A Curvature-Norm Based Centroid Initialized Distance Regularized Level Sets for Nuclear Segmentation in Histopathological Images

P.M. Shivamurthy¹*, T.N. Nagabhushan¹ and Vijaya Basavaraj²

¹Sri Jayachamarajendra College of Engineering, Mysuru, India.
²JSS Medical College and Hospital, Mysuru, India.

http://dx.doi.org/10.13005/bpj/1496

(Received: 02 June 2018; accepted: 04 August 2018)

Nuclear pleomorphism is considered to be one of the most significant shape based feature adapted in grading the cancer through the pathological studies of the H&E stained tissue slides. Microscopic study of manually extracting the feature is highly laborious and misleads the pathologists during grading. Digitization of the slides has given rise to various segmentation approaches to extract the nuclei shape to assess the degree of pleomorphism. Here, a novel approach of initializing and evolving the distance regularized level sets (DRLS) for the detection and segmentation of the nuclei has been presented. In this work, two major objectives have been achieved. First, a novel geometric approach has been devised for the detection of centroids of each nuclei in the occluded region and second, a shape prior model has been presented for the extraction of gradient information through morphological operations. The multiple level set implementation of the DRLS contours are initialized using the centroids detected and driven through the gradient computed. The proposed method has been experimented over the images of benign and malignant breast cancer tissue obtained from BeakHis dataset. A quantitative analysis of the results have shown that a 97% of object detection accuracy and 78% of overlap resolution has been achieved through the proposed model. A comparative study with that of geodesic active contours have indicated an improvement in the segmentation accuracy measure of 9-10 pixel difference.

Keyword: Nuclear pleomorphism, Curvature-norm, Geometric centroid, shape-prior model.

Extraction of the significant biomarker play a significant role in the diagnosis and prognosis of the cancer. Pathologists have considered the nuclear pleomorphism as an important shape based biomarker in the process of staging the disease. A manual delineation of the nuclear boundary fetches the pleomorphic features based on the shape detected through microscopic observation of the histopathological slides. This is a very laborious and time consuming activity due to the heterogeneity of the tissue objects observed on the slides. This will mislead the diagnosis process resulting in wrong staging. This may also, be due to inter and intra-observer variabilities existing among the pathologists.

Many automated staging systems have been proposed following the advent of digitization of the slides. The digitized image is considered as a 2-D scene \( I_s \), represented as a matrix \( M \), consisting of pixel intensity values of RGB components. It is defined as \( I_s = (M, \omega) \), where \( \omega = \chi(x,y) \) is the pixel intensity function representing a vector \( \omega \in M \), consisting of intensity levels of red, green and blue components. Various imaging techniques have been presented in the literature, addressing the object detection, segmentation and followed
by classification. Segmentation of nuclei is an important phase in the process of extracting the pleomorphic features. Since the boundary to be extracted is irregular and presents discontinuity, many low level approaches fail in segmenting the nuclei to the required accuracy. Hence, due to complex morphological features and heterogeneity of the image, segmentation of the nuclei is considered to be a challenging task.

**Existing Literature**

Most of the works presented in the literature have highlighted significance of various features and similarity measures in addressing the object detection in an overlapped region and segmenting the irregular boundary. An unsupervised segmentation has been presented based on the features computed using magnitude and spectra in the frequency domain. A morphologically seeded watershed based method has been proposed to extract the overlapped nuclei. In a work presented, a Gaussian based hierarchical voting and repulsive balloon model has been used for a cell segmentation. An improved hybrid active contour model driven by both boundary and region information is used for an effective nuclear segmentation. In the work presented by, a multi-scale radial line scanning has been proposed to delineate the boundaries of nuclei detected using Laplacian of Gaussian kernels. An integrated model of adaptive morphology and curvature scale space has been used to segment the overlapped cells. A color decomposition based active contours and a sparse shape prior and occlusion constraint based levels sets have also been proposed for a robust nuclei segmentation. A large feature set based adaboost classifier technique has been used to perform nuclear detection. There are methods based on deep convolution networks applied to perform nuclear segmentation. The basic idea is to evolve the contour \( u \), which is represented as a polynomial function in a level set functional model. The level set function

\[
A_{\text{avg}} = \frac{\text{area of nuclei}}{\text{mean area of the nuclei computed using the shape prior model}}
\]

... (1)

where \( A_{\text{region}} \) is the area of the occluded region and \( A_{\text{avg}} \) is the mean area of the nuclei computed using the shape prior model. Through this approach an approximate centroid of each nuclei present in the occluded region is obtained, hence detecting the existence of nuclei.

Second, the segmentation of the nuclei boundary through the evolution of contours implemented as multiple level sets of distance regularized level set (DRLS) function. Active contours, originally proposed as energy minimizing deformable models have been considered to be most effective in segmenting the irregular boundaries. The basic idea is to evolve the contour \( u \), which is represented as a polynomial function in a level set functional model. The level set function...
$\Phi E(u)$ is represented as a partial differential equation as given in Eq.2

$$\frac{\partial u}{\partial t} + f_s(\nabla |\nabla u|) = 0 \tag{2}$$

Where $f_s$ is the function which computes the gradient information. The evolution of the contour towards the object boundary is controlled by the gradient information, obtained as the function of $E_r$ and $E_u$ representing the energy gradient of image and the contour respectively. As the contour evolves towards the boundary, the energy difference reduces to null value. Various active contour models have been proposed in the existing literature emphasizing the importance of gradient computation to drive the contours effectively. An active shape model based on a statistical approach constrained by point distribution\textsuperscript{20}, has been presented to drive the active contours effectively\textsuperscript{21}. A multiple level set implementation based on both region and edge gradient is also presented\textsuperscript{22}. Geodesic active contours\textsuperscript{23} have been shown to be quite effective segmentation approach\textsuperscript{24}. A novel method of computing adaptive energy and integrating the shape, region and the boundary features have been presented\textsuperscript{25,26}. A region gradient based active model has also been proposed\textsuperscript{27}, which is based on an energy minimizing model\textsuperscript{28}.

In this research, an improved edge based distance regularized level set (DRLS)\textsuperscript{29} active contour model has been adapted. Here, two important terms viz., forward and backward diffusion effects of contour evolution have been integrated in a distance regularization model. It results in reduced initialization with fewer iterations of contour evolution. The DRLS model is as shown in Eq.3

$$\frac{\partial u}{\partial t} = \mu \text{div} (\eta_{pr} \nabla \Phi) - \frac{\partial E}{\partial \Phi} \tag{3}$$

Here the distance regularization term is as shown in Eq.4

$$\frac{\partial \Phi}{\partial t} = \mu \text{div} (\eta_{pv} \nabla \Phi) \nabla \Phi \tag{4}$$

**Table 1.** List of symbols and notations used in this research

| Symbol | Description | Symbol | Description |
|--------|-------------|--------|-------------|
| $I$    | 2-D image scene | $M$    | 2-D image grid |
| $\chi$ | Pixel intensity function | $\omega$ | RGB vector |
| $C$    | Set of centroids | $c_i$  | Centroid of each nuclei |
| $X_i$  | Points of intersecting norms | $N_c$  | Set of norms orthogonal to the tangent over the curve |
| $T_i$  | Set of the tangents drawn on the boundary points of the curve ($B_i$) | $A_{\text{ROI}}$ | Total area of ROI |
| $A_{\text{avg}}$ | Average nuclei area | $\phi(u)$ | Level set function of the contour $u$ |
| $f_{sv}$ | Gradient function | $D$   | Diffusion rate of DRLS |
| $F_{\text{roi}}$ | Foreground region of interest | $I$  | Energy gradient |
| $s_{sv}$ | Signal frequency of the image | $N$  | Noise component |
| $D_{\text{ef}}$ | Denoised signal co-efficient | $\beta, \gamma$ | Structural elements for erosion and dilation |
The term $D = \mu \eta_p \nabla^2 \phi$ is the diffusion rate controlled by a positive or negative potential value $p \eta$ indicating the forward and backward diffusion. The second term is the derivative of the external energy.

The basic idea of the work presented in this research is to obtain the region of interest by adapting a shape prior model to morphologically extract the foreground regions $F_{roi}$. A novel geometrical approach of obtaining the set of norms $n_i \in N_c$ to the curvature orthogonal to a tangent $t_i$ drawn over each $i^{th}$ boundary point, is adapted to compute the set of centroids $C_g$ representing the geometrical centers of each nuclei present in the region. The resultant of morphological processing based on the proposed shape prior is adapted to compute the energy gradient $I_g$, which represents the external driving force for the contour evolution. Hence, the DRLS contours implemented as multiple level sets are initialized at the centroids detected and made to evolve using the shape prior based gradient computed. Subsequent sections provide the description of the data set used and a detailed discussion of the various stages of the methodology, followed by experimentation and result analysis.

**MATERIALS AND METHODS**

In this section, a description of the data set used followed by a detailed discussion on each stages of the methodology is presented.

**Description of the dataset**

The digitized images of H&E stained histopathological slides have been obtained from a standard collection provided by BreakHis dataset\(^3^\). The dataset represents the samples of surgical open biopsy(SOB) of benign and malignant breast cancer tissues. A total of 200 images representing adenosis of benign and ductal carcinoma of malignant samples, at a zooming level of 400x have been chosen for experimentation. Fig.1 shows the sample images of both clinical representations.

The images represented as a 2-D grid $M$, are considered at a resolution of 700x460. The various stages of methodology, as discussed in the

---

**Algorithm 1: Object-Seg-Contour**

| INPUT | Input RGB Image |
|-------|-----------------|
| OUTPUT| Active contours segmenting the objects of interest |

1. begin
2. // Enhancement of Image quality using filtering
3. DWT-wiener-filter();
4. // Computation of geometrical measure to fix the shape prior
5. Compute.shape.prior.measures();
6. // Morphological operation using shape priors
7. erosion.dilation.reconstruction();
8. // Curvature-Norm technique to detect the number of centroids
9. for $x$ on the curvature-boundary do
10. | draw-tangent();
11. | compute-tangent-norm();
12. end
13. // Based on the density of the intersecting norms, compute the centroids using k-means
14. k-means();
15. // DRLS initialization and evolution
16. Initialize-contour();
17. while contour == objectboundary do
18. | Evolve-contour();
19. end
20. end
following section, have been applied to achieve the above listed objectives. Throughout this research many notation have been used as listed in Table 1.

**Proposed methodology**

In this section, various stages of the proposed method has been presented in detail. Algorithm 1 shows the complete illustration of the various phases involved.

**Image enhancement**

The H&E stained image inherently presents both low and high frequency noise components, due to staining and zooming errors. Hence, Wiener filter is considered to be the promising technique in eliminating both the components. The frequency sub-bands, $s_i$, are separated from noise components $N=(s_{low}, s_{high})$ and a denoised signal co-efficient $D_{cf}$ can be computed as shown in Eq. 5:

$$D_{cf} = \mu + \frac{s^2 - N}{\sigma^2} (s_i - \mu)$$

...(5)

Where $\mu$ and $\sigma^2$ represents mean and variance of the signal component computed as shown in Eq. 6 and 7:

$$\mu = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} s_{ij}$$

...(6)

**Fig. 2.** Results of curvature-norm method for centroid detection (a) Occluded region (b) norms on the curvature points and (c) The centroids computed.

**Fig. 3.** The results of each stages of the methodology (a) Original Image (b) Region of interest using k-means (c) morphological enhancement (d) foreground region of interest with centroids detected (e) The occluded regions (f) DRLS initialized at the centroids (g) Evolving contours (h) Final Segmentation.
\[ \sigma^2 = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} (s_{ij} - \mu_{ij})^2 \]  
\[ s_{low} = \sum_{j=-\infty}^{\infty} s_{ij}(2^{n-j}) \]  
\[ s_{high} = \sum_{j=-\infty}^{\infty} s_{ij}(2^{n-j}) \]

A discrete wavelet transform (DWT) of the original signal component splits the signal into various spatial bands and separates both low and high frequency noise bands from the image as shown in Eq. 8 and 9.

The idea of integrating Weiner filter with DWT results in an efficient filtering of the noise components\(^\text{31}\). It has also been shown that DWT results in reduced over-segmentation\(^\text{32}\).

Computation of Shape prior

In this phase, the foreground region of interest \( F_{roi} \) is extracted using a suitable shape prior model representing the structural element \( \beta \) and \( \gamma \), to compute the area of the nuclei objects to be detected. The shape prior model is as shown as shown in Eq. 10 and 11.

\[ \beta = \mu_D - \sigma_D \]  
\[ \gamma = \mu_D + a \sigma_D \]

The parameters \( \mu_D \) and \( \sigma_D \) are the mean and standard deviation of the nuclei diameter, which is derived from the mean area of the foreground objects obtained from the outcome of the clusters generated using k-means algorithm. The area computed by \( \beta \) and \( \gamma \) are used with erosion and dilation process respectively. The parameter \( \pm \) corresponds to the thresholding factor for dilation. These morphological operations generates the foreground scene from which the image gradient \( I_g \) is computed. Binarization of the same results in the extraction of ROI.

Centroid detection

After obtaining the ROI from the previous stage, the detection of the existence of the nuclei is performed in this stage. As presented earlier, the novel curvature-norm technique is adapted to extract the centroid of the nuclei. Fig. 2 shows the outcome of the method over an occluded region. It shows the generation of norms over the boundary points of the curve and finally showing the centroid points of the number of nuclei present in the region.

The computation of the centroids is achieved using k-means clustering applied over the intersecting points of the norms. Here, the value of \( k \) is computed as shown in Eq. 1. Finally, the DRLS contours are initialized as multiple level set functions at the centroids detected and

|      | SN  | SP  | PPV | OR  |
|------|-----|-----|-----|-----|
| DRLS-CN | 0.97 | 0.74 | 0.82 | 0.78 |
| DRLS   | 0.92 | 0.66 | 0.81 | 0.70 |

Fig. 4. Charts showing the object detection accuracy and overlap resolution by comparing DRLS-CN and DRLS with actual count
the evolution of the same is guided by the image gradient $I_g$, which is computed as discussed above.

RESULTS AND DISCUSSION

The proposed methodology is experimented over 200 images selected from both benign and malignant samples in the dataset chosen. Fig. 3 shows the results of each stages and the final outcome of the segmentation approach proposed.

The proposed method DRLS with curvature-norm initialization (DRLS-CN) outperforms the DRLS without curvature-norm initialization. **Quantitative analysis**

The efficacy of the proposed methodology has been studied using following two quantitative measures. First, object detection and occlusion resolution measures and, second, segmentation accuracy based on boundary error metrics. The object detection measures are computed using sensitivity (SN), specificity (SP), positive predictive value (PPV), and the overlap resolution (OR). Based on the ground truth, the above measures are computed using true positive (TP), true negative (TN), false positive (FP) and false negative (FN). Since, the manual delineation performed by the pathologist is tedious, only 40 samples have been considered for quantitative analysis. A comparative results of DRLS segmentation with and without curvature-norm initialization is presented in Table 2.

The chart shown in Fig. 4 indicates the other two measures of object detection and overlap resolution viz., actual count (AC) and detected count (DC). These measures are computed by taking the average of 20 randomly chosen objects.

Second, the measures of segmentation accuracy is computed using following two metrics. They are Hausdorff distance (HD) and Mean absolute distance (MAD) as shown in Eq. 12 and Eq. 13.

![Fig. 5. Charts showing the comparison of Hausdorff distance (HD) between DRLS-CN and GAC-CN](image1)

![Fig. 6. Charts showing the comparison of Mean Absolute distance (MAD) between DRLS-CN and GAC-CN](image2)
The key factor for computing the above measures is the distance in terms of pixel difference between the manual delineation performed over the object boundary and the final contour. Since the manual delineation is a tedious task, pathologists have randomly chosen 20 objects for ground truth generation. These measures have been plotted in the charts shown in Fig.5 and Fig.6, in comparison with the Geodesic active contours driven by curvature-norm initialization. The proposed DRLS-CN has shown a very less pixel difference of utmost 4 pixels in contrast with that of GAC-CN, which measures in a range of 2-14 pixels. It is clearly evident from the result, that DRLS-CN outperforms GAC-CN in terms of segmentation accuracy.

CONCLUSION

The work presented in this research, has been able to address the importance of extracting the pleomorphic features for the purpose of diagnosis and prognosis of the cancer disease. An improved active contour technique has been adapted to perform the challenging task of segmenting the nuclei from an occluded region of a digitized H&E stained image. Here, a novel curvature-norm technique is devised to compute the geometric centroid of the occluded object and the DRLS contours are initialized at those centroids to evolve towards the object boundary to segment the nuclei shape resulting in pleomorphic features. Initially, the region of interest is extracted using a novel shape prior model, which is also used to compute the image gradient, representing the external energy in driving the DRLS contour efficiently towards the object boundary. Hence, this work has been able to present two novel techniques. First, The shape prior model and second, the curvature-norm technique for centroid detection. The results of the segmentation have been compared with other techniques and found to be quite promising in terms of object detection, overlap resolution and also with respect to segmentation accuracy. Further, the results can be extended for the post segmentation classification process.

REFERENCES

1. A. M. Khan, H. El-Daly, E. Simmons, and N. M. Rajpoot, “A hybrid magnitude-phase approach to unsupervised segmentation of tumor areas in breast cancer histology images,” J. Pathol. Inform., 4(7): (2013).
2. J. Shu, H. Fu, G. Qiu, P. Kaye, and M. Ilyas, “Segmenting overlapping cell nuclei in digital histopathology images,” in Proc. IEEE 35th Annu. Int. Conf. Eng. Med. Biol. Soc., Osaka, Japan, 3–7 Jul, pp. 5445–5448 (2013).
3. Zhang, X., Xing, F., Su, H., Yang, L. and Zhang, S., High-throughput histopathological image analysis via robust cell segmentation and hashing. Medical image analysis, 26(1), pp.306-315 (2015).
4. Jing, J., Wan, T., Cao, J. and Qin, Z., An improved hybrid active contour model for nuclear segmentation on breast cancer histopathology. In Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on (pp. 1155-1158). IEEE (2016).
5. Xu, H., Wang, H., Berendt, R., Jha, N. and Mandal, M., November. Automated nuclear segmentation in skin histopathological images using multi-scale radial line scanning. In Healthcare Innovation Point-Of-Care Technologies Conference (HI-POCT), 2016 IEEE (pp. 175-178). IEEE (2016).
6. Xu, H., Lu, C., Berendt, R., Jha, N. and Mandal, M., Automatic Nuclear Segmentation Using Multiscale Radial Line Scanning With Dynamic Programming. IEEE Transactions on Biomedical Engineering, 64(10), pp.2475-2485 (2017).
7. Wang, P., Hu, X., Li, Y., Liu, Q. and Zhu, X., Automatic cell nuclei segmentation and classification of breast cancer histopathology images. Signal Processing, 122, pp.1-13 (2016).
8. Prasath, V.S., Fukuma, K., Aronow, B.J. and Kawanaka, H., Cell nuclei segmentation in glioma histopathology images with color decomposition based active contours. In Bioinformatics and Biomedicine (BIBM), 2015 IEEE International Conference on (pp. 1734-1736). IEEE (2015).
9. Zhang, P., Wang, F., Teodoro, G., Liang, Y., Brat, D. and Kong, J., Automated level set segmentation of histopathologic cells with sparse shape prior support and dynamic occlusion constraint. In Proceedings. IEEE International Symposium on Biomedical Imaging, p. 718). NIH Public Access (2017).
10. J. Vink, M. V. Leeuwen, C. V. Deurzen,
21. S. Ali and A. Madabhushi, “Active contour for overlap resolution using watersheds based initialization (acorew): Applications to histopathology,” in Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on, 614–617, IEEE (2011).

22. S. Ali and A. Madabhushi, “Segmenting multiple overlapping objects via a hybrid active contour model incorporating shape priors: applications to digital pathology,” in SPIE Medical Imaging, 7962W–7962W, International Society for Optics and Photonics (2011).