Spring 2020

Automation of Cervical Cancer Cytology

Julian Donovan

Automation of Cervical Cancer Cytology

Dr. Martin Cenek

Julian Donovan

Glossary

The following alphabetically ordered list of terms serves as a reference to clarify the definition of bolded text throughout the rest of the paper.

Active Contour Models - An **energy functional** based segmentation technique focused on finding a curve within the input image's gradient which minimizes the sum of its **energy functions**. With a typical **Active Contour Model**, this can be thought of as finding the contour outlining the steepest edges within an image.

Deep Convolutional Neural Network - A variety of **supervised learning** models capable of performing object detection on images.

Dice Coefficient - An expression of image segmentation accuracy expressed as twice the overlapping area divided by the total number of pixels in both images

Double Thresholding - A component of **Canny edge detection**, taking place after **non-maximum suppression.** In an effort to remove noisy edge data, low and high pixel intensity thresholds are defined. Any pixel whose intensity proves greater than the high filter is kept as a strong pixel edge. Any pixel whose intensity falls between the high and low filters is kept as a weak pixel edge. Weak pixel edges are further considered in blob analysis.

Edge Tracking - A component of Canny edge detection, occurring after **Double Threshold**. Purposed to handle the weak edge pixels, as calculated in the previous step, edge tracking removes weak edge pixels which are not directly connected (by one of its eight adjacent pixels) to strong edge pixels.

Canny Edge Detection - A multi-stage process used to detect prominent edges in a given image. Phases involved include image smoothing via a **gaussian blur**, calculating an image **gradient** to detect edges, applying **non-maximum suppression** to isolate the most significant edges, and utilizing both **double thresholding** and **edge tracking** to remove **noisy** edge data.

Energy Functionals - A mathematical function across a specific section of an image domain purposed to produce a meaningful value motivating some logical consideration. A typical example might be a function which produces smaller values the smoother a known contour becomes. By seeking to minimize such a function, one is able to iteratively generate a smoother contour.

Focal Plane - The distance between the camera and each cell at which the sharpest focus is obtained

Gaussian Blur - An operation involving the application of a Gaussian function across an image, resulting in a blurred version of the original input

Gaussian Distributions - A continuous probability distribution following the general form $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-u}{\sigma})^{23}}$

Gradient - A mapping of the directional changes in color intensity of an image

Level-Set Function - A mathematical function representing a cross-section of a three dimensional image at some given depth. Commonly used in image segmentation applications to model optimization of two dimensional boundaries by interpreting the third dimension as time

Noise - Random variation of color and brightness in images due to errors during image production. For the purposes of this paper, noisy input also includes non-cellular artifacts like blood, mucus, and inflammatory cells

Non-Maximum Suppression - A component of Canny edge detection, following the image **gradient** calculation. Considering edge-steepness, Non-Maximum Suppression seeks to isolate local maxima. By looking at the image **gradient**, where darker pixel values correspond with sharper edges, edge pixels may be traversed as follows: Compare the intensity of each pixel with neighboring pixels, if the intensity of the current pixel is the largest of all its neighbors keep it. Otherwise, remove the current edge pixel.

Super Pixel - A group of pixels known to share one or more common characteristics, like color intensity

Supervised Learning - A machine learning approach in which the model learns from data which has been labelled with the correct answer. In most contexts, this involves manual effort to determine and label each expected result.

Unsupervised Learning - A machine learning approach in which the model learns from pre-existing trends within the data. In consequence, results can be more unpredictable.

Watersheds - A transformation applied to a grayscale image intended to isolate prominent edges. Many different implementations exist.

Wavelet Transforms - A mathematical function purposed to divide waveform data into meaningful components. May be used in context with three dimensional image applications, like pap-smears with overlapping cells, to produce a single output image where all subjects are in focus.

Introduction

Cancer cytology describes the field of study concerned with the understanding and diagnosis of cancerous cells. As applied to cervical cancer, such diagnostic methods often take the form of a Pap smear: A procedure to test for cervical cancer involving the collection and analysis of cells from the patient's cervix. Once the cervix has been swabbed, the collected cells are deposited onto a cell slide where they may be manually reviewed under microscope by a trained lab technician [3]. With each cell needing to be individually considered, 100,000 - 200,000 cells typically existing on a single slide, and only 10 cancerous or precancerous cells being required to diagnose a patient, the Pap smear process proves both exhaustive and difficult. Further rigor is introduced by the high degree of overlap often found amongst cervical cells as well as the presence of mucus, blood, and inflammation. A combination of all these factors results in high rates of **false negatives**, affecting between 20 and 30 percent of all screenings [1].

Considering the slow speed and lacking accuracy which Pap smear tests currently offer, automating the detection and classification of cervical cancer cells remains a highly discussed problem within the realm of microscopic image analysis [4]. The goal being to create computer software capable of accepting a cell slide image as input, and returning a segmented version where all cell boundaries have been individually isolated. Recognizing the success of a particular approach introduced by Zhi Lu et al. in their publication *An Improved Joint Optimization of Multiple Level Set Functions for the Segmentation of Overlapping Cervical Cancer Cells*, this paper outlines an effort to convert Zhi Lu's Matlab code to a performant C++ implementation known as Cytology Portal. With Zhi Lu's code achieving an average accuracy of 94 percent when considering a **Dice Coefficient** of 0.7 [7], it is our hope that through optimizing and publically distributing such open-source code that Cytology Portal may function as an accessible decision support tool for all. Serving to automate the task of manual cell identification for hospitals and medical research facilities, Cytology Portal will speed up the pap smear process by providing lab technicians with a list of individually segmented cells - sortable by dynamically configurable criteria (i.e. size) - for each cell slide analyzed.

Background

Of the various approaches taken to the problem of automating cervical cancer cell segmentation, two overarching groups may be defined: **Supervised** and **unsupervised** algorithms [4]. The work performed by Zhi Lu et al. in their paper *An Improved Joint Optimization of Multiple Level Set Functions for the Segmentation of Overlapping Cervical Cancer Cells* falls into the latter of these two categories, with his architecture proving capable of segmenting cervical cancer cells without the use of annotated cervical cancer cell images [7]. Specifically, Zhi Lu et al. rely upon a **level-set** implementation, where the boundaries of each cell are modelled as mathematical functions describing logical considerations (i.e. boundary overlap). By evolving the **level-set** such that each of these functions is optimized, cell boundaries may be effectively calculated [7]. Though several other examples of **unsupervised** approaches exist - including **watersheds** [9], and **active contour models** [6] - of all the publications we surveyed Zhi Lu et al. proved most accurate when considering high degrees of cellular overlap [4], [7].

Though Zhi Lu's work proves precise, performance remains an issue. Even with the optimizations which Cytology Portal seeks to implement, we recognize that execution will likely remain less than optimal as Zhi Lu's architecture involves many complicated phases. **Supervised** algorithms, like **Deep Convolutional Neural Networks** (D-CNNs), represent a solution to this problem. As demonstrated in several publications, these models prove capable of providing efficient and accurate segmentation results for cells with 40% boundary overlap or less [5], [8]. However, being a **supervised** model such architectures require labelled cervical cancer cell slides which are quite difficult to source due to their time consuming and technical nature. As such, we have chosen to implement our architecture based on Zhi Lu's **unsupervised** segmentation pipeline in hopes that Cytology Portal may later support the development of a D-CNN through generating synthetically labelled cell slides and providing performance/accuracy benchmarks to gauge future improvements.

Methodology

Overview

This section seeks to outline the process by which cell segmentation occurs, detailing non-technical descriptions of each segmentation phase as well as technical explanations of the Cytology Portal codebase. A bulleted summary of these phases and their major sub-phases has been included below.

1. **Scene Segmentation** - Outlining cell clumps, nuclei and initial cell boundaries
    a. Image Preparation
    b. Cell Clump Segmentation
        i. Super Pixel Map (Quick Shift)
        ii. Super Pixel Edge Map (Canny Edge Detection)
        iii. Clump Segmentation (Gaussian Mixture Model)
    c. Nuclei Segmentation (Maximum Stable Extremal Regions)
    d. Initial Cell Segmentation
        i. Shape Prior Calculation
2. **Cell Segmentation** - Determining final cell boundaries
    a. Cell Segmentation (Binary/Unary Distance-Regulated Level Set Equations)
    b. Re-Compute Shape Priors each Iteration

Scene Segmentation

Scene segmentation seeks to isolate clumps of cells within the initial image. Upon completion of the scene segmentation phase, the architecture will have identified clump boundaries, cell nuclei, and initial boundary predictions for individual cells. This process consists of five primary stages in order: image preparation, cell clump segmentation, nuclei segmentation, initial cell segmentation, and shape prior calculation.

Image Preparation

For best results, Cytology Portal expects all cell slide image input to utilize mono-layer preparation and extended depth of field (EDF) image transformations. Mono-layer preparation techniques help to ensure cells are deposited on the slide without overlap, this functions to reduce ambiguity between cell boundaries. EDF transformations utilize **wavelet transforms** to ensure that all cells are in focus regardless of **focal plane** depth (Figure 1). In order for cell boundaries to be properly recognized all cells must be in focus. All images inputted to Cytology Portal are expected to follow a PNG file format. Both colored and grayscale images are accepted by Cytology Portal, as all images are converted to grayscale at runtime. A trivially difficult sample image meeting each of the previous criteria has been included below (Figure 2).

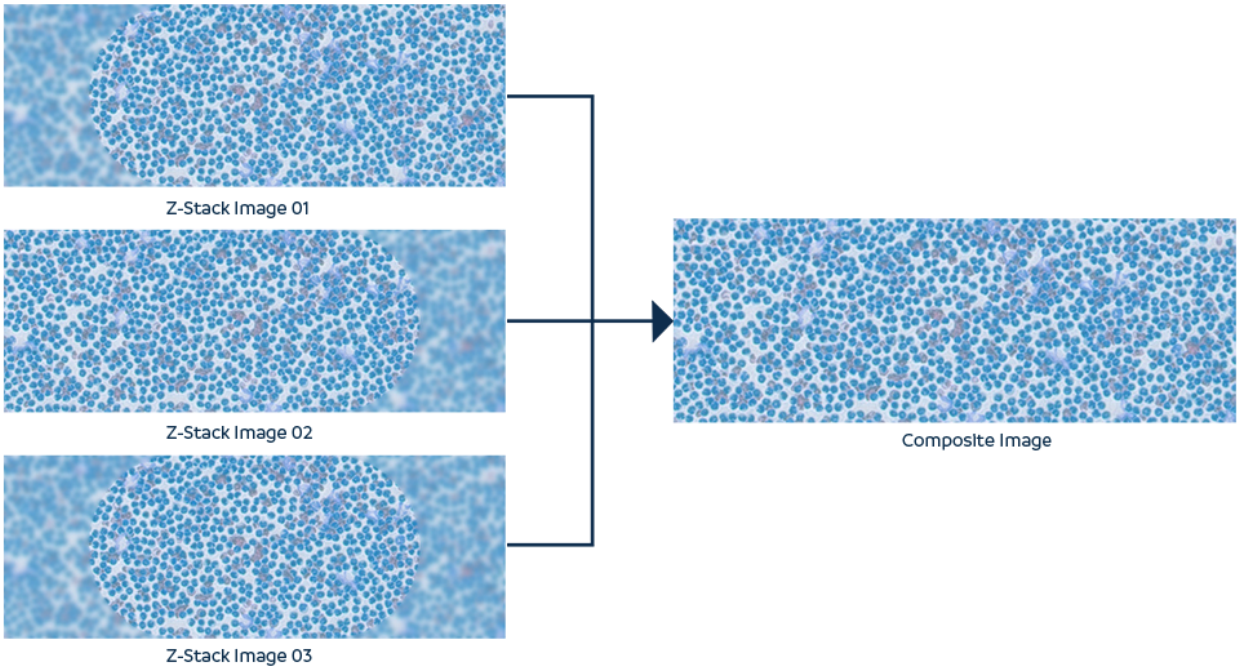

Figure 1. Several images of varying focus being composed through EDF such that all cells are in focus [10]

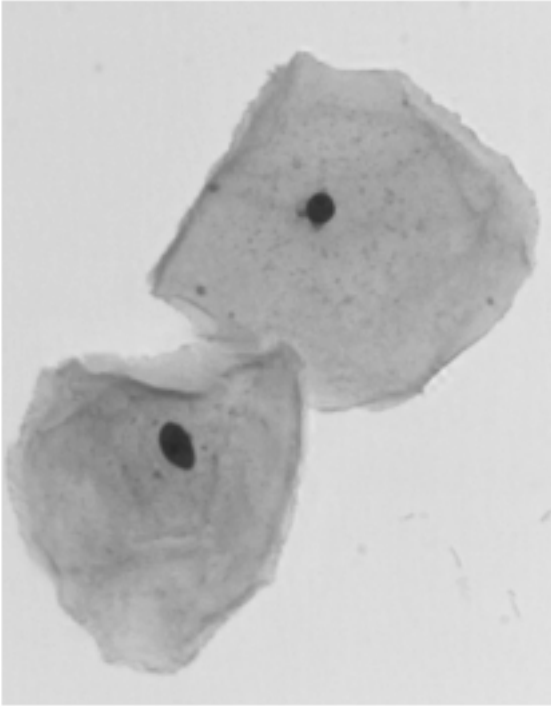

Figure 2. Sample input for Cytology Portal

Cell Clump Segmentation

Cell clump segmentation involves the isolation of overlapping cells from their background. By separating these cell clumps, future phases in the segmentation pipeline may focus where cells are known to exist instead of pointlessly looking for nuclei or cell boundaries within image whitespace. This process of cell clump segmentation consists of three steps listed in order: super pixel mapping, super pixel edge mapping, and clump segmentation.

Super Pixel Map (Quick Shift)

In order to begin determining where clump boundaries are located, Cytology Portal uses a quick shift algorithm to group pixels together based on their color intensities. The idea being to find areas of similar color values such that the mode intensity of each group is maximised. The end result is an image broken into like-pixel-color chunks, known as **super pixels**. In context with cervical cell segmentation, since pap smear cell slide backgrounds are known to be whitish with relatively little variation in intensity, the image background is effectively linked together as a

cohesive chunk. This provides a starting point from which cell clump segmentation may occur (Figure 3).

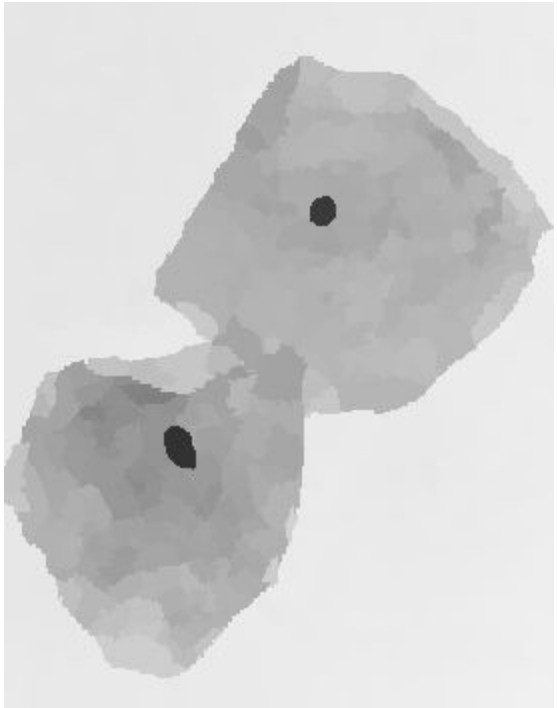

Figure 3. Super pixel map of the sample input image (Figure 2) as calculated via the quick shift algorithm

Super Pixel Edge Map (Canny Edge Detection)

Though the quick shift algorithm helps to group areas of like-coloration in a way which renders clump boundaries apparent to the human eye, a computer has no such intuition. Instead, neighboring super-pixel edges must be considered to distinguish clump boundaries from random **noise**. Known as an edge detector, this process involves locating areas within an image where the color **gradient** abruptly shifts. In specific, Cytology Portal uses a **Canny edge detector**. This **Canny edge detector** functions by smoothing the image with a **Gaussian blur**, calculating the image **gradient** of the smoothed image, removing all non-local maxima with **non-maximum suppression**, removing obvious noisy edges with **double thresholding**, and carefully excluding possible noisy edges with **edge tracking**. The end result is an image where most background information has been removed and the majority of edges which remain pertain to either clump or cell boundaries (Figure 4).

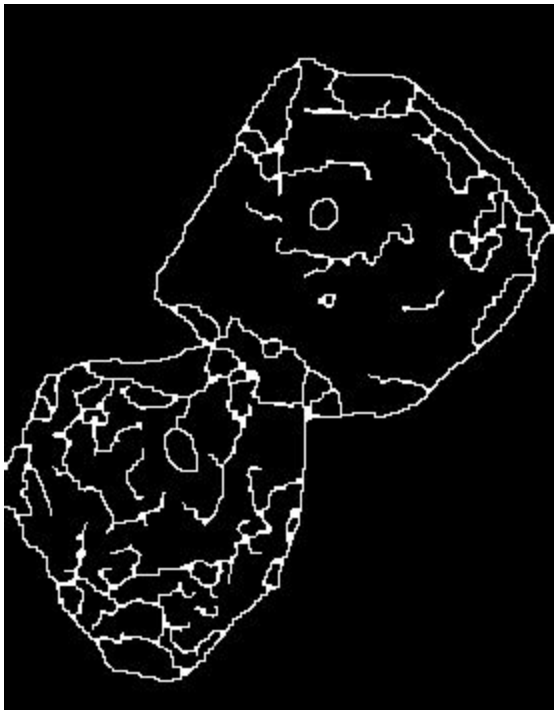

Figure 4. Super pixel edge map of the sample input image (Figure 2) as calculated via Canny edge detection

Clump Segmentation (Gaussian Mixture Model)

Having simplified the input image to a set of prominent cell and clump edges, Cytology Portal may calculate clump boundaries through the use of a binary classification algorithm. Purposed to classify each pixel into one of two classes (clump pixels or background pixels) this process is performed by an unsupervised Gaussian Mixture Model. Gaussian Mixture Models define a probabilistic method to cluster an arbitrary number of node classes based upon **Gaussian distributions**. However, before such a model may be used, initial pixel classes must first be assigned.

Cytology Portal generates these initial class assignments through the construction of a convex hull. Given a set of points, a convex hull is defined as the smallest polygon capable of containing them all. In this context, the convex hull seeks to include all edge pixels from the previous phase's super pixel edge map (Figure 6). All pixels within the convex hull are initially

assigned to the clump class, while all pixels outside the convex hull are assigned to the background class.

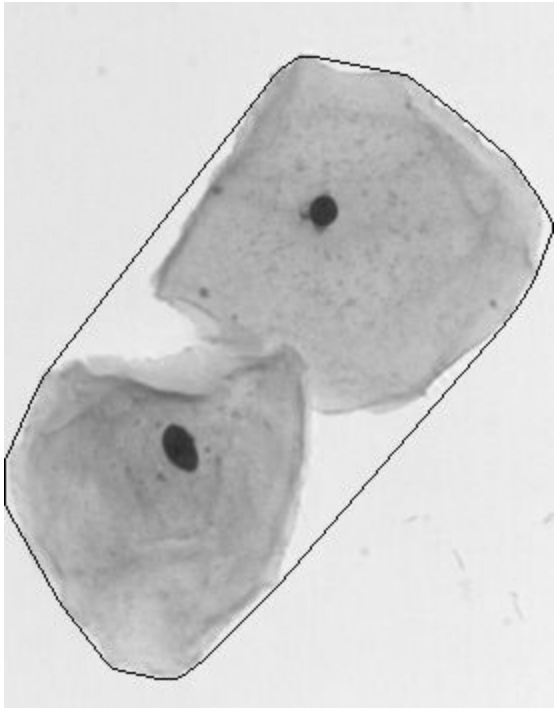

Figure 6. Convex hull of the sample input image (Figure 2)

Having met the prerequisite of initial pixel class assignments, two binary **Gaussian Mixture Models** (GMM) may now be used to learn clump and background classifications across the input image (Figure 2). After each training iteration, the joint **likelihood map** created by aggregating the output of the two models is used to fine tune each GMM. Once the joint **likelihood map** has stabilized, with no significant changes occurring between training iterations, the output values are used for clump and background class assignment. In an effort to remove **noisy** data, all clumps of size less than a user defined threshold are removed (Figure 7).

Figure 7. Binary classification of the sample input image (Figure 2) as calculated via Gaussian Mixture Model

Nuclei Segmentation (Maximum Stable Extremal Regions)

With clump segmentation completed, Cytology Portal is now ready to search within each clump for nuclei based on their typically uniform texture and high color intensities. Operating under the assumption that nuclei do not overlap, the Maximum Stable Extremal regions algorithm may be used to detect stable connected components via usage of **level set methods**. Essentially, this entails searching each cell clump (Figure 7) within the input image (Figure 2) such that all areas of high pixel intensity with regular, elliptical shape are marked as nuclei (Figure 8).

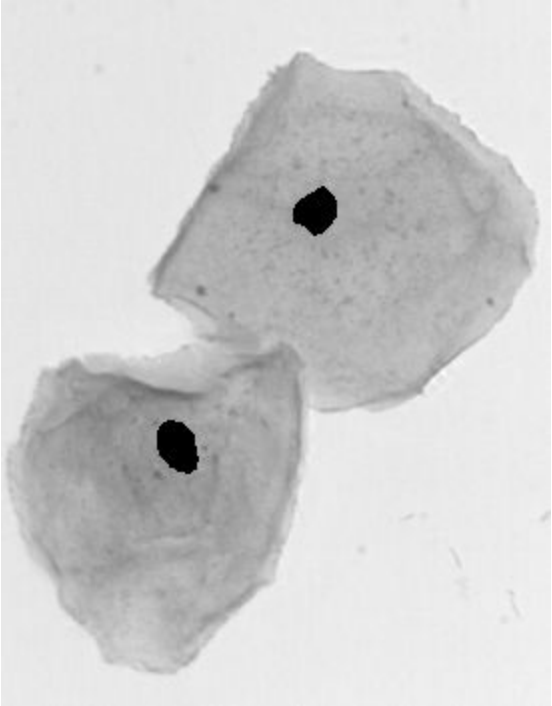

Figure 8. Segmented nuclei from the sample input image (Figure 2) as calculated via Maximum Stable Extremal Regions

Initial Cell Segmentation

Initial cell segmentations revolve around the empirical observation that the majority of cytoplasmic contours are located on pixels of similar distance from their contextual nuclei. Using this knowledge and the results from the previous step, the clump boundary is iteratively traversed such that each edge pixel is assigned to the nearest nuclei. Prior to any assignment, it is verified that a straight line connecting the prospective nucleus to the clump edge pixel can be drawn. This ensures no holes exist within the initial segmentation. Once all clump boundaries have been assigned, missing cell boundary information may be filled in. For any cell whose boundary proves incomplete, a circular radius equal to the distance to the nearest nucleus (within the clump) is used to fill in the gaps (Figure 9).

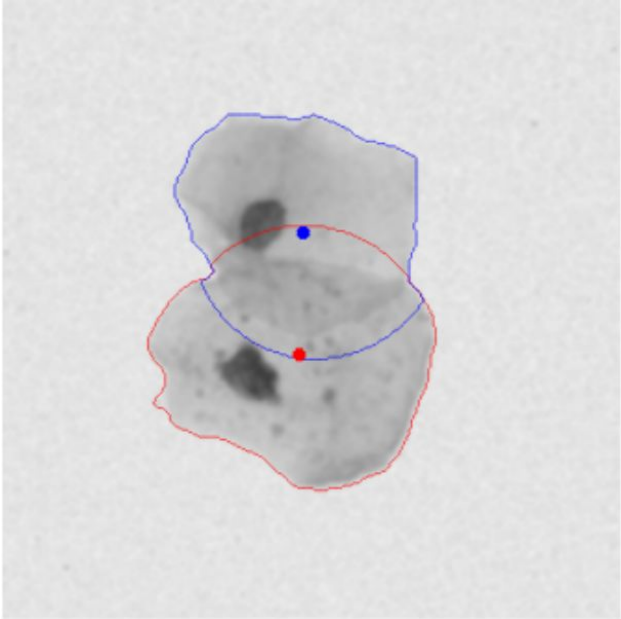

Figure 9. Initial cell segmentation example, not produced by Cytology Portal [7]

Shape Prior Calculation

Implemented as an array of pixel values, of equal size to the input image (Figure 2), a **shape prior** is defined for each cell. This construct proves useful in the cell segmentation phase, where high pixel intensities (within the **shape prior)** repulse the growing cell boundary and low pixel intensities attract it. Purposed to motivate cell boundary growth towards the edge of each initial segmentation, Cytology Portal instantiates **shape priors** such that each cell's nuclei and the image background are given high pixel intensities whereas the interior of each cell is defined by a gradient of decreasing pixel values (Figure 10). A clump prior, representing the aggregation of all cell shape priors is formed by taking an aggregate max of all shape priors within a given clump.

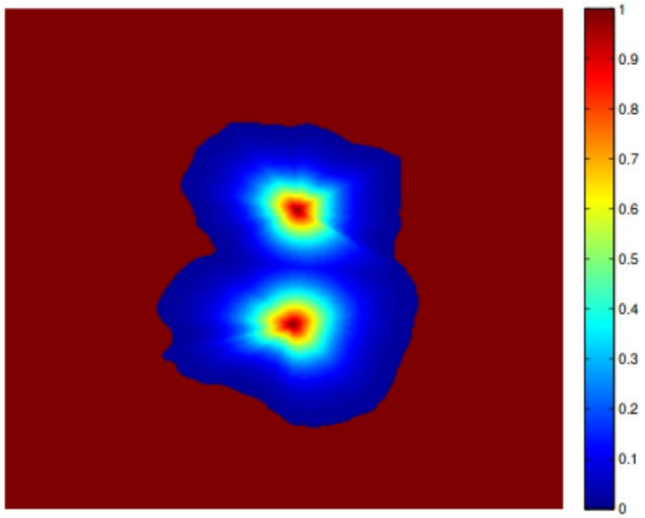

Figure 10. Clump prior example, not produced by Cytology Portal [7]

Cell Segmentation

Cell segmentation seeks to accurately divide an input image of many cells into many separate images of individual cells. As implemented in Cytology Portal, such segmentation is performed through an iterative process where a boundary is evolved outwards based on **energy functional** considerations and **shape prior** constraints. Broken into two distinct phases of cell segmentation and re-computation of shape priors, this process iteratively loops until cell segmentations converge.

Cell Segmentation (Binary/Unary Distance-Regulated Level Set Equations)

Using the initial cell segmentations and **shape priors** calculated within the scene segmentation phase, better cell boundary estimates may be iteratively computed. Relying on two dimensional, distance-regulated **level-set functions**, the boundaries of each cell may be expressed in terms of unary and binary **energy functionals**. These **energy functionals** are mathematical formulas serving to represent considerations regarding boundary smoothness, boundary area, and inter-cellular overlap between neighboring cells. As **level-set functions** evolve over time, growth is driven such that the total sum of all **energy functionals** is minimized. Effectively, this

results in a set of cell boundaries which are suitably smooth, of regular shape, and feature minimal inter-cellular overlap (Figure 11).

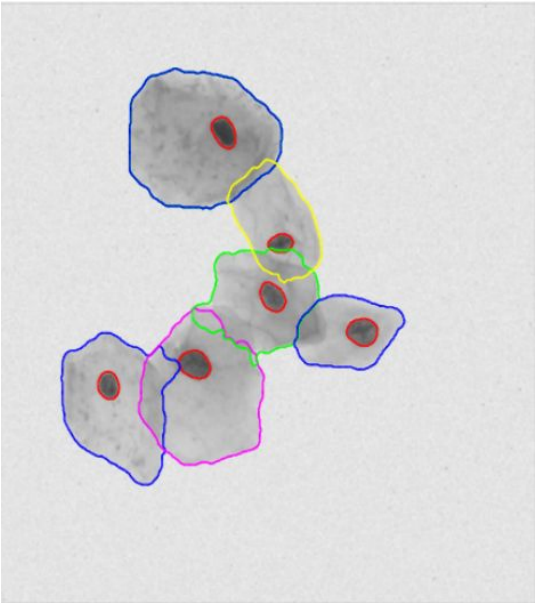

Figure 11. Cell segmentation example, not produced by Cytology Portal [7]

Re-Compute Shape Priors

To ensure smooth evolution of all **level-set functions** all shape priors are re-computed after each iteration. This process functions identically to that described in the scene segmentation shape prior section listed above.

Analysis

Regretfully, as of this time Cytology Portal's cell segmentation phase is still in development and as such the architecture has yet to produce functional results. However, recognizing that the scene segmentation phase is fully functional, prospective performance improvements for this subsection may still be estimated. To this end, the scene segmentation implementations for both Zhi Lu's architecture and Cytology Portal have been run against a 100 image random subsampling of the Herlev dataset [2]. These results have been recorded below (Table 1).

Table 1. Average Runtime per Scene Segmentation Subphase in Seconds (s)

	Average Cytology Portal Runtime (s)	Average Zhi Lu Runtime (s)
Cell Clump Segmentation	34.95557	13.532147
Nuclei Segmentation	0.650355	0.779124
Initial Cell Segmentation	331.887159	3827.354640
Scene Segmentation **(Total)**	367.493084	3841.665911

As demonstrated, Cytology Portal provides a performance improvement of 1045% when considering the scene segmentation phase of cervical cancer image segmentation. Though this does not guarantee any performance shifts for the cell segmentation phase, it validates the central thought of this research effort: that reimplementation of Zhi Lu's matlab code in performance-optimized C++ will yield significant runtime speedups. Interestingly, noting that on average our cell clump segmentation ran 21 seconds slower than Zhi Lu's architecture, it is suggested that further performance optimizations may be possible. Specific attention will be paid to such improvements as the cell segmentation phase is implemented in an effort to ensure that the Cytology Portal will be able to function as an efficient, accurate decision support tool for cervical cancer analysis.

References

1. C. Arratoon, “ASSESSMENT OF TECHNIQUES FOR CERVICAL CANCER SCREENING,” *Canadian Coordinating Office for Health Technology Assessment*, May 1997.
2. M. Arya, N. Mittal, and G. Singh, “Clustering Techniques on Pap-smear Images for the Detection of Cervical Cancer,” *Journal of Biology and Todays World*, vol. 07, no. 01, 2018.
3. E. Bengtsson and P. Malm, “Screening for cervical cancer using automated analysis of PAP-smears,” *Computational and mathematical methods in medicine*, 2014. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3977449/. [Accessed: 6-Apr-2020].
4. T. Conceição, C. Braga, L. Rosado, and M. J. M. Vasconcelos, “A Review of Computational Methods for Cervical Cells Segmentation and Abnormality Classification,” *International journal of molecular sciences*, 15-Oct-2019. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6834130/. [Accessed: 23-Apr-2020].
5. Y. Huang, H. Zhu, P. Wang, and D. Dong, “Segmentation of Overlapping Cervical Smear Cells Based on U-Net and Improved Level Set,” *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, 2019.
6. P. Kumar, S. L. Happy, S. Chatterjee, D. Sheet, and A. Routray, “An unsupervised approach for overlapping cervical cell cytoplasm segmentation,” *2016 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, 2016.
7. Z. Lu, G. Carneiro, and A. P. Bradley, “An Improved Joint Optimization of Multiple Level Set Functions for the Segmentation of Overlapping Cervical Cells,” *IEEE Transactions on Image Processing*, vol. 24, no. 4, pp. 1261–1272, 2015.
8. Y. Song, L. Zhang, S. Chen, D. Ni, B. Li, Y. Zhou, B. Lei, and T. Wang, “A deep learning based framework for accurate segmentation of cervical cytoplasm and nuclei,” *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2014.
9. A. Tareef *et al*., "Multi-Pass Fast Watershed for Accurate Segmentation of Overlapping Cervical Cells," in *IEEE Transactions on Medical Imaging*, vol. 37, no. 9, pp. 2044-2059, Sept. 2018.

10. “Index,” *optrascan*. [Online]. Available: https://www.optrascan.com/solutions/cytology-imaging. [Accessed: 23-Apr-2020].