AN IMPROVED HYBRID ACTIVE CONTOUR MODEL FOR NUCLEAR SEGMENTATION ON BREAST CANCER HISTOPATHOLOGY

Juan Jing† Tao Wan†∗ Jiajia Cao‡ Zengchang Qin‡∗

† School of Biological Science and Medical Engineering, Beihang University, China
‡ Intelligent Computing & Machine Learning Lab, School of ASEE, Beihang University, China

ABSTRACT

Segmentation of nuclei on breast cancer histopathological images is considered a basic and essential step for diagnosis in a computer-aided diagnosis framework. Nuclear segmentation remains a challenging problem due to the inherent diversity of cancer biology and variability of nucleus appearance. We present an automatic nuclear segmentation method using an improved hybrid active contour (AC) model driven by both boundary and region information. The initialization of segmentation based on morphological operations and watershed allows for generation of initial closed curves and reduction in computational load of curve evolution for AC model. Color gradients are computed to capture image gradients along the nucleus border. The AC segmentation scheme is performed in a coarse-to-fine fashion which can help to solve the problem of multiple object overlap in an image scene. Segmentation performance was evaluated on various breast cancer histopathological images with different grades and was compared with the other popular AC models, suggesting that the improved hybrid AC model can be used to build an accurate and robust nuclear segmentation tool.

Index Terms— hybrid active contour model, nuclear segmentation, breast cancer, histopathology

1. INTRODUCTION

Histopathological analysis is a common clinical procedure for diagnosing the presence, type, and progression of breast cancer. Analysis methods that are routinely performed by pathologists, such as identification of the most progressed regions and examination of nuclear morphology, can be tedious and are hampered by observer variability [1]. There is a need for computer-aided diagnosis (CAD) to potentially tackle the existing problems. The segmentation of nuclei in breast cancer histopathological images is the first and critical step in cancer diagnosis in a CAD framework. Nuclear segmentation remains a very challenging problem particularly for routinely stained hematoxylin and eosin (H&E) sections. Many of the challenges could arise from:

- Tissue variations due to imperfections in the staining process;
- Heterogeneous appearance inside nucleus resulting in over-segmentation;
- Low contrast and weak boundaries between background and objects;
- Nucleus overlap leading to under-segmentation.

A large variety of techniques for automatic segmentation of nuclei in breast cancer histopathological images have been proposed [2]. Active contour (AC) models are among the most popular methods in various applications. The AC methods can be categorized as boundary-based and region-based schemes. The boundary-based approaches, such as geodesic active contours [3], have become popular on account of their reliable performance when strong object gradients are present and fast computation. Region-based approaches, such as the Chan-Vese model [4], rely on statistical information derived from foreground and background regions. Both methods have disadvantages which limit their ability in dealing with object occlusions and scene overlaps. Ali and Madabhushi introduced an integrated boundary and region based active contour model along with shape information to segment nuclear and glandular structures on digitalized breast histopathological images [5]. This motivates our work on building a hybrid active contour scheme for segmentation and multiple object overlap resolution.

In this work, we present an improved hybrid active contour model (see Fig.1) to automatically segment nuclei in breast cancer histopathological images. Watershed based automatic initialization is able to generate initial curves close to the actual nuclear boundary and speed up the computation of the subsequent AC segmentation. Color gradients are computed to better capture image gradients at the margin of nucleus for use with the hybrid AC model. This is particularly useful when nuclei have diffuse boundaries. To prevent the problem of single contour enclosing multiple overlapping nuclei, The AC based segmentation is performed in a coarse-to-fine fashion in which the independent nuclei are well segmented in the coarse segmentation phase, while the potential overlapping regions containing multiple nuclei,
identified by measuring the region properties, are further segmented in the fine segmentation phase. To overcome intensity inhomogeneity within the regions, we incorporate the local image information into the hybrid AC model to perform an accurate and robust nuclear segmentation.

The rest of this paper is organized as follows. The AC based segmentation method is presented in Section 2. The experimental results and discussion are given in Section 3. Section 4 concludes the paper.

2. METHODOLOGY

2.1. Data Preparation

We collected 89 H&E stained breast cancer histopathological images from 89 patient studies diagnosed using the Nottingham grading system. The dataset contains 22 low, 49 intermediate, and 18 high grade breast cancers. All slides were digitized via a whole slide scanner (Motic©, Xiamen, China) at 40× magnification (1µm/pixel resolution). A pixel-based color-map quantile normalization method [2] was utilized to reduces the differences in tissue samples due to variation in staining and scanning conditions.

2.2. Initialization

We presented an integrated boundary and region-based active contour model in a level set formulation with an automatic initialization based on morphological operations and watershed. First, we created a binary image via a simple 2-class clustering method. The morphological operators were then used to remove small objects and isolate individual elements.

The watershed transformation was applied to obtain the estimated nuclear boundaries. Finally, we computed Euclidean distance on the watershed-generated binary mask to obtain an initial delineation and initialize the level sets. Fig.2(c) shows a better segmentation result using our initialization method in comparison with the uniform spaced placement of the initial contour (Fig.2(b)). Further, our initialization method yielded 20% savings in computational times compared to the uniform spaced initialization scheme.

![Fig. 2. Watershed initialization (a) and uniform spaced initialization (b) with corresponding segmentation results (c) and (d), respectively.](image-url)

2.3. Hybrid AC Segmentation

The presented hybrid AC model incorporates both boundary- and region-based information and comprises two phases (coarse and fine) to achieve accurate nuclear segmentation. The contours that segment the nuclear boundaries are represented using the level set method, and are evolved by minimizing the variational energy functional expressed as [6]:

$$
\varepsilon(\phi) = -\alpha \int_{\omega} (I - \mu) H(\phi) d\omega + \beta \int_{\omega} G |\nabla H(\phi)| d\omega
$$

where $\phi$ is the zero set of embedding function representing the active contour $C = \{x|\phi(x) = 0\}$, $I$ is the image to be segmented, $H(\phi)$ is the Heaviside function, $\omega$ represents image domain, $G = G(|\nabla I|$ is the image gradient, and $\alpha$ and $\beta$ are predefined weights to balance the two terms. The first term is a region term where $\mu$ is a predefined parameter indicating the lower bound of the gray level of the nucleus. The region term guides the contours to enclose the regions with grayscale values greater than $\mu$. The second term is the geodesic active contour function in the level-set formulation. In our AC model, we computed RGB color gradients, instead of gradients in the grayscale space, to capture the image gradients at the nuclear border. The color gradient is defined as [7]: $G = \sqrt{\lambda_x - \lambda_\gamma}$, where $\lambda_x = (G_1 + G_2 \pm \sqrt{(G_1 - G_2)^2 + 4G_3^2})/2$, $G_1 = (\partial I/\partial x)^2 + (\partial I/\partial y)^2$, $G_2 = \partial I/\partial x \cdot \partial I/\partial y$, $G_3 = \partial I/\partial y$, $\partial I/\partial x$, $\partial I/\partial y$, $I_R$, $I_G$, and $I_B$ denote RGB components. The AC energy functional $\varepsilon(\phi)$ was accordingly modified by incorporating RGB color information in both region and boundary terms. Fig.3 illustrates that color gradients
In this paper, we presented an improved hybrid active contour scheme for automatic segmentation of nuclei in the histopathological images.

3. EXPERIMENTAL RESULTS

3.1. Experimental Design

We compared the segmentation method with the other two popular AC models. The Chan-Vese model utilized the global image statistics inside and outside the evolving curve rather than the gradients on the boundaries [4]. Zhang’s model is a hybrid AC method that combined both boundary and region information and was developed in a level-set framework [6]. We used the optimal parameter settings for both models reported in [4, 6]. The parameters of the presented AC model were empirically tuned to yield the best segmentation results. In all the experiments, we used $\alpha = 0.01$, $\beta = 1.0$, $\mu = 0.40 - 0.55$ in the coarse segmentation phase and $\alpha = 0.05$, $\beta = 1.0$, $\mu = 0.47 - 0.50$ in the fine segmentation phase.

The window length of the median filter ranged from 3 to 7 determined by the size of the detected objects.

3.2. Results

The qualitative evaluation of our segmentation method was visually performed across all 89 histopathological images with different breast cancer grades. The segmentation results are illustrated in Fig.4. Our segmentation method outperformed Zhang’s model [6] and the Chan-Vese model [4] in terms of good nuclear separation and accurate curve evolution. Figures 4(b), 4(g), 4(l) show the coarse segmentation in which the majority of nuclei have been separated while some overlapping nuclei are still connected. The evolving curves continue to split on these regions and successfully stop on the boundaries of clumped nuclei when meet the termination criteria (see Figures 4(c), 4(h), 4(m)). It is can be clearly seen that the Chan-Vese model (see Figures 4(e), 4(j), 4(o)) fails to handle the object overlap problem and converges to the outer boundaries enclosing multiple clumped nuclei. Zhang’s model (see Figures 4(d), 4(i), 4(n)) provides good performance in nuclear segmentation. However, the model also shows its limitation in dealing with overlapping nuclei. Further, we evaluated the segmentation performance of our model on various images with different cancer grades. Fig.4(m) demonstrates the ability of our method to segment high grade cancer images as it provides more accurate and smooth boundaries compared with the other two models.

All the three models were implemented by Matlab R2013a platform on a computer with Intel Duo E7500 2.94GHz CPU, 4GB RAM. The average processing time per 300 nuclei is 23 seconds for the coarse segmentation, 33 seconds for the fine segmentation, 96 seconds for Zhang’s model, and 59 seconds for the Chan-Vese model.

4. CONCLUDING REMARKS

In this paper, we presented an improved hybrid active contour scheme for automatic segmentation of nuclei in the histopathological images.
histopathological images. This coarse-to-fine segmentation approach allows for accurate separation of multiple overlapping nuclei enclosed in a single contour. The segmentation method was evaluated in context of segmenting nuclei on low, intermediate, and high grade breast cancer histopathology. Our results showed that our model achieved better segmentation performance and faster computation compared with the popular AC models. In future work we intend to leverage the improved AC model for nuclear and cell segmentation in other domains within digital pathology.

5. REFERENCES


