Optimizing the Cervix Cytological Examination based on Deep Learning and Dynamic Shape Modelling

Afaf Tareef[a,*], Yang Song[a], Heng Huang[b], Yue Wang[c], Dagan Feng[a], Mei Chen[d,e], Weidong Cai[a]

[a] *Biomedical and Multimedia Information Technology (BMIT) Research Group, School of Information Technologies, University of Sydney, Australia.*
[b] *Department of Computer Science and Engineering, University of Texas at Arlington, USA.*
[c] *Bradley Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, USA.*
[d] *Computer Engineering Department, University of Albany State University of New York, USA*
[e] *Robotics Institute, Carnegie Mellon University, USA*

Abstract

The task of segmenting nuclei and cytoplasm in Papanicolau smear images is one of the most challenging tasks in automated cervix cytological analysis owing to the high degree of overlapping, the multiform shape of the cells and their complex structures resulting from inconsistent staining, poor contrast, and the presence of inflammatory cells. This article presents a robust continuous variational segmentation framework based on convolutional neutral network and a learned shape prior enabling an accurate analysis of overlapping cervical mass. The shape prior is dynamically modelled during the segmentation process as a weighted linear combination of shape templates from an over-complete shape dictionary under sparsity constraints. We provide quantitative and qualitative assessment of the proposed method using two databases of 153 cervical cytology images, with 870 cells in total, synthesised by accumulating real isolated cervical cells to generate overlapping cellular masses with a varying number of cells and degree of overlap. The experimental results have demonstrated that our methodology can successfully segment nuclei and cytoplasm from highly overlapping mass. Our segmentation is also competitive when compared to the

*Corresponding author
Email address: atar8654@uni.sydney.edu.au (Afaf Tareef)

state-of-the-art methods.

Keywords: Overlapping cell segmentation, convolutional neural network, feature learning, sparse reconstruction, level set evolution.

1. Introduction

Cervical cancer is the fourth most common cause of cancer death in women worldwide, with more than 270 thousand deaths every year [1]. Fortunately, this cancer can be easily detected and preventable in its early stage by Pap smear test, where a sample of cells is collected from the vagina and the neck of the uterus, and examined under a microscope, to identify the abnormalities of cell number, shape, and size. Pap test is currently a manual screening method performed by a cytologists or pathologist, hence, it is a tedious and time-consuming process. The accurate segmentation of cervical nuclei and cytoplasm is the most challenging step toward developing an automatic machine-assisted screening and diagnosing system for cervical cancer, which is complicated by the complex structure of cervical cells, high overlapping degree, and presence of mucus, blood, and inflammatory cells in Pap smears [2]. Figure 1 shows a snapshot of microscopic images obtained from the Pap smear test, and samples of overlapping cervical cells. It is observed that there are many overlapping cells with fuzzy contour, making it extremely difficult and time consuming for humans to delineate the cells, and also challenging to design a fully automatic segmentation method.

Recently, various methods have been proposed in the literature to segment the isolated, touching, and overlapping cervical cells. These cervical cell segmentation methods can be classified into two main categories based on the segmented cellular components: (1) nuclei segmentation methods to detect only the nuclei boundaries in isolated or overlapping smear cells, such as [3, 4, 5, 6, 7, 8, 9]; (2) both nuclei and cytoplasm segmentation methods to delineate the nucleus and cytoplasm contours in either isolated or overlapping smear cells [10, 11, 12, 13, 14, 15].

Several methods for segmenting nuclei of isolated or partially overlapping cells, were proposed based on active contours [3], level set [6], watershed transform [4, 9], unsupervised classification [5], deep learning with Graph Partitioning [8], and shape modeling [7, 16]. For instance, Bergmeir et al. [6] introduced a semi-automatic approach to segment nuclei in high-resolution Pap images under the expert control, in which nuclei candidates were determined by Canny edge detector and Hough transform, and then processed by a level set evolution. A nuclei shape modeling approach, combining local features of the nuclei boundary and a priori knowledge of the expected nuclei shape trained on isolated nuclei images, is proposed [7] to segment the overlapping nuclei. Another shape modeling approach is introduced in [16] to approximate the nuclei shape based on nucleus-level information and ellipse fitting method. These approaches, however, do not segment the cell cytoplasm, whose shape and size are substantial information for accurate diagnosis.

A number of methods have been adopted to segment single and touched cytoplasm from cervical smear images. Earlier researchers in this field used thresholding techniques, such as [17], which often resulted in unsatisfactory results due to the complex structure of cervix cells from poor contrast and variable staining. Marker-based and multi-scale watersheds have also been used to segment the cytoplasm [12]. However, it could be difficult to find a representative marker for each cell and result in over-segmentation. Unsupervised classification is another option that has been applied to single cell segmentation [11]. Other widely used segmentation methods include active contour models (ACM) with edge- and region-based models [10, 18] due to their ability to recover closed object boundaries with pixel accuracy. However, all of these techniques extract the whole cellular mass consisting of a number of cells, which are however insufficient for shape analysis.

In the last few years, the development of a complete segmentation techniques for both nuclei and associated cytoplasm from overlapping cells has drawn the attention of research groups over the world. Several segmentation methods for partially overlapping cells, have been designed based on edge enhancement tech-

niques [19], geodesic active contour [20, 21], watershed transform [15, 13, 22], sliding band filter [23], and cell shape formation [24, 25]. However, these methods cannot work with highly overlapping cells. The segmentation of highly overlapping cells is still the most challenging problem, in which further research must be undertaken to validate automatic schemes providing precise segmentation of overlapping nuclei and cytoplasm.

Recently, the first "Overlapping Cervical Cytology Image Segmentation Challenge" [26] was held in conjunction with the IEEE International Symposium on Biomedical Imaging (ISBI) in 2014. The challenge covered two issues, the automatic nuclei detection, and individual cytoplasm segmentation from cellular mass, with different numbers of cells and overlapping ratios. There are two approaches, along with two versions of the baseline method developed by the challenge organizers, passed the challenge with promising results. The first challenge winner was a Voronoi diagram-based segmentation method proposed by Ushizima et al. [27]. In this approach, nuclear narrow-band seeding and graph-based region growing were applied for nuclei detection, whereas Voronoi diagrams were used to segment the cytoplasm of the overlapping cells. However, this approach separates the overlapping cells with straight lines, which do not represent the true cell boundaries. The second challenge winner [28] and the baseline methods [14, 29] were based on the regularized level set evolution with elliptical shape assumption. An enhanced version of the second approach was presented in [30] with a star shape prior for optimizing the segmentation performance of overlapping cells. These approaches proved that incorporating shape priors into the parametric segmentation procedures can obviously enhance the segmentation results of overlapping cells. However, these shape priors are too simplified to represent the actual shape of the cervical cells.

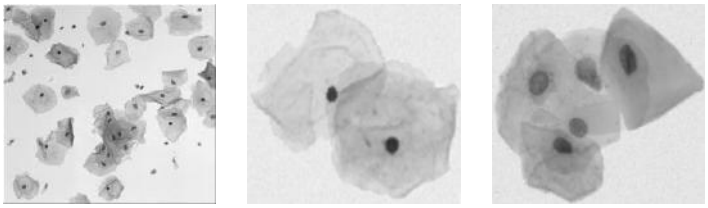
In our preliminary version of the method presented in this paper [31], we proposed dynamically generated shape prior based on sparse approximation to segment the overlapping cytoplasm. Sparse approximation (SA) has emerged as an effective solution to many problems, such as pattern recognition [32], face recognition [33, 34] and object tracking [35, 36]. Recently, SA has been applied

to medical image segmentation [37, 38, 39, 40, 41] to infer/refine the contours of organs with typical shapes, e.g., the lung and liver. It is more challenging to derive such these approaches for cervical cell segmentation due to the large variation in shapes and sizes of cells. In [31], we design an approximated shape prior that is incorporated into a continuous variational method and iteratively update during evolution to generate the best representation of cells. In addition, prior to the cytoplasm segmentation stage, we employed the multilayer perceptron (MLP) neural network with ten hand-engineered shape and texture features for cellular component classification. One drawback of this classification method is that the hand-engineered features might not generalize well to new datasets, when there are different contrast degrees, noise levels, or different nuclei characteristics.

In this paper, we present an automatic method based on deep learning and dynamic shape model for segmenting the individual nuclei and cytoplasm from a large cellular mass with overlapping cells. The proposed segmentation method first partitions the Pap smear images into three cellular components, i.e., background, nuclei, and cytoplasmic mass, using a convolutional neural network (CNN); and then separates the individual cytoplasm inside each cellular mass in a variational segmentation framework with a learned shape prior that is iteratively updated. Different from our conference paper [31], we enhance the cellular component classification method with a learning-based approach. This approach helps to promote the nuclei detection performance and address the drawback of sensitivity and poor generalization to new datasets. Subsequently, the accuracy of overlapping cytoplasm segmentation is also improved and the method could better handle the difficult cases. In addition, we also provide more methodological details of the shape learning procedures, and more thorough experimental results and discussion.

The rest of this paper is organized as follows. In Section 2, the proposed segmentation methodology is presented. In Section 3, the image datasets used in our experiments and evaluation metrics are described. Experimental results and discussion are given in Section 4. Finally, conclusions and future work are

presented in Section 5.



(a) (b) (c)

Figure 1: (a) Pap Smear Image, and samples of the cervical overlapping cells from two datasets: (b) the test dataset of [14], and (c) ISBI challenge test dataset [26].

2. Methodology

The workflow of the proposed learning methodology is illustrated in Figure 2. The proposed method consists of two main phases: cellular components classification with region clustering and learned features based on convolutional network, and 2) individual cytoplasm segmentation from overlapping cellular mass with Voronoi segmentation and dynamic shape prior-based level set evolution.

2.1. Cellular Components Separation

This phase aims to divide the image into initial cellular components: background, nuclei, and cellular cytoplasmic masses without separating the cytoplasm of different cells. To do this, an image is first partitioned into fixed-size patches, and then CNN-based feature learning and classification are performed on the patches to identify the various cellular components.

2.1.1. Generation of image Patches

This step begins with region clustering, where the Pap image is tessellated into superpixels taking into account the intensity similarities and spatial proximity. There are many clustering methods in the literature. In this article, the simple linear iterative clustering technique (SLIC) [42] is adopted due to

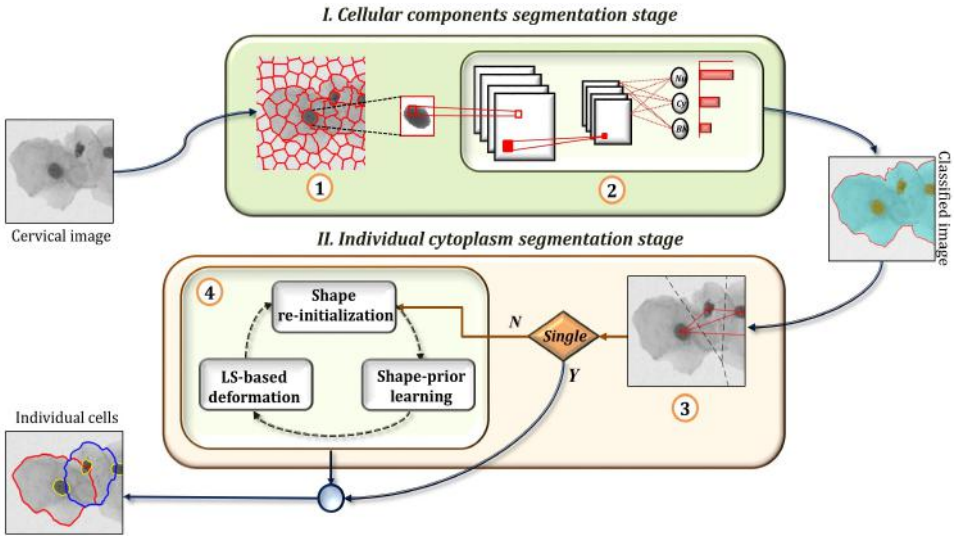


Figure 2: The workflow of the proposed learning methodology. (I) is the 3-class cellular components classification stage, including (1) cellular patches generation and (2) CNN-based classification. (II) is the individual cytoplasm segmentation stage, including (3) Voronoi segmentation and (4) learned shape prior-based evolution.

two reasons: 1) it can produce superpixels with approximately equal sizes and regular shapes, 2) it is computationally more efficient compared to many other clustering algorithms. However, the parameter setting for this algorithm is very important and have to be carefully tuned to get the best results. There are two major parameters: the superpixel size (S) and the regularizer (r). S controls the size of each superpixel, whereas r controls the shape regularity. A smaller r leads to superpixels with more similar pixels but irregular shape, whereas a larger r provides smaller superpixels with regular shape (i.e., elliptical, hexagonal, and zigzag shape in our case). In our experiments, the S and r parameters were set empirically to 25 and 0.01, respectively.

Figure 3 shows the importance of choosing appropriate parameters to get more meaningful clustering. As seen, with suitable parameters, the superpixels of different cellular categories can be easily distinguished by its shape and intensity characteristics (e.g., the nuclear superpixel has an elliptical shape and

a lower intensity than the surrounding cytoplasm superpixels). However, we observed that some Pap images have isolated dark pixels located at the image border, which are gathered in separate vertical or horizontal line-shaped superpixels. These superpixels should belong to the background; however, they are not consistent with the actual background superpixels, leading to misclassification. To ensure consistency of each cellular category, we choose to remove the image border by cropping four pixels at each four directions of the image.

Then, to highlight the difference between Nuclei, cellular mass, and background, a refined superpixels map is generated. Specifically, Gaussian lowpass filter of size 20 with standard deviation of 1, and histogram equalization with 0.0005 threshold, are applied on the superpixels map to remove noise, and highlight the intensity difference. Next, the median value of each superpixel is assigned to all pixels in the region, thereby, all pixels belonging to the same category have rather similar texture. A fixed-size square patch is then generated by bounding box and bicubic interpolation algorithm [43], with a superpixel centralizes the patch and bordered with a clear background to highlight the superpixel shape information. Finally, the generated patches are passed as input images for feature learning and classification step. The generated patches have distinguishable shape and the intensity (e.g., Figure 4 (Input image)), which helps to get an accurate prediction of the patch category.

2.1.2. Classification with Learned Features

The next step is to classify the image patches into nuclei, background, or cellular cytoplasmic masses. The classification process is conducted using a convolutional neutral network (CNN) model. CNN is a type of deep learning algorithm, that shown to be effective for many tasks in computer vision, such as image classification [44, 45], face recognition [46], and handwriting character recognition [47]. The main advantage of the CNN model is its reliability in learning discriminative properties directly from the raw image and generating a prediction of image category, eliminating the need for traditional hand-crafted feature extractor, which is typically computationally intensive and requires prior

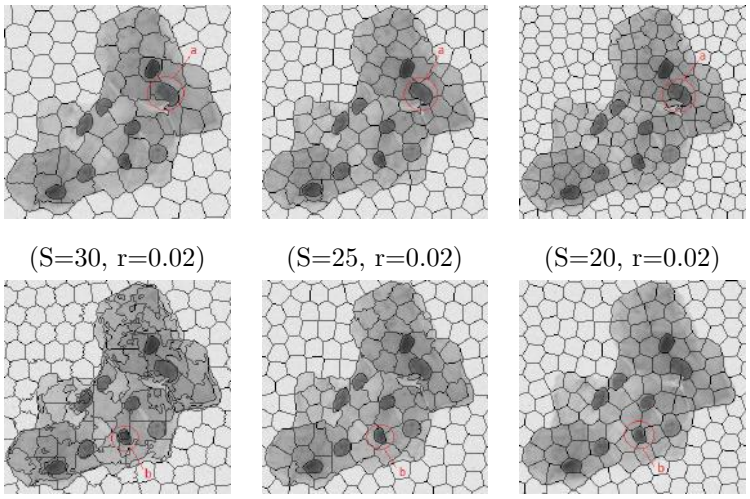


Figure 3: Influence of the SLIC parameters in cellular component clustering.

knowledge of the classification problem.

The standard CNN consists of several layers including an input layer of raw image, hidden layers which are usually convolutional and pooling layers, and an output layer representing the classes. The convolutional layer is the feature extractor consisting of a set of feature maps, computed by applying a series of convolution kernels to its input and passing the result through a sigmoid function. Pooling or subsampling layer reduces the resolution of the feature maps, thus, reducing the number of parameters, the memory consumption, and the required computation.

We designed a 7-layer CNN for our classification problem. The input layer contains ($50 \times 50 = 2500$) neurons corresponding to the image patches. The hidden layers include a convolutional layer of 20 convolutional filters (feature maps) with (5×5) pixel kernel window applied over the input patch, and Rectified Linear Unit (ReLU) layer [48] to increase the nonlinearities in the network, making the decision function more discriminative [49]. The next layer is a maximum pooling (max-pooling) layer with (2×2) subsampling ratios without overlap. According to [50], a maximum pooling improves generalization to data, and

leads to faster convergence rate comparing with other subsampling operations by selecting superior invariant features. Next is a fully-connected layer connecting every neuron in the max-pooling layer to each of its neurons. Then, a softmax layer is used to represent the categorical probability distribution over the three cellular component categories. Finally, the classification output layer of three neurons assigns a label, presenting one of the three cellular component classes , to the input patch.

Figure 4 displays the design of our CNN model, with the filter numbers and sizes. The CNN was trained with the stochastic gradient descent (SGD) with momentum algorithm. The training consists of 30 epochs, conducted with 45 training images, each with 441 patches. With the 3-way softmax function, the class of each patch is predicted. The output of the classification step is a map of labels indicating the nuclei (Nu), cytoplasmic mass (Cy), and background (BK) superpixels. In addition, we consider that during the clustering step, some boundary pixels belonging to two different cellular components may be clustered into a single superpixel. Therefore, we employed the level set evolution [56] to refine the regions contour, with 10 iterations for cellular mass, and 5 iterations for nuclei regions.

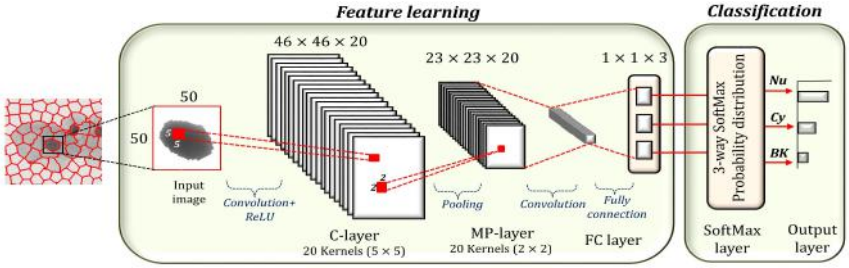


Figure 4: The structure of the employed CNN for cellular component separation, with a single convolutional (C-layer) and ReLU layer, max-pooling layer (MP-layer), fully connected (FC) layer, softmax and classification output layer, to predict whether the input patch is likely belonging to a nucleus (Nu), cytoplasm(Cy), or background (BK).

2.2. Individual Cytoplasm Segmentation

In Pap smear images, there are a large number of overlapping cells with complex structures and multiform shape, making the segmentation of individual cells a challenging problem. The main goal of this phase is to provide the actual contour for each individual cell inside the cellular mass. To this end, two segmentation stages are performed for each detected nucleus: initial segmentation by Voronoi diagram with Delauney Triangulation, followed by final segmentation based on level set with learned shape prior.

2.2.1. Initial Segmentation based on Voronoi Diagram

Extracting the initial rough contours of the individual cytoplasm is the basis for our next variational segmentation process. The initial segmentation is performed by computing the Voronoi polygons of the image seeds, which are determined as the intensity weighted centroids of the detected nuclei. Voronoi diagram (VD) [52] generates polygons of intensity values using information from the external boundary vertices of Delaunay triangulation (DT), or so-called J-triangle or junction triangle, between the seeds. DT are first established between the seeds (i.e., nuclei centroids), then, the convex polygon of boundary points that have the shortest Euclidean distances to the corresponding nucleus centroid, are computed (see Figure 5).

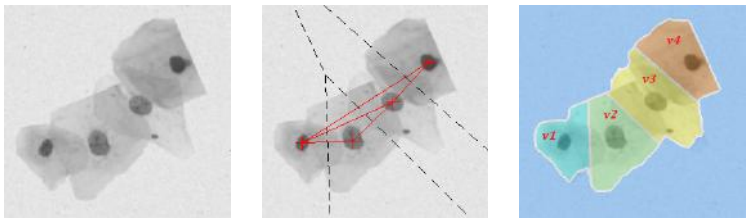
More formally, Given a set of 2D points representing the coordinates of n nuclei centroids $C = \{C_1, C_2, ..., C_n\}$, and $C_i, C_j \in C$ where $i \neq j$, then, the Voronoi cell V of a point C_i is the set of all the points that are closer to C_i than to any other nuclei centroids. This can be written as:

$$V(C_i) = \cap_{1 \leq j \leq n, j \neq i} \{p|\ d(p, C_i) \leq d(p, C_j)\} \tag{1}$$

where $d(p, C_i)$ is the Euclidean distance between the pixel p and corresponding centroid C_i. The set of all n V is called the Voronoi diagram VD of the given 2D points. To generate the rough Voronoi cells as shown in Figure 5 (c), the boundary points between the cellular mass and background are assigned to the nearest nucleus, and combined with the surrounding Voronoi diagram lines (i.e.,

black dash line in Figure 5 (b)), providing that the generated cell is completely inside the cellular mass (i.e., the cell must be convex).

By the end of this step, the segmentation process for the isolated cells is completed, and the Voronoi cell represents the individual cell. The Voronoi cells of overlapping cells are dilated with a disk of radius equals 20 to increase the search areas for the final shape prior-based segmentation process.



(a) (b) (c)

Figure 5: The rough segmentation process: (a) the original Pap image, (b) the Delauney triangle (red solid line) and the Voronoi polygons (black dash line) using four nuclei as seeds, and (c) the initial segmentation for each cell, i.e., $v1$, $v2$, $v3$, and $v4$.

2.2.2. Final segmentation based on level set with learned shape prior

This stage aims to reconstruct the final accurate boundaries of the individual cells based on the initial Voronoi segmentation of overlapping cells. Figure 6 illustrates the processes of the final segmentation stage, where (I) shows the training phase establishing the reference shape and the shape dictionary, and (II) represents the segmentation phase with prior-shape level set evolution based on the obtained Voronoi cells and pre-generated shape dictionary.

$\Rightarrow$ *Training Phase*

The final segmentation stage starts by generating an over-complete shape dictionary using a training set of annotated cell images. The shape vector for each cell is represented by the coordinates of a set of boundary points P enclosing

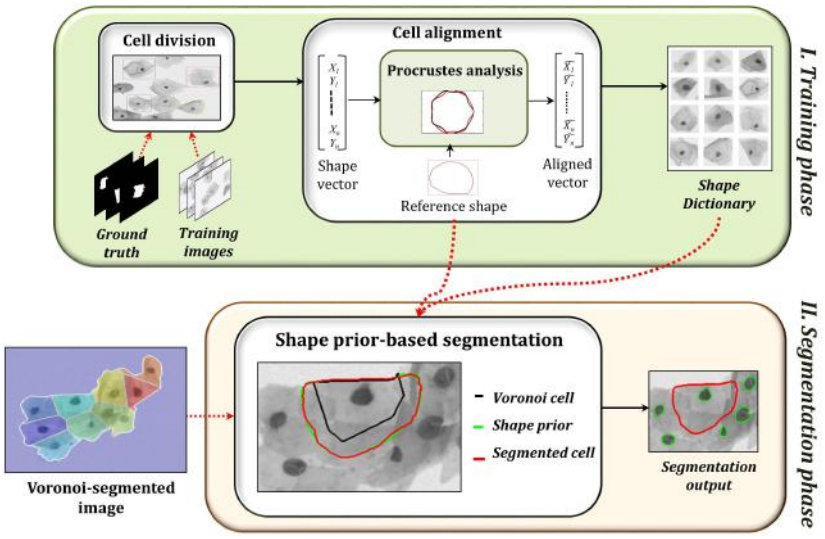


Figure 6: The final segmentation stage, with (I) training phase generating the shape dictionary, (II) segmentation phase based on pre-trained dictionary and Voronoi cells obtained from initial segmentation stage.

the cell, and given as following:

$$P = \{p_i \ |(r_i, \theta_i), \ i = \{1, ..., 360\}\} \tag{2}$$

Here (r_i, θ_i) indicates the polar coordinates of the point p_i, r_i indicates the radian between p_i and the corresponding nucleus center, $\theta_i \in [1, 2\pi]$ is the angle between the radial line crossing p_i and a reference radial axis. The radial line used for contour points retrieving has a pre-defined length determined based on the cell size. For our experiments, the line length is set to 100. There are three possible exceptional cases in this process. First, if the radial line crosses two or more points, then, the nearest point to the corresponding nucleus is retrieved. Second, if the radial line does not cross any point at some angles (this may occur when some nuclei are missed during classification, thus, the obtained Voronoi cell has a large size of two or more actual cells), then the last retrieved point p_{i-1} is rotated by angle $(\theta_i - \theta_{i-1})$ to complete the missed contour portion. Third, if the estimated contour point by the second case is located outside image border,

then, the border point at the same angle is retrieved instead.

After forming the shape vectors, they are transformed into the coordinate system of the mean cell shape using Procrustes analysis [51] to remove the geometrical translation, scale, and rotation effects. The shape dictionary is then represented by a matrix $\mathcal{D} \in \mathbb{R}^{M \times N}$, where each column refers to a single shape represented by $M/2$ boundary points with each point denoted by its xy coordinates, and N is the total number of training cell shapes in the dictionary $\mathcal{D}$, i.e., 270 cell shapes from the training dataset. The shape dictionary and the mean reference shape are passed to the next segmentation phase, consisting (See Figure 6).

$\Rightarrow$ *Segmentation Phase*

In segmentation phase, three shape-driven deformation steps are repeated (i.e., outer deformation loop equals 2): shape re-initialization, shape prior learning, and level set evolution with the learned shape prior (see Figure 2 (4)).

• Shape re-initialization. The shape re-initialization process is important for maintaining the shape approximation in the right track. This step generates the inputs for the next two steps, which are the shape vector with xy coordinates of the test cell contour for prior learning step, and the corresponding smooth cell mask for level set evolution step. At the beginning of this stage, the input cell c for this step is the Voronoi cell obtained from the initial segmentation. For the next outer iterations, the output cell from the level set evolution is the input of the re-initialization step. In this step, the shape vector v is generated with Eq. (2) as described in 'Training phase'. The shape re-initialization excludes the points outside the reasonable cell area, by means of the radial line length used for retrieving contour points. The smooth cell mask is generated by XOR-combining of input shape points with the boundary points of the cellular mass that are closer to the cell-in-focus than any other cells. These shape points are then connected by the moving average filtering (MAF) [53] and closing operation to guarantee the shape regularity and connectivity. The convex mask

generated by these connected points builds the initial level set function for next level set evolution step.

• Shape prior learning. In this step, the shape prior is established based on the shape vector v and the shape dictionary $\mathcal{D}$. Specifically, each shape vector $v \in \mathbb{R}^M$ is sparsely reconstructed as a weighted linear combination of a few templates from $\mathcal{D}$, as follows:

$$s = \arg\min_{s} \kappa \|s\|_1 + \|v - \mathcal{D}s\|_2^2 \tag{3}$$

where $\kappa > 0$ is a regularization parameter automatically selected, and s is a sparse weighting vector with few significant entries corresponding to the most representative shape templates and their weights in approximating v computed by the temporally correlated multiple sparse Bayesian learning (T-MSBL) algorithm [54]. The shape prior vector is then obtained by $v_{\mathcal{P}} = \mathcal{D}s$.

Finally, the shape prior mask $B_{\mathcal{P}}$, used to build the prior level set function, is generated and connected by MAF and closing operations, and then Procrustes transformed to the coordinate system of the input cell c. For some cases, the shape points of $v_{\mathcal{P}}$ are not sufficient to generate a connected mask, as the distance between some sequent points in the reconstructed vector become higher than being padded by the earlier connection process. To handle this issue, a binary image of $v_{\mathcal{P}}$ points is used to generate a convex hull image, used as initial mask by edge-based active contour to segment the image into foreground and background, where the foreground represents the shape prior mask.

• Shape prior-based level set evolution. In this step, a separate level set function (LSF): ϕ_i is built for each initial cell shape $i \in \{1, ..., N\}$, where N is the number of initialized cell shapes. Let $\phi(x, y, t) : \Omega \to \mathbb{R}$ represents a 2D time dependent LSF on the image domain Ω. Then, the energy function $E(\{\phi_i\}_{i=1}^N)$ is constrained by several terms including shape prior, area, data-

driven, regularization terms. $E(\phi_i)$ can be written as:

$$E(\phi_i) = \lambda_1 E_{\mathcal{P}}(\phi_i) + \lambda_2 E_{\mathcal{A}}(\phi_i) + \lambda_3 E_{\mathcal{D}}(\phi_i) + \lambda_4 E_{\mathcal{R}}(\phi_i) \tag{4}$$

where $\lambda_1, \lambda_2 \in \mathbb{R}$ and $\lambda_3, \lambda_4 > 0$ are constant weights balancing the contribution of each energy term, with λ_1=0.15, λ_2=-5, λ_3=4, and λ_4=0.2 for 20 iterations, for our experiments. The first term $E_{\mathcal{P}}(\phi_i)$ is the shape prior term [55] that constrains the possible cell shape, and is defined as follows:

$$E_{\mathcal{P}}(\phi_i) = \int_{\Omega} g(x) H(-\mathcal{P}(\phi_i(x))) dx \tag{5}$$

where g is the stopping function defined as $g(x) = 1/1 + (|\nabla G_\sigma I|^2)$, G_σ is the Gaussian kernel with standard deviation σ, and I is the image on a domain Ω. $H(.)$ is the Heaviside function, and $\mathcal{P}(\phi_i)$ is the level set function of the shape prior $B_{\mathcal{P}}$ generated in the last step, where $\mathcal{P}(x) > 0$ if the pixel x is inside $B_{\mathcal{P}}$ boundaries and $\mathcal{P}(x) < 0$ otherwise.

$E_{\mathcal{A}}(\phi_i)$ is the area term computing the segmentation area of $\phi_i < 0$ used to speed up the motion of the zero level contour in the level set evolution. $E_{\mathcal{A}}(\phi_i)$ is defined as:

$$E_{\mathcal{A}}(\phi_i) = \int_{\Omega} g(x) H(-\phi_i(x)) dx \tag{6}$$

$E_{\mathcal{D}}(\phi_i)$ is the data-driven term driving the segmenting curve to the object boundaries by having a lower energy when the zero level contour of ϕ_i is located at the cell boundaries. $E_{\mathcal{D}}(\phi_i)$ is defined as:

$$E_{\mathcal{D}}(\phi_i) = \int_{\Omega} g(x) \delta(\phi_i(x)) |\nabla \phi_i(x)| dx \tag{7}$$

where $\delta(.)$ is the Dirac delta function. $E_{\mathcal{R}}(\phi_i)$ is the regularization term [56], which ensures the smoothness of the segmentation boundaries by maintaining the signed distance property $|\nabla \phi| = 1$, and is defined as follows:

$$E_{\mathcal{R}}(\phi_i) = \int_{\Omega} p(|\nabla \phi_i(x)|) dx \tag{8}$$

where $p : [0, \infty) \rightarrow \mathbb{R}$ is a potential or energy density function, e.g., $p(e) = 0.5(e-1)^2$ [56]. Finally, the energy functional E is minimized by solving the gradient descent flow for each LSF $\{\phi_i\}_{i=1}^N$, as follows:

$$\begin{aligned}\frac{\partial \phi_i}{\partial t} &= -\frac{\partial E(\{\phi_i\}_{i=1}^N)}{\partial \phi_i} \\ &= -\lambda_1 \frac{\partial E_{\mathcal{P}}(\phi_i)}{\partial \phi_i} - \lambda_2 \frac{\partial E_{\mathcal{A}}(\phi_i)}{\partial \phi_i} - \lambda_3 \frac{\partial E_{\mathcal{D}}(\phi_i)}{\partial \phi_i} - \lambda_4 \frac{\partial E_{\mathcal{R}}(\phi_i)}{\partial \phi_i}\end{aligned} \tag{9}$$

where $\partial E(\{\phi_i\}_{i=1}^N)/\partial \phi_i$ is the Gâteaux derivative of the functional E with respect to $\{\phi_i\}_{i=1}^N$. At the end of this stage, the false cell segmentation is determined based on the morphological characteristics of the cytoplasm candidates. In particular, the cytoplasm candidates that are smaller or larger than being a cell (i.e., not in the range [2000, 16000] pixels for the ISBI dataset), or those have irregular-shape (i.e., eccentricity larger than 0.95) are replaced with the initial segmentation of the cell.

3. Materials and Experiments

3.1. Image datasets

In this study, we used two databases from the ISBI 2014 "Overlapping Cervical Cytology Image Segmentation Challenge" [26]. Both datasets consist of synthetic cervical cytology images, with a varying number of cells and degree of cell overlap, which were generated using isolated cells from non-overlapping fields of view (FOV) images. For each FOV, A stack of images from multiple focal planes was acquired with a focal depth separation of $1\mu m$, and then converted to a single extended depth of field (EDF) image where all cellular components are in focus. The nuclei, cytoplasm and the background regions for isolated cells from samples of EDF images were manually annotated. Then, synthetic images of 512×512 pixels were constructed by applying a random rigid transform (i.e., rotation, translation and scale) and a random linear brightness transform on the annotated isolated cells, and located them on the synthetic

image using an alpha channel (sampling from 0:88 to 0:99), providing that they were overlapped with varying overlap coefficients [2].

The first dataset was provided by preliminary version of the baseline method of the ISBI 2014 challenge [14]. This dataset consists of 18 gray-scale cervical cytology images, each with 2 to 5 cells of different degrees of overlap, with 60 cells in total. The second dataset is the ISBI 2014 challenge dataset [26], consisting of 135 synthetic cervical cytology images (i.e., 45 training images and 90 test images, with 810 cells in total), where the number of overlapping cells is varied from 2 to 10 and the overlap coefficient between pairs of cells is in one of the following ranges: [0, 0.1], [0.1, 0.2],[0.2, 0.3], [0.3, 0.4], [0.4, 0.5].

3.2. Evaluation Metrics

We conducted quantitative and qualitative evaluations of the proposed method for nuclei and cytoplasm segmentation. In this study, the training dataset from ISBI challenge was used to train the CNN classifier and generate the shape dictionary, whereas the test datasets of the baseline method [14] and the ISBI challenge were used to evaluate our method performance. The proposed method was compared with the preliminary version of the baseline method [14] on the first dataset, and compared with the results of the ISBI challenge winners: Ushizima et al. [27], and Nosrati et al [28] and their newly proposed method [30], and the later version of the baseline method [29] on the ISBI challenge dataset. The nuclei segmentation results were also compared with the results of our preliminary version [31] on both datasets.

1) Nuclei segmentation. To assess the nuclei segmentation results, we used the criteria developed by Gençtav et al. [12] as defined in Eq. (10). The ground truth regions (R_{GT}) were used to categorize all segmented regions (R_{Seg}) into true positive detection TP_R (correctly classified as nucleus), false positive detection FP_R (classified as nucleus, but is in fact cytoplasm or background), or false negative detection (classified as non-nucleus, but is in fact nucleus). For each true detection instance, the numbers of true positive pixels (TP_p), false

positive (FP_p), and false negative (FN_p) pixels were counted.

$$\frac{R_{Seg} \cap R_{GT}}{R_{Seg}} > 0.6 \quad \text{and} \quad \frac{R_{Seg} \cap R_{GT}}{R_{GT}} > 0.6 \tag{10}$$

Then, the object-level precision ($prec_o$) and recall (rec_o), and pixel-level precision ($prec_p$) and recall (rec_p) were computed as:

$$prec_o = \frac{\#\ \text{of correctly segmented regions}}{\#\ \text{of all segmented regions}} = \frac{\sum TP_R}{\sum R_{Seg}} \tag{11}$$

$$rec_o = \frac{\#\ \text{of correctly segmented regions}}{\#\ \text{of all objects in the ground truth}} = \frac{\sum TP_R}{\sum R_{GT}} \tag{12}$$

$$prec_p = \frac{\#\ \text{of correctly segmented pixels}}{\#\ \text{of all segmented pixels}} = \frac{TP_p}{TP_p + FP_p} \tag{13}$$

$$rec_p = \frac{\#\ \text{of correctly segmented pixels}}{\#\ \text{of all pixels in the ground truth}} = \frac{TP_p}{TP_p + FN_p} \tag{14}$$

2) Cytoplasm segmentation. Using the evaluation code provided by ISBI challenge [26], the performance of cytoplasm segmentation was computed over the 'good' cell segmentations [57], for which the cell segmentation has a Zijdenbos similarity index (ZSI) above a threshold of 0.7. ZSI was computed as:

$$ZSI = 2\ \frac{|R_{GT} \cap R_{Seg}|}{|R_{GT}| + |R_{Seg}|} \tag{15}$$

where R_{GT} and R_{Seg} denote the ground truth and segmented regions, respectively, and $|.|$ denotes the number of pixels in the region. In addition, pixel-based true positive rate (TP_p) and false positive rate (FP_p) were also computed. The other segmented cells having a ZSI below the threshold were reported as false negatives, and used to compute the object-level false negative rate (FN_o).

4. Experimental Results

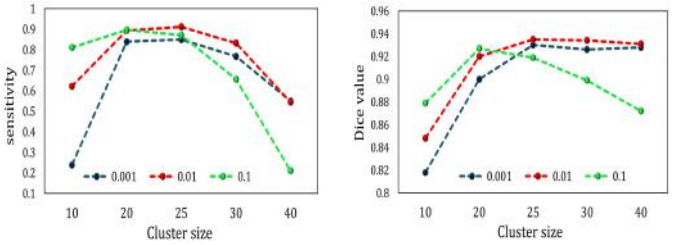
4.1. Quantitative evaluation

1) Nuclear segmentation: The nuclear segmentation of our proposed method is assessed in terms of object-level and pixel-level, and the results are shown in Table 1 and Table 2, respectively. Our object-level segmentation was the best among the other methods on the two datasets. The proposed method yielded a high improvement over [14] on the first dataset, with $prec_o$ of 0.98 and 29% improvement. Our $prec_o$ is also higher than the $prec_o$ obtained by our preliminary version [31]. Moreover, the object-level recall was 0.96 with 6% improvement compared with 0.90 obtained by [14]. Our precision and recall on the second dataset are 0.99 and 0.91 (i.e., with on average improvement 4% and 2%, respectively) over [27, 28, 29, 31]. Our method missed only two true nuclei out of 60 nuclei in the first dataset, and 48 nuclei out of 570 in the second dataset.

Furthermore, Table 2 shows that our method had the highest pixel-based recall rec_p and ZSI values on both datasets. Also, our pixel-based precision on the first dataset, 0.97 ($\pm$0.04) is similar to the $prec_p$ obtained by [14]. For the second dataset, Ushizima et al. [27] had the highest precision of 0.97 ($\pm$0.05), but it however had the lowest recall. The proposed method had a 8% better recall over Ushizima et al.'s method [27] and also a 3% better ZSI. A high recall, indicating high true positive rate, is more important than high precision since each nucleus represents a cell in the next stage. The pixel-based segmentation performance of our previous work and this work is rather similar due to using the clustering technique in classification, thereby, the whole superpixel, with the same number of pixels in both versions, is either correctly classified or misclassified. As mentioned earlier in Section 2.1.1, the SLIC parameters affect the performance of nuclei detection. To explain this, Figure 7 shows the influence of different SLIC parameters on the object-level and pixel-level nuclei segmentation results (without nuclei contour refinement) in term of sensitivity and Dice values. As shown, the selected parameters in our experiments provide

the best nuclei segmentation performance.

455 The object-level and pixel-level segmentation results proved the effectiveness of the cellular component separation stage of our proposed method. Tessellating the Pap smear image into small superpixels classified by deep learning model has the ability to successfully differentiate the nuclei regions. Deep convolutional neural network showed to be a reliable method to feature-learn and classify the 460 single-category patches in an accurate and fully automatic manner.



(a) (b)

Figure 7: The influence of different SLIC parameters on the (a) object-level and (b) pixel-level nuclei segmentation in term of sensitivity and Dice values.

Table 1: Quantitative object-level nuclei segmentation results.

Methods	$prec_o$	rec_o
Test dataset of [14]		
Baseline[a] [14]	0.69	0.90
preliminary version [31]	0.97	0.96
Our method	0.98	0.96
ISBI test dataset		
Ushizima [27]	0.959	0.895
Nostrati [28]	0.903	0.893
Baseline[b] [29]	0.977	0.883
preliminary version [31]	0.983	0.909
Our method	0.994	0.911

Table 2: Quantitative pixel-level nuclei segmentation results (using the mean and standard deviation result of each measure).

Methods	$prec_p$	rec_p	ZSI
Test dataset of [14]			
Baseline[a] [14]	0.97(±0.04)	0.88(±0.08)	0.92(±0.04)
preliminary version [31]	0.98(±0.03)	0.90(±0.08)	0.93(±0.04)
Our method	0.97(± 0.04)	0.91(± 0.08)	0.93(± 0.03)
ISBI test dataset			
Ushizima [27]	0.97(±0.05)	0.87(±0.07)	0.91(±0.04)
Nostrati [28]	0.90(±0.10)	0.92(±0.09)	0.90(±0.05)
Baseline[b] [29]	0.94(±0.08)	0.91(±0.08)	0.92(±0.05)
preliminary version [31]	0.93(±0.06)	0.95(±0.06)	0.94(±0.04)
Our method	0.94(±0.06)	0.95(±0.06)	0.94(±0.04)

2) Cytoplasmic segmentation: Table 3 shows a comparison of the cytoplasm segmentation performance with [14] on the first dataset, and with [27, 28, 30, 29] on the second dataset. The table showed that our proposed method achieved the highest ZSI and TP_p on the both datasets. The obtained ZSI on the first dataset was 0.93, compared with 0.88 obtained by [14]. For the second dataset, the obtained ZSI and TP_p were 0.90 and 0.95, with 1-3% and 1-12% improvement, respectively. In addition, the FN_O obtained by the proposed method on the first dataset is zero, which indicates that our method successfully segmented all cells in this dataset. Our optimal object-level true positive detection TP_o of 1.00 led to an increase of the FP_p value (i.e., 0.004) over that obtained by [14], which was 0.002. However, this FP_p is still small and has minimal impact on the reliability of our method. The FN_o value obtained on the second dataset is also better than [29]. These high ZSI and TP_p values on the two datasets demonstrated the capability of our method to accurately segment the cytoplasm from highly overlapping cells in different cervical image datasets.

Table 4 shows the cytoplasm segmentation performance of the proposed

Table 3: Quantitative results of the cytoplasm segmentation (using the mean and standard deviation result of each measure).

Methods	ZSI	TP_p	FP_p	FN_o
Test dataset of [14]				
Baseline[a] [14]	0.88(±0.08)	0.92(±0.10)	0.002(±0.005)	0.21(±0.29)
Our method	0.93(±0.06)	0.92(±0.09)	0.004(±0.005)	0.00(±0.00)
ISBI test dataset				
Ushizima [27]	0.87(±0.08)	0.83(±0.13)	0.001(±0.002)	0.17(±0.21)
Nosrati [28]	0.87(±0.08)	0.90(±0.09)	0.005(±0.004)	0.14(±0.17)
Nosrati [30]	0.88(±0.08)	0.93(±0.09)	0.005(±0.004)	0.11(±0.17)
Baseline[b] [29]	0.89(±0.08)	0.91(±0.10)	0.003(±0.005)	0.32(±0.29)
Our method	0.90(±0.08)	0.95(±0.07)	0.005(±0.004)	0.21(±0.24)

method over the range of ZSI thresholds $\{0.5, 0.6, 0.8\}$, comparing with the results of [14] on the first dataset and [28] on the second dataset. The table shows that our method achieved better ZSI and TP_p than the other methods at different ZSI thresholds. The proposed method had a stable improvement in performance on the first dataset, where the minimum obtained ZSI for the segmented cells was 0.71. In addition, a consistent improvement in performance was showed on the second dataset, with high ZSI and TP_p rates for all 'good' segmented cells. The best ZSI obtained by our method was 0.991 on the first dataset and 0.997 on the second dataset, which is close to the ideal segmentation.

Furthermore, the segmentation performance of the proposed method, in terms of ZSI and FNo, was assessed with respect to varying number of cells and overlapping degree, and the results were displayed in Figure 8. It is found from Figure 8 that the performance of the proposed method was promising with high ZSI values, i.e., above 0.90, when the number of cells in a cellular mass was three or less. Likewise, the segmentation of the cellular mass of four to six cells with overlapping degree less than 0.4 was shown to be good. However, when the overlapping ratio became more than 0.4 or the number of cells exceeded seven

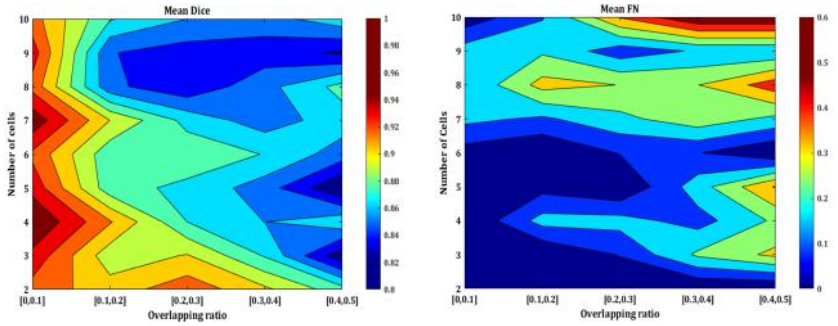
Table 4: Quantitative results of the cytoplasm segmentation at different ZSI thresholds.

Methods	ZSI >**0.5**	ZSI >**0.6**	ZSI >**0.8**
Test dataset of [14]			
Baseline[a] [14]	ZSI=.83,TP_p=.88, FP_p=.003,FN_o=0.02	ZSI=.85,TP_p=.89, FP_p=.002,FN_o=0.09	ZSI=.91,TP_p=.93, FP_p=.001,FN_o=0.34
Our method	ZSI=.93,TP_p=.92, FP_p=.004,FN_o=.00	ZSI=.93,TP_p=.92, FP_p=.004,FN_o=.00	ZSI=.93,TP_p=.93, FP_p=.004,FN_o=.02
ISBI test dataset			
Nosrati [28]	ZSI=.87,TP_p=.87, FP_p=.003,FN_o=.01	ZSI=.86,TP_p=.87, FP_p=.003,FN_o=.02	ZSI=.90,TP_p=.88, FP_p=.002,FN_o=.24
Our method	ZSI=.88,TP_p=.94, FP_p=.006,FN_o=.15	ZSI=.89,TP_p=.94, FP_p=.005,FN_o=.17	ZSI=.92,TP_p=.95, FP_p=.003,FN_o=.33

cells in the cellular mass, the ZSI was reduced to less than 0.85. Once the number of cells in the mass exceeded eight cells, the proposed method became not sensitive to the changes in the number of cells, but still sensitive to the change in the overlapping degree. Moreover, the FN_o of the proposed method was less than 0.1 as long as the overlapping ratio was less than 0.2 and the number of cells is less than seven. However, the FN_o was significantly increased when the number of cells is increased and the overlapping ratio is higher than 0.3. This is due to the insufficient number of isolated boundaries to predict the cell shape based on the training templates. In our proposed method, the number of cells, less than nine, did not have a great influence on the FN_o values, as the success of our deformation process depends mainly on the presence of the shape of the segmented cell in the shape dictionary. According to our observation, the elliptical shape cells in the dataset had more accurate segmentation than irregular shape cells regardless the degree of overlap (See Figure 9 in Section 5.2).

4.2. Qualitative evaluation

The qualitative experiment consists of a visual inspection of the nuclei and cytoplasm segmentation results. Some of our segmentation results with different



(a) (b)

Figure 8: Functional analysis of the proposed method in terms of Dice (in range [0:8; 1]) and FNo (in range [0; 0:6]). The diagrams show Dice and FNo as a function of number of cells (y axis) and cell overlap (x axis).

accuracy level, comparing with the results of Voronoi segmentation used in [27], and the ISBI baseline method [29], were displayed in Figure 9. It can be seen that the proposed method provided a precise estimation of the nuclei contour and the cytoplasmic mass contour, obtained in the first segmentation stage. The proposed method also succeeded in providing a good estimation for the cytoplasm boundaries inside the overlapping regions.

The visual comparison with the segmentation results of Voronoi diagram (e.g., [27]) and the baseline method[29], demonstrated the superiority of our approach in extracting the boundaries of overlapping cells. As shown in the figure, the Voronoi diagram yielded straight lines splitting the cytoplasm between pairs of cells, which obviously is not realistic. The proposed method could efficiently correct this limitation, and constructed more realistic cell shape. Our segmentation results were also better than the ISBI baseline results, which demonstrating that incorporating the proposed dynamically generated shape prior in a variational framework exhibited better performance than elliptical shape prior used by the baseline method.

Overall, the quantitative and qualitative results demonstrate that our proposed learned shape prior-based variational method succeeded in improving the

object-level and pixel- level segmentation performance of individual cytoplasm, over the elliptical shape prior [14, 28, 29] and the star shape prior [30], hence it was dynamically constructed based on the most representative shape templates from similar cells in the training dataset.

4.3. Computational complexity

Table 5 shows the computational time for the proposed method and the compared approaches. Our method was implemented in Matlab with non-optimized code running on a PC with Intel Core i5 3.2 GHz and 8 GB RAM. The average computational time of our proposed method was ∼30 seconds per image, with around 8 seconds for the cellular component classification, and 22 seconds for the individual cell segmentation. The proposed approach was 33 times faster than the baseline method [14] whose average computational time was ∼1000 seconds per image. The proposed method and the baseline method needed longer time than [27, 28]. This was mainly due to the classification step and more iterations of level set evolution, which, on the other hand, led to better performance than other methods.

Our cellular component separation stage took 8.81(±0.29) seconds to categorize all patches of a single image. In general, neutral networks are computationally intensive, because of iterative updating of a large number of parameters several times to minimize error and produce a precise model. However, CNNs are inherently parallel algorithms, hence, Graphics Processing Units (GPUs) can be used to dramatically reduce computation time needed for training. To speed up our method, we employed the CNN implementation of the MATLAB Computer Vision System Toolbox^{TM}, which provides a high-performance GPU-based CNN model.

The sparse-shape segmentation stage needed 22.18(±11.06) seconds to segment all cells of a single image. For each nucleus in a single image, there are two loops; the outer sparse-shape segmentation loop including the re-initialization, shape learning, and level set evolution, as shown in Figure 2 (4), and the inner iterations for prior-based level set evolution. Lets represent the number of nuclei

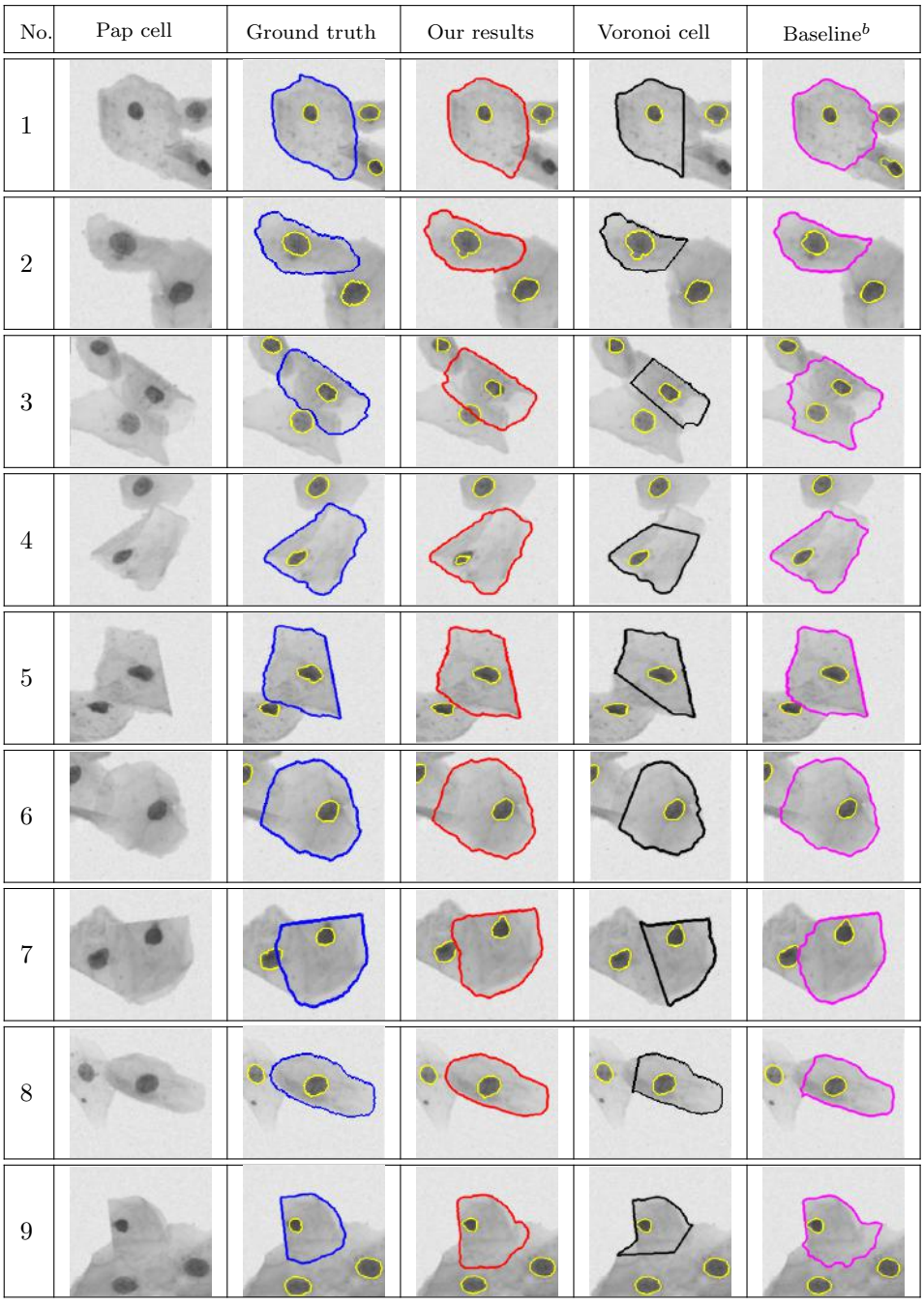


Figure 9: Qualitative segmentation results of overlapping cervical cells with different shapes. From left to right: Original cervical cell image, the ground truth cell contours, our segmentation, the Voronoi segmentation used by [27], and baseline[b] [29] segmentation respectively.

in the image as M, the number of outer iterations by N, and the inner iterations by K. Then, the complexity of the second stage of our approach can be written in term of Big O notation as $O(MNK)$. The complexity of our classification stage is $O(C)$, given that C is the number of superpixels in each image.

Table 5: Time complexity of the proposed and ISBI methods.

Methods	**T**ime/Image	**C**omputer specification
Ushizima [27]	12 sec.	Cray XC30 supercomputer, 12-core Intel, CPU 2.4GHz, 64GB RAM
Nosrati [28]	16.7 sec.	PC, CPU 3.40 GHz,16 GB RAM
Baseline [14]	$\sim$ 1000 sec.	PC, CPU 2.7 GHz, 40 GB RAM
Our method	$\sim$ 30 sec.	PC, CPU 3.20 GHz, 8 GB RAM

4.4. Failure cases

In spite of the superior segmentation results shown in Figure 9, our method still has some failure cases, such as those displayed in Figure 10. The segmentation results of the baseline method [29] have also been displayed in Figure 10 to show that these cases are difficult for other segmentation methods as well. The main failure segmentation cases obtained by our approach were caused by one of three reasons; (1) the initial segmentation obtained by Voronoi segmentation is not accurate (e.g., cell 1), (2) there is a lack of cell shape templates in the training shape dictionary (e.g., cell 2), and (3) the isolated boundary of the cell is not sufficient to guide the shape deformation by the training shapes (e.g., cell 3). Therefore, the segmentation results of our method can be further enhanced by establishing more reliable rough segmentation, and including more cell shapes in the training templates. To tackle the third issue, edge detection can be used to extract some robust boundaries from different directions to be used along with the isolated boundaries in order to give a good indication of the actual cell shape. Referring to cell 4 in Figure 10, if we are able to get a part of the boundary from the flat side, the cell shape can be easily estimated. These observations also show that our proposed method can get an optimizing

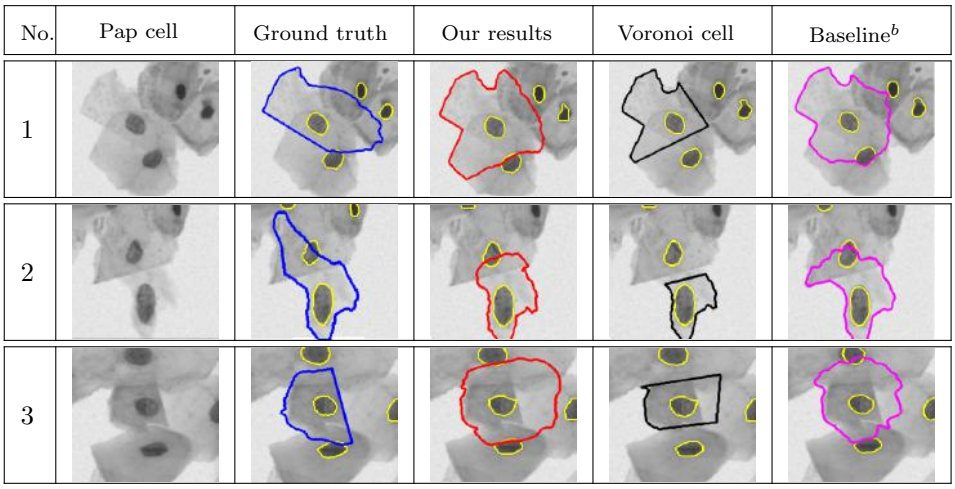


Figure 10: Failure segmentation because of the bad Voronoi segmentation (cell 1), the odd cell shape (cell 2), or Non-existence of enough isolated boundary (cell 3). From left to right: Original cervical cell image, the ground truth cell contours, our segmentation, the Voronoi segmentation used by [27], and baseline[b] [29] segmentation respectively.

performance if it is applied on cells with a typical shape with small variations, such as red blood cells which have semi-spherical shape.

5. Conclusions and Future Work

Cervical cell segmentation is a prerequisite to analyze cell-by-cell information toward optimizing the cervix cytological examination. One of the major challenges here is how to segment overlapping cells, which are typically presented in Pap smear. This paper addresses this issue and introduces a learning shape-driven variational method to provide an accurate localization and delineation of both nuclei and cytoplasm of highly overlapping cells. The proposed method employed the convolutional neural network with distinct features automatically learned from image patches to separate the nuclei, isolated cytoplasm, and cellular patches of overlapping cells. In the next stage, a dynamic shape prior generated from shape templates, and incorporated with regularized level set evolution is used to delineate the contour of overlapping cells. The segmen-

tation results of the proposed method were analyzed with both quantitative and qualitative evaluations, showing better results compared to the state-of-the-art methods using two cervical cell databases with a total of 870 cells. The experimental results indicated that the patch-wise CNN-based nuclei detection is more reliable and accurate than other state-of-the-art methods, even in presence of poor quality and similarity among different cellular patches. Moreover, our dynamic shape-driven variational method could successfully separate the individual cytoplasm of a mass with highly overlapping cells. We expect our proposed method to generalize effectively to arbitrary microscopy image types with overlapping cells, such as blood cell images, and this will be our future work.

We believe that this approach is ready to support and give a great help for the pathologist. The proposed approach provides an accurate segmentation for the overlapping cells, which is a time-consuming and error-prone task to the pathologists when performed manually. Our future work will include the classification of segmented cells to normal or abnormal cells based on morphological features such as cytoplasm and nuclei size and shape, which is another important question in automated cervix cytological examination.

References

[1] W. H. Organization, et al., Who guidance note: comprehensive cervical cancer prevention and control: a healthier future for girls and women.

[2] Z. Lu, G. Carneiro, A. Bradley, D. Ushizima, M. S. Nosrati, A. Bianchi, C. Carneiro, G. Hamarneh, Evaluation of three algorithms for the segmentation of overlapping cervical cells, IEEE Journal of Biomedical and Health Informatics.

[3] M. Hu, X. Ping, Y. Ding, Automated cell nucleus segmentation using improved snake, in: International Conference on Image Processing (ICIP), Vol. 4, IEEE, 2004, pp. 2737–2740.

[4] C. Jung, C. Kim, Segmenting clustered nuclei using H-minima transform-based marker extraction and contour parameterization, IEEE Transactions on Biomedical Engineering 57 (10) (2010) 2600–2604.

[5] C. Jung, C. Kim, S. W. Chae, S. Oh, Unsupervised segmentation of overlapped nuclei using bayesian classification, IEEE Transactions on Biomedical Engineering 57 (12) (2010) 2825–2832.

[6] C. Bergmeir, M. García Silvente, J. M. Benítez, Segmentation of cervical cell nuclei in high-resolution microscopic images: A new algorithm and a web-based software framework, Computer Methods and Programs in Biomedicine 107 (3) (2012) 497–512.

[7] M. E. Plissiti, C. Nikou, Overlapping cell nuclei segmentation using a spatially adaptive active physical model, IEEE Transactions on Image Processing 21 (11) (2012) 4568–4580.

[8] Y. Song, L. Zhang, S. Chen, D. Ni, B. Lei, T. Wang, Accurate segmentation of cervical cytoplasm and nuclei based on multiscale convolutional network and graph partitioning, IEEE Transactions on Biomedical Engineering 62 (10) (2015) 2421–2433.

[9] M. E. Plissiti, C. Nikou, A. Charchanti, Combining shape, texture and intensity features for cell nuclei extraction in Pap smear images, Pattern Recognition Letters 32 (6) (2011) 838–853.

[10] K. Li, Z. Lu, W. Liu, J. Yin, Cytoplasm and nucleus segmentation in cervical smear images using radiating GVF snake, Pattern Recognition 45 (4) (2012) 1255–1264.

[11] T. Chankong, N. Theera-Umpon, S. Auephanwiriyakul, Automatic cervical cell segmentation and classification in pap smears, Computer Methods and Programs in Biomedicine 113 (2) (2014) 539–556.

[12] A. Gentav, S. Aksoy, S. nder, Unsupervised segmentation and classification of cervical cell images, Pattern Recognition 45 (12) (2012) 4151–4168.

[13] A. Kale, S. Aksoy, Segmentation of cervical cell images, in: International Conference on Pattern Recognition (ICPR), IEEE, 2010, pp. 2399–2402.

[14] Z. Lu, G. Carneiro, A. P. Bradley, Automated nucleus and cytoplasm segmentation of overlapping cervical cells, in: Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer, 2013, pp. 452–460.

[15] N. Béliz-Osorio, J. Crespo, M. García-Rojo, A. Muñoz, J. Azpiazu, Cytology imaging segmentation using the locally constrained watershed transform, in: Mathematical Morphology and Its Applications to Image and Signal Processing, Springer, 2011, pp. 429–438.

[16] H. A. Phoulady, M. Zhou, D. B. Goldgof, L. O. Hall, P. R. Mouton, Automatic quantification and classification of cervical cancer via adaptive nucleus shape modeling, in: Image Processing (ICIP), 2016 IEEE International Conference on, IEEE, 2016, pp. 2658–2662.

[17] H.-S. Wu, J. Gil, J. Barba, Optimal segmentation of cell images, in: IEE Proceedings on Vision, Image and Signal Processing, Vol. 145, IET, 1998, pp. 50–56.

[18] J. Fan, R. Wang, S. Li, C. Zhang, Automated cervical cell image segmentation using level set based active contour model, in: International Conference on Control Automation Robotics & Vision (ICARCV), IEEE, 2012, pp. 877–882.

[19] A. Tareef, Y. Song, W. Cai, D. Feng, M. Chen, Automated three-stage nucleus and cytoplasm segmentation of overlapping cells, in: International Conference on Control Automation Robotics & Vision (ICARCV), IEEE, 2014, pp. 865–870.

[20] N. M. Harandi, S. Sadri, N. A. Moghaddam, R. Amirfattahi, An automated method for segmentation of epithelial cervical cells in images of ThinPrep, Journal of Medical Systems 34 (6) (2010) 1043–1058.

[21] T. Guan, D. Zhou, Y. Liu, Accurate segmentation of partially overlapping cervical cells based on dynamic sparse contour searching and gvf snake model, IEEE Journal of Biomedical and Health Informatics 19 (4) (2015) 1494–1504.

[22] J. Zhang, Z. Hu, G. Han, X. He, Segmentation of overlapping cells in cervical smears based on spatial relationship and overlapping translucency light transmission model, Pattern Recognition.

[23] M. M. Fatima alias Niraimathi, V. Seenivasagam, Radial tracing method of cytoplasm segmentation in overlapped cervical cell images, IETE Journal of Research 61 (4) (2015) 402–410.

[24] A. Tareef, Y. Song, M.-Z. Lee, D. D. Feng, M. Chen, W. Cai, Morphological filtering and hierarchical deformation for partially overlapping cell segmentation, in: International Conference on Digital Image Computing Techniques and Applications (DICTA), IEEE, 2015, pp. 1–7.

[25] S. N. A. M. Kanafiah, Y. Jusman, N. A. M. Isa, Z. Mohamed, Radial-based cell formation algorithm for separation of overlapping cells in medical microscopic images, Procedia Computer Science 59 (2015) 123–132.

[26] Overlapping Cervical Cytology Image Segmentation Challenge ISBI 2014, `http://cs.adelaide.edu.au/~carneiro/isbi14_challenge/`.

[27] D. Ushizima, A. Bianch, C. Carneiro, Segmentation of subcellular compartiments combining superpixel representation with voronoi diagrams, in: ISBI Overlapping Cervical Cytology Image Segmentation Challenge, IEEE, 2014, pp. 1–2.

[28] M. Nosrati, G. Hamarneh, A variational approach for overlapping cell segmentation, in: ISBI Overlapping Cervical Cytology Image Segmentation Challenge, IEEE, 2014, pp. 1–2.

[29] Z. Lu, G. Carneiro, A. P. Bradley, An improved joint optimization of multiple level set functions for the segmentation of overlapping cervical cells, IEEE Transactions on Image Processing 24 (4) (2015) 1261–1272.

[30] M. Nosrati, G. Hamarneh, Segmentation of overlapping cervical cells: A variational method with star-shape prior, in: IEEE International Symposium on Biomedical Imaging (ISBI), IEEE, 2015.

[31] A. Tareef, Y. Song, W. Cai, H. Huang, Y. Wang, D. Feng, M. Chen, Learning shape-driven segmentation based on neural network and sparse reconstruction toward automated cell analysis of cervical smears, in: International Conference on Neural Information Processing, Springer, 2015, pp. 390–400.

[32] J. Wright, Y. Ma, J. Mairal, G. Sapiro, T. S. Huang, S. Yan, Sparse representation for computer vision and pattern recognition, Proceedings of the IEEE 98 (6) (2010) 1031–1044.

[33] W. Ou, X. You, D. Tao, P. Zhang, Y. Tang, Z. Zhu, Robust face recognition via occlusion dictionary learning, Pattern Recognition 47 (4) (2014) 1559–1572.

[34] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, Y. Ma, Robust face recognition via sparse representation, IEEE Transactions on Pattern Analysis and Machine Intelligence 31 (2) (2009) 210–227.

[35] X. Mei, H. Ling, Robust visual tracking and vehicle classification via sparse representation, IEEE Transactions on Pattern Analysis and Machine Intelligence 33 (11) (2011) 2259–2272.

[36] S. Zhang, H. Yao, H. Zhou, X. Sun, S. Liu, Robust visual tracking based on online learning sparse representation, Neurocomputing 100 (2013) 31–40.

[37] S. Zhang, Y. Zhan, M. Dewan, J. Huang, D. N. Metaxas, X. S. Zhou, Deformable segmentation via sparse shape representation, in: Medical Im-

age Computing and Computer-Assisted Intervention (MICCAI), Springer, 2011, pp. 451–458.

[38] C. Florin, N. Paragios, G. Funka-Lea, J. Williams, Liver segmentation using sparse 3D prior models with optimal data support, in: Information Processing in Medical Imaging, Springer, 2007, pp. 38–49.

[39] Y. Gao, S. Liao, D. Shen, Prostate segmentation by sparse representation based classification, Medical Physics 39 (10) (2012) 6372–6387.

[40] G. Gill, M. Toews, R. R. Beichel, Robust initialization of active shape models for lung segmentation in CT scans: A feature-based atlas approach, International Journal of Biomedical Imaging 2014.

[41] T. Tong, R. Wolz, P. Coupé, J. V. Hajnal, D. Rueckert, A. D. N. Initiative, et al., Segmentation of mr images via discriminative dictionary learning and sparse coding: Application to hippocampus labeling, NeuroImage 76 (2013) 11–23.

[42] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, S. Susstrunk, SLIC superpixels compared to state-of-the-art superpixel methods, IEEE Transactions on Pattern Analysis and Machine Intelligence 34 (11) (2012) 2274–2282.

[43] R. Keys, Cubic convolution interpolation for digital image processing, IEEE Transactions on Acoustics, Speech, and Signal Processing 29 (6) (1981) 1153–1160.

[44] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, in: Advances in Neural Information Processing Systems, 2012, pp. 1097–1105.

[45] D. Ciregan, U. Meier, J. Schmidhuber, Multi-column deep neural networks for image classification, in: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, IEEE, 2012, pp. 3642–3649.

[46] S. Lawrence, C. L. Giles, A. C. Tsoi, A. D. Back, Face recognition: A convolutional neural-network approach, IEEE Transactions on Neural Networks 8 (1) (1997) 98–113.

[47] F. Lauer, C. Y. Suen, G. Bloch, A trainable feature extractor for handwritten digit recognition, Pattern Recognition 40 (6) (2007) 1816–1824.

[48] V. Nair, G. E. Hinton, Rectified linear units improve restricted boltzmann machines, in: Proceedings of the International Conference on Machine Learning (ICML), 2010, pp. 807–814.

[49] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556.

[50] D. Scherer, A. Müller, S. Behnke, Evaluation of pooling operations in convolutional architectures for object recognition, in: International Conference on Artificial Neural Networks, Springer, 2010, pp. 92–101.

[51] C. Goodall, Procrustes methods in the statistical analysis of shape, Journal of the Royal Statistical Society. Series B (Methodological) (1991) 285–339.

[52] T. Fu, X.-t. Yin, Y. Zhang, Voronoi algorithm model and the realization of its program, Computer Simulation 23 (2006) 89–91.

[53] S. W. Smith, "Moving average filters", in The Scientist and Engineer's Guide to Digital Signal Processing, California Technical Pub. San Diego, 1997, Ch. 15, pp. 277–284.

[54] Z. Zhang, B. D. Rao, Sparse signal recovery with temporally correlated source vectors using sparse bayesian learning, IEEE Journal of Selected Topics in Signal Processing 5 (5) (2011) 912–926.

[55] M. Rousson, N. Paragios, Shape priors for level set representations, in: European Conference on Computer Vision (ECCV), Springer, 2002, pp. 78–92.

[56] C. Li, C. Xu, C. Gui, M. D. Fox, Distance regularized level set evolution and its application to image segmentation, IEEE Transactions on Image Processing 19 (12) (2010) 3243–3254.

[57] P. Radau, Y. Lu, K. Connelly, G. Paul, A. Dick, G. Wright, Evaluation framework for algorithms segmenting short axis cardiac MRI, The MIDAS Journal-Cardiac MR Left Ventricle Segmentation Challenge.