extended with new features. Additionally, we plan to further investigate the variance in the results between the different folds. The standard deviation of the accuracy in our experiment is 7.25%, which indicates some level of sensitivity to the choice of training data. This is unsurprising given the wide variety of cell appearances in histopathological images, but could potentially be addressed using active selection schemes to incorporate the widest variety of examples to be used in training. Nonetheless, we believe this work represents a solid step toward a more complete segmentation correction system.

4. CONCLUSION

Computer-Aided Diagnosis is an important component of the modern fight against breast cancer. CAD systems rely on low-level cell information, including cell segmentation results, to make important decisions about the nature and severity of cases. While there exist many current methods for cell segmentation, they often struggle with the large variety of images present in breast cancer cases. Our method of automatic assessment of segmentation provides a useful starting point for a more complete system of automatic segmentation correction. Via efficiently calculated discriminating features, we are able to provide a usable degree of certainty about segmentation quality, which can then be passed to further stages in this system. Without ties to a particular method of segmentation, our automated correction system can provide worthwhile improvements to any current or future method of cell segmentation, regardless of modality, as a post-processing step after those methods. It can thus be fitted as a small but useful part of the larger CAD ecosystem, and can provide substantive improvements to the quality of cell segmentation, and thus the quality of CAD systems as a whole. In the future, we intend to complete this system via development of an automated correction scheme which uses the output of our current classification

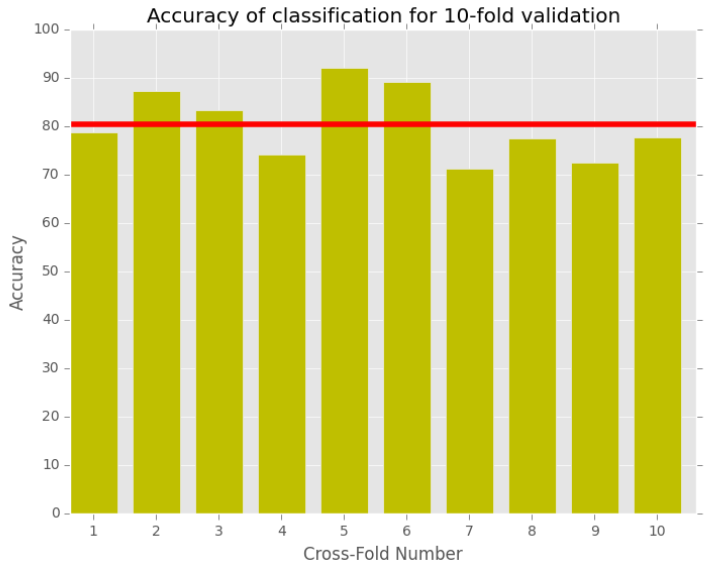

Figure 5. The results of a 10-fold cross-validation experiment using IUPUI data.[‡]1 the The red line indicates the average accuracy across the 10 folds.

system. By identifying common forms of segmentation failure in the identified incorrect segments, we can make efficient and simple corrections that salvage the output of existing methods to be useful to the larger system.

ACKNOWLEDGMENTS

The majority of this work was supported by the National Science Foundation Award #1156822, as part of a Research Experience for Undergraduates (REU) Site, hosted at the University of North Carolina at Charlotte and the Video and Image Analysis Lab.

REFERENCES

[1] "Breast cancer statistics." http://www.cdc.gov/cancer/breast/statistics/. Accessed: 2014-11-30.

[2] Doi, K., "Computer-aided diagnosis in medical imaging: Historical review, current status and future potential," *Journal of the Computerized Medical Imaging Society* **31**, 198 – 211 (2007).

[3] Rangayyan, R. M., Ayres, F. J., and Desautels, J. L., "A review of computer-aided diagnosis of breast cancer: Toward the detection of subtle signs," *Journal of the Franklin Institute* **344**(34), 312 – 348 (2007). Special Issue: Medical Applications of Signal Processing, Part I.

[4] Jiang, Y., Nishikawa, R. M., Schmidt, R. A., Metz, C. E., Giger, M. L., and Doi, K., "Improving breast cancer diagnosis with computer-aided diagnosis," *Academic Radiology* **6**(1), 22 – 33 (1999).

[5] Zhang, X., Liu, W., and Zhang, S., "Mining histopathological images via hashing-based scalable image retrieval," in [*International Symposium on Biomedical Imaging*], IEEE (2014).

[6] Zhang, X., Liu, W., Dundar, M., Badve, S., and Zhang, S., "Towards large-scale histopathological image analysis: Hashing-based image retrieval," *IEEE Transaction on Medical Imaging* **34**(2), 1–11 (2015).

[7] Foran, D. J., Yang, L., et al., "Imageminer: a software system for comparative analysis of tissue microarrays using content-based image retrieval, high-performance computing, and grid technology," *JAMIA* **18**(4), 403–415 (2011).

[8] Basavanhally, A. N., Ganesan, S., Agner, S., Monaco, J. P., Feldman, M. D., Tomaszewski, J. E., Bhanot, G., and Madabhushi, A., "Computerized image-based detection and grading of lymphocytic infiltration in HER2+ breast cancer histopathology," *TBME* **57**(3), 642–653 (2010).

[9] Doyle, S., Agner, S., Madabhushi, A., Feldman, M., and Tomaszewski, J., "Automated grading of breast cancer histopathology using spectral clustering with textural and architectural image features," in [*ISBI*], 496–499, IEEE (2008).

[10] Naik, S., Doyle, S., Agner, S., Madabhushi, A., Feldman, M., and Tomaszewski, J., "Automated gland and nuclei segmentation for grading of prostate and breast cancer histopathology," in [*ISBI*], 284–287 (2008).

[11] Hafiane, A., Bunyak, F., and Palaniappan, K., "Fuzzy clustering and active contours for histopathology image segmentation and nuclei detection," in [*Advanced concepts for intelligent vision systems*], 903 –914 (2008).

[12] Yu, W., Lee, H. K., Hariharan, S., Bu, W., and Ahmed, S., "Evolving generalized voronoi diagrams for accurate cellular image segmentation," *Cytometry Part A* **77**(4), 379 –386 (2010).

[13] Wu, Z., Gurari, D., Wong, J. Y., and Betke, M., "Hierarchical partial matching and segmentation of interacting cells," in [*MICCAI*], 389 –396 (2012).

[14] Nath, S. K., Palaniappan, K., and Bunyak, F., "Cell segmentation using coupled level sets and graph-vertex coloring," in [*MICCAI*], 101 –108 (2006).

[15] Yang, X., Li, H., and Zhou, X., "Nuclei segmentation using marker-controlled watershed, tracking using mean-shift, and kalman filter in time-lapse microscopy," *IEEE Trans. on Circuits and Systems I: Regular Papers* **53**(11), 2405–2414 (2006).

[16] Yan, P., Zhou, X., Shah, M., and Wong, S., "Automatic segmentation of high-throughput rnai fluorescent cellular images," *T-ITB* **12**, 109 –117 (January 2008).

[17] Qi, X., Xing, F., Foran, D., and Yang, L., "Robust segmentation of overlapping cells in histopathology specimens using parallel seed detection and repulsive level set," *IEEE Trans. on Biomedical Engineering* **59**, 754 –765 (mar. 2012).

[18] Sanati, S. and Allred, D., "The progression of pre-invasive to invasive cancer," in [*Pre-Invasive Disease: Pathogenesis and Clinical Management*], ch. 5, Springer (2011).

[19] Dundar, M., Badve, S., Bilgin, G., Raykar, V., Jain, R., Sertel, O., and Gurcan, M., "Computerized classification of intraductal breast lesions using histopathological images," *Biomedical Engineering, IEEE Transactions on* **58**, 1977–1984 (July 2011).

[20] Sokol Petushi, Fernando Garcia, M. M. H. C. K. and Tozeren, A., "Large-scale computations on histology images reveal grade-differentiating parameters for breast cancer," *BMC Medical Imaging* **6**, 14 (2006).

[21] Demir, C. and Yener, B., "Automated Cancer Diagnosis Based on Histopathological Systematic Images: A Systematic Survey," (2005).

[22] Arteta, C., Lempitsky, V., Noble, J. A., and Zisserman, A., "Learning to detect cells using non-overlapping extremal regions," in [*International Conference on Medical Image Computing and Computer Assisted Intervention*], Ayache, N., ed., *Lecture Notes in Computer Science*, 348–356, MICCAI, Springer (2012).

[23] Gurcan, M. N., Boucheron, L. E., Can, A., Madabhushi, A., Rajpoot, N., and Yener, B., "Histopathological image analysis : a review," *IEEE Reviews in Biomedical Engineering* **Vol.2**, 147–171 (October 2009).

[24] Lempitsky, V., Vedaldi, A., and Zisserman, A., "A pylon model for semantic segmentation," in [*Advances in Neural Information Processing Systems*], (2011).