Assessment and Elimination of Inflammatory Cell: A Machine Learning Approach in Digital Cytology

Jing Ke
Computer Science and Engineering
Shanghai Jiao Tong University
Shanghai, China
kejing@sjtu.edu.cn

Junwei Deng
School of Information
University of Michigan
Ann Arbor, United States
junweid@umich.edu

Yizhou Lu
Computer Science and Engineering
Shanghai Jiao Tong University
Shanghai, China
yizhou_lu@sjtu.edu.cn

Dadong Wang
Quantitative Imaging Research Team
Data61, CSIRO
Sydney, Australia
dadong.wang@csiro.au

Yang Song
Computer Science and Engineering
University of New South Wales
Sydney, Australia
yang.song1@unsw.edu.au

Huijuan Zhang
Department of Pathology
Peace Maternity and Child Health Hospital, SJTU
Shanghai, China
zhanghj815@sjtu.edu.cn

***Abstract*—In automatic cytology image diagnosis, the false-positive or false-negative often come up with inflammatory cells that obscure the identification of abnormal or normal cells. These phenotypes are presented in the similar appearance in shape, color and texture with cells to detect. In this paper, to evaluate the inflammation and eliminate their disturbances of recognizing cells of interests, we propose a two-stage framework containing a deep learning based neural network to detect and estimate the proportions of inflammatory cells, and a morphology based image processing architecture to eliminate them from the digital images with image inpainting. For performance evaluation, we apply the framework to our collected real-life clinical cytology slides presented with a variety of complexities. We evaluate the tests on sub-images cropped from 49 positive and 49 negative slides from different patients, each at the magnification rate of 40×. The experiments shows an accurate profile of the coverage of inflammation in the whole slide images, as well as their proportion in all the cells presented in the image. Confirmed by cytotechnologists, more than 96.0% of inflammatory cells are successfully detected at pixel level and well-inpainted in the cytology images without bringing new recognition problem.**

I. Introduction

The automated detection of objects of interests, mainly referring to abnormalities and lesions in medical images, is a key part of diagnosis and efficiently saves the labor-intensive analyzing work for clinicians in large-sized whole slide images. There have been several publications in the literature to study the detection, segmentation and classification of cervical cells in cytology images. In these computer-aided methods, the detection of inflammatory cells is seldom investigated as inflammation is often considered not to be a severe problem. Inflammatory cells are not included in widely used public dataset like Herlev [1] or HEMLBC [2] dataset as a major subtype of cervical cells for classification study, where a list of seven subtypes named superficial, intermediate and columnar as normal cells, and mild, moderate, severe and carcinoma as abnormal cells are focused for cancer recognition. However, the research in the cervical cytology diagnosis suggest that cervical inflammation may be associated with high-grade lesions and may be a cofactor for high-grade cervical lesions in women infected with oncogenic HPV [3]. Typically, the inflammatory cells are extensively spread over the whole slides and may even cover up to 75% of other normal or abnormal cells, and brings another difficulty in automatic diagnosis that they are prone to be considered as false-positive or false-negative to obscure the identification of abnormal or normal cells, and the only existing and robust way to eliminate them from the image is to filter them with chemical methods.

In this paper, we propose a framework to evaluate the inflammation of patient from cytology images and eliminate their disturbances of recognizing a phenotype in the further diagnosis. This work makes two major contributions to the field: 1) We propose an end-to-end convolutional neural network (CNN) architecture to jointly detect and segment inflammatory cells and other phenotypes. We evaluate the coverage of all the inflammatory cells in the whole slide image. Due to the variance in cell density in liquid-based cervical (LBC) slides, we also estimate their proportion of all the cells in the image excluding background area. 2) We remove them naturally from the images that they will not be wrongly catalogued to a certain phenotype of interest, therefore the false-positive or false-negatives are significantly reduced in the further cancer diagnosis. Based on the extraction of morphological features, we perform pixel-wise level corrections to the results from the first stage by involving more aggregative inflammatory

978-1-7281-3857-2/19/$31.00 ©2019 IEEE

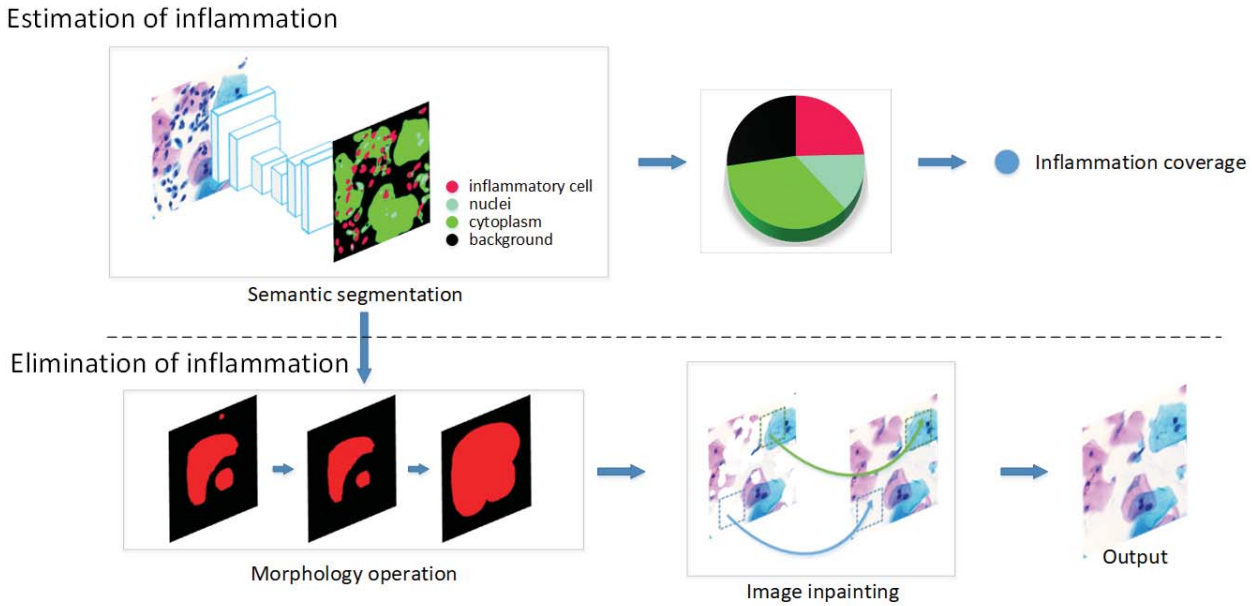

Fig. 1. Workflow of the proposed model in CNN-based detection and conditional enhancement, morphological correction and image impainting for inflammatory cells.

cells falsely identified and segmented to other subtypes, and vice versa exclude wrongly predicted pixels. To fill the vacancies left by the removal of inflammatory cells and provide natural and reasonable imagery, as well as to avoid further potential recognition problem by using single-color mask, we design a biased PatchMatch algorithm for a predictable image-inpainting. To the best of our knowledge, this is the first work that provides a framework to robustly estimate and eliminate inflammatory cells from complex digital scenarios based on real-life clinical dataset. Generally, such suppression of noise or false-positives is often achieved by chemical methods with higher cost of labor and time.

The rest of the paper is organized as follows. In section 2, we discuss the challenges in detection and classification tasks in the medical image analysis. In section 3, we introduce the methodologies applied for inflammation assessment and the approaches to eliminate inflammatory cells for normal/abnormal cell identification. In section 4, we present the final experimental results performed on the clinical cytology images with very complex scenarios. In section 5, we discuss our current and future work of cancer diagnosis based on artificial intelligence.

II. Related Work

Before the breakthrough of neural networks in the 1990s, medical image processing and analysis are mainly performed by handcrafted features extraction for about 20 years. When supervised and unsupervised techniques are applied for medical analysis, it is an almost standard procedure that the systems are trained by annotated example data for feature vector extraction and the trainings are performed by computers. Recently in computer vision, deep convolutional networks have become the top choice of techniques for their remarkable achievement in classification accuracy, and now are also very broadly used in lesion detection and classification in medical image analysis system. Among the recently deep learning based models, we see some outstanding novel architectures stood out from ImageNet Large Scale Visual Recognition Challenge, like AlexNet [4], GoogleNet/Inception [5], VGG-Net [6], ResNet [7] and models afterward with deeper layers and more complex architectures. However, with simple architecture, VGG-16 and ResNet are still very popular in medical analysis with transfer learning [8] or learning from scratch. However, we still employ handcrafted features in biomedical image analysis mainly due to two challenges. Firstly, high accuracy is often attributed to the large annotated dataset while in biomedical domain, the challenge is not only the availability of image data itself, but also the acquisition of relevant annotations for these images from clinicians. Secondly, the boundaries of cells or tissues cannot be as clear as those from natural objects hence researchers' knowledge over the target objects will contribute to the classification and segmentation results. Recent searches also prove that, experimentally, when high sensitivity and specificity are demanded simultaneously, the extraction of discriminant features in specific cell or tissue can be better acquired with hybrid learning approaches [9], [10].

In the past few years we see many successful applications of deep learning models used to assist diagnosis in medical domain [11]–[13]. In cervical cell segmentation or classification [10], [14]–[16], some approaches have been proposed based on deep learning architectures, some are incorporated with handcrafted features like Laplacian of Gaussian (LoG) filter

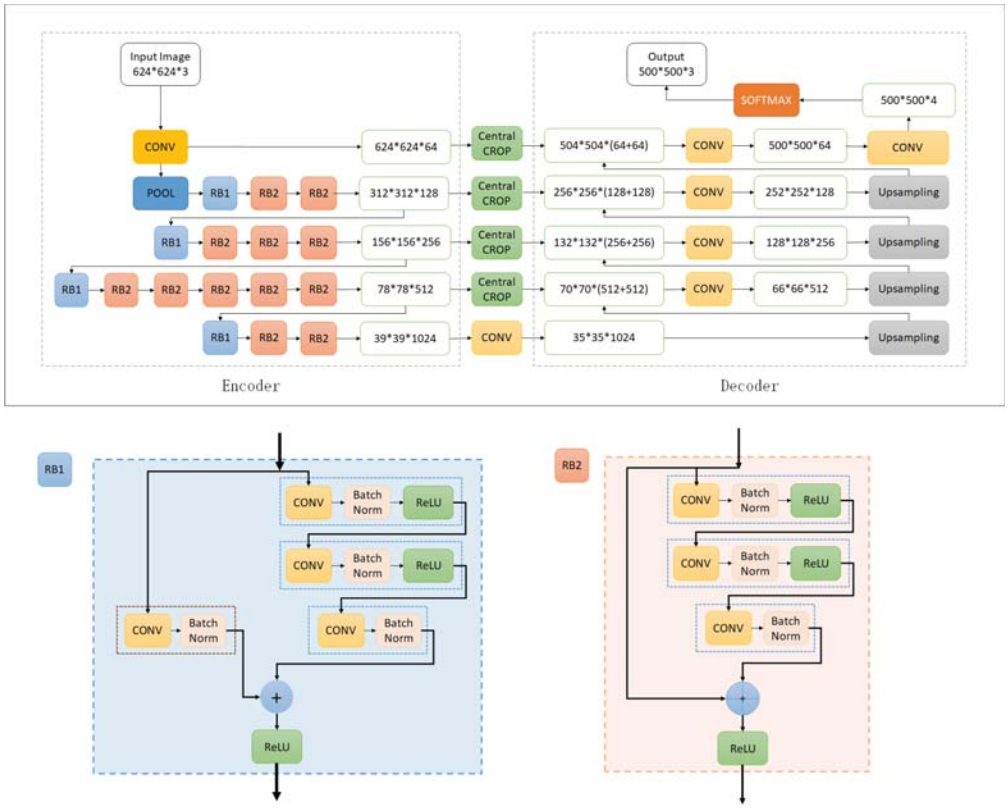

Fig. 2. A hybrid architecture of ResNet blocks as encoder and U-Net ascending part as decoder.

[17] and graph partitioning [18] for overlapping cells. Some Although great efforts have been invested into each branch in image recognition and processing, there are few researches into the assessment of inflammation in cytology images, nor methods to eliminate their influence the the identification of abnormal/normal cells.

III. Methodology

The whole framework contains two main functional parts. Firstly, a hybrid ResNet and U-Net architecture is proposed for joint detection and segmentation of nuclei and inflammatory cells, and the proportion of inflammation is estimated. Secondly, morphological operations are applied to correct the boundaries of inflammatory cells are a biased image-inpainting algorithm is designed for a robust removal of these cells in the cytology image. In this way, we significantly reduced the false negatives in normal cell detection, as well as false positives in abnormal cell detection in the further diagnosis.

A. Assessment of inflammation in Cytology Images

Fully convolutional networks (FCNs) efficiently make dense predictions for per-pixel tasks from supervised pre-training [19]. U-Net [11] extends the convolutional autoencoder and combines skip connections with the feature maps located before the bottleneck layer. To inherit the advantages, at the first stage, we apply residual blocks in the encoder path for feature extraction and object location, where the original high-dimension is halved and channels are doubled from the code. We retain the decoder path of U-Net to recover the data from the code combining high-resolution features from the encoding path with upsampling, and output prediction for classification and segmentation. We predict four catalogues of inflammatory cells, nuclei, cytoplasm and background with pixel-level labeled segmentation map. The encoder and decoder are demonstrated in Fig.2.

We do not estimate any boundary to separate overlapped cells as inflammatory are often aggregated and expected to be removed all together in the processing. On the contrary, the boundaries of inflammatory cells are also adaptively expanded by our loss function, and a higher confidence is assigned to inflammatory cells when background or cytoplasm are valued with the highest probability. The confidence rate will not be changed for the binary classification between inflammatory cells and nuclei hence the abnormal cells are preserved in the further cancer detection task. The loss function is described as:

$$l(\theta) = LW^T \cdot (-LX^T + \log(\sum_{j=1}^{n} \exp(X_j)) \qquad (1)$$

where $L = (l_1, l_2, ..., l_n)$ is the one-hot-encoded label for one pixel in the order of (background, inflammatory cell, cell nucleus, cytoplasm) $W = (w_1, w_2, ..., w_n)$ is the weight for each

class. $X = (x_1, x_2, ..., x_n)$ is the output of the network before the softmax layer. In the detection and segmentation task, n is set to 4 and $W = (1, 2, 2, 1)$ for expected segmentation results.

The tests are performed on real-life clinical data we collected with a variety of complexities of 49 positive and 49 negative liquid-based cytology (LBC) images at the magnification rate of 40x and each around 56,000×56,000 pixels. We pixel-wisely annotate a very small sample of only ten 2000×2000 sub-images for training. The tested whole slides are cropped to patches of 2000×2000 and 500×500, among which 100 slides are selected and balanced in subtypes. Confirmed by cytotechnologists, about 92.0% to 97.5% inflammatory cells are detected and segmented by our proposed method at pixel level, and the fluctuation of accuracy is up to the complexity of the scenarios from different samples.

B. Elimination of Inflammatory Cells from Digital Images

Mathematical morphological operators are designed for the extraction of image objects or structures as the designers can use their knowledge about the shape and geometrical properties of the target image objects [20]. Based on a known shape called "structuring element" (SE) B, erosion and dilation are two basic operators used to erode away or oppositely gradually enlarge the boundaries of regions of foreground pixels.

As there are high similarities between nuclei and inflammatory cells as objects to detect, there are unavoidable false result of nested contour by any neural network. In this case, we use erosion operation to eliminate such false detection and segmentation results of inflammatory cells nested in nuclei, or in the cytoplasm or background. And oppositely, we use dilation operation to enlarge the area of inflammatory cells to aggregate more overlapped ones, and hence more blurring pixels around the boundaries that have been falsely segmented to other classes, are catalogued properly into the target group.

In this workflow, we perform a morphological opening, that is erosion followed by dilation, to the masked region generated by our CNN. Cytologically, cells are often presented with the shapes close to circle, so we use a circle-shaped SE of radius r. The original masked region after the prediction from neural network is expressed as f, for any pixel position $\mathbf{x}$, we apply erosion $\varepsilon_B(f)$:

$$[\varepsilon_B(f)](\mathbf{x}) = \min_{\mathbf{b}\in B} f(\mathbf{x}+\mathbf{b}) \tag{2}$$

followed by dilation $\delta_B(f)$ with a slightly enlarged r:

$$[\delta_B(f)](\mathbf{x}) = \max_{\mathbf{b}\in B} f(\mathbf{x}+\mathbf{b}) \tag{3}$$

to get opening $\gamma_B(f)$:

$$\gamma_B(f) = \delta_B[\varepsilon_B(f)] \tag{4}$$

We present a demonstration of the workflow in Fig.3, where we have (a) as the input image, (b) presents the segmentation map of probability, (c) is the predicted four-catalogue segmentation result from neural network, (d) is our target for mathematical morphological operation, (e) shows the erosion result, (f) is the dilation result based on erosion, and the shape of SE is also demonstrated.

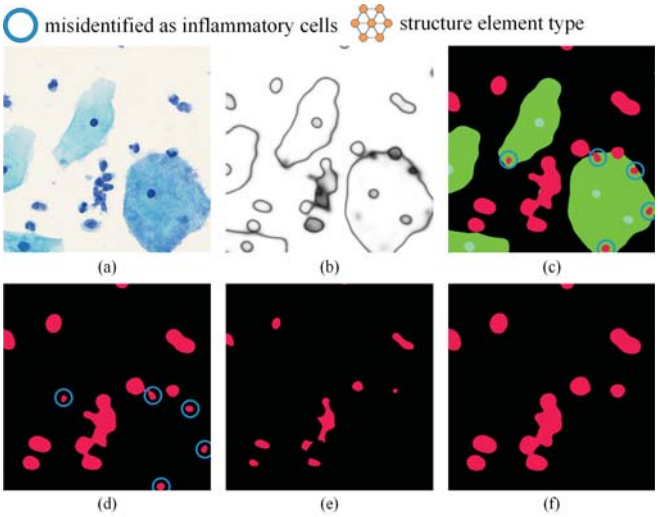

Fig. 3. Workflow of morphological operations, where (a) the input image, (b) the segmentation map of probability, (c) the predicted four-catalogue segmentation map from neural network with adaptive confidence adjustment, where the circled area are wrongly defined as target cells, (d) target cells (red) for morphological operations, (e) erosion result, (f) dilation result.

In order to fill up the holes naturally after noise removal, we apply image-inpainting methods for a reasonable image rather than filling it with single color to cause further recognition obstacles. Previously, deep learning approaches have focused on rectangular regions located around the center of the image, and often rely on expensive post-processing. The core PatchMatch algorithm [21] quickly finds correspondences between small square regions or patches of an image, naturally adapted to image inpainting problem and applicable to irregular masks. We apply this algorithm to fill the holes after the removal of unwanted content with two optimizations of a biased prediction and a faster convergence.

The inpainting process is based on building an image pyramid to obtain the better result. The reconstruction of the image is performed through each layer for a coarse-to-fine gradual resize processing. The algorithm is based on a bidirectional distance measure for completeness and coherence to minimize the distance function:

$$\mathrm{Dist}ance(src, dst) = \overbrace{\frac{1}{H*W}\sum_{s\subset src}\min_{d\subset dst} D(s,d)}^{d_{complete}(src,dst)} + \overbrace{\frac{1}{H*W}\sum_{d\subset dst}\min_{s\subset src} D(d,s)}^{d_{cohere}(src,dst)} \tag{5}$$

where src is the input image and dst is the output image, both sized $[H, W]$ as width and height of the image, and metric $D(a, b)$ evaluates the distance between the patches centered at a and b. The core PatchMatch algorithm contains multiple iterations of propagation and random search, each performed twice to minimize the $d_{complete}$ and d_{cohere} in one iteration.

The *completeness* term ensures that the output image extract as much visual information from the input as possible and the *coherence* term ensures that the output is coherent with the input where artifacts are penalized. Formally, the distance measure is defined simply as the sum of the average distance of all patches in src to their most similar patches in dst and vice versa [22].

As the input data of the algorithm, we have an original image src and its "holes" of $mask$, both sized $[H, W]$ as width and height of the image. The $downSample$ operation in the original PatchMatch algorithm gets an average pixel value of 2×2 patch, hence any boundary pixels of the mask region can be valued by a pixel from one of the layers in the pyramid. The $patchMatch$ is the edited core algorithm to fetch patches from image B to fill masked region in image A, and in this way we reach to a map of the best position by multiple iterations. The $weightVote$ decides the value of the pixels based on the weighted distances. We summarize the algorithm in Algorithm 1.

Algorithm 1 Image Inpainting

input: src, $mask$, H, W
output: dst
// Build image pyramid:
$l\_num \leftarrow \min(\log_2 H, \log_2 W)$
$p_0 \leftarrow src, mask$
for $i \leftarrow 1 : l\_num - 1$ **do**
  $p_i \leftarrow downSample(p_{i-1})$
// Inpaint every layer from coarse to fine:
$src, dst \leftarrow p_{l\_num-1}$
for $i \leftarrow l\_num - 1 : 0$ **do**
  // EM-like algorithm to bidirectionally inpaint:
  for $EM_{iter} \leftarrow 1 : 2i + 1$ **do**
    $o\_sd \leftarrow patchMatch(src, dst)$
    $o\_ds \leftarrow patchMatch(dst, src)$
    if $EM_{iter} == 2i + 1$ and $i\ != 0$ **then**
      $src \leftarrow p_{i-1}$
    $dst \leftarrow$ weightVote(o_sd, o_ds, src, dst)

However, as no architecture can provide 100% correct segmented results, there are unavoidable pixels from target objects falsely catalogued to other classes in the image. In this case, unwanted tissues might be inpainted by the original PatchMatch algorithm. As the false positives or noise reduction task cannot involve any nuclei or inflammatory cells as new content, we need to make a biased prediction to exclude such possibilities. Consequently, we optimize the algorithm to a biased PatchMatch algorithm to robustly exclude new noises in inpainting. To conditionally select suitable structure, we set a threshold to the pixels involved in downsampling processing to the next layer in the pyramid. As the materials we prefer to fill is either from background or cytoplasm, we set a color range of inflammatory cells and nuclei that the pixels in the restricted regions are not selected to the downsampling processing for pyramid building. The Equation 6 illustrates the downsample processing, where A is the image for size-decrease, K the sliding frame kernel of size $N \times N$, and $thr = (r_t, g_t, b_t)$ sets the threshold to exclude unwanted visual contents.

$$\text{Downsample}(i, j, A, K) = \frac{\sum_{k_x=1}^{N}\sum_{k_y=1}^{N} K_{k_x,k_y} * \max(A_{(i+k_x-\frac{N}{2})(j+k_y-\frac{N}{2})}, thr)}{\sum_{k_x=1}^{N}\sum_{k_y=1}^{N} K_{k_x,k_y}} \tag{6}$$

As a result, the boundary of the masks will not be filled with any unwanted visual structures in any scale hence it is ensured that the image reconstruction will contain only expected contents, and a typical comparison between original PatchMatch and this biased optimization is shown in Fig.4.

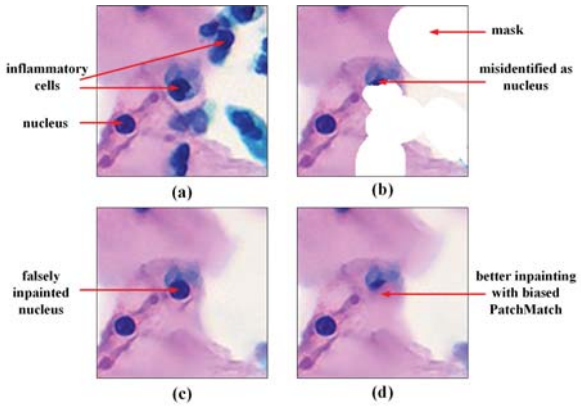

Fig. 4. Inpainting results with optimized algorithm. (a) the original image, (b) mask region, (c) a falsely inpainted result by the original PatchMatch algorithm(2009), and (d) result of our optimized biased PatchMatch algorithm.

The core PatchMatch algorithm contain three major stages of mask initialization, iterative propagation and random search, which is described in Algorithm 2, where A is the target image and B is the source image of size $[H, W]$. A_{mask} and B_{mask} are the masked region of A and B, w is the randomly search window and α is a ratio between search windows size. r is the current index of layer in pyramid. l is the result nearest neighbor map.

The original PatchMatch algorithm is used to reconstruct the masked region in image A by finding its nearest neighbour from image B. For the initialization of i, j in the masked region, the nearest neighbor of image A is randomly associated with an un-masked spot from image B and the un-masked region of image A are reserved without update. Then we bidirectionally scan all the masked regions and conditionally copy their nearest neighbor from $(i-1, j)$, $(i, j-1)$,$(i+1, j)$ or $(i, j+1)$. In the original PatchMatch algorithm, random search step is performed in the whole image area. However, due to the noise-reduction task of our work, we propose a constrained search to narrow down w within the neighboring cytoplasm or background. The random search window size is described as:

Algorithm 2 PatchMatch

input: $A, B, H, W, w, \alpha, A_{mask}, B_{mask}, r$
output: l

//Initialization
for $(i,j) \in [1,H] \times [1,W]$ **do**

$$l_{i,j} \leftarrow \begin{cases} (i,j) & \text{if } (i,j) \notin A_{mask} \\ (x,y) \in [1,H] * [1,W] \backslash B_{mask} & \text{if } (i,j) \in A_{mask} \end{cases}$$

//Iteration
for $NFF_{iter} \leftarrow 1 : 1 + r$ **do**
//Proporgation
for $(i,j) \in A_{mask}$ **do**

$$s_p \leftarrow \begin{cases} \{l_{i,j} \quad l_{i+1,j} \quad l_{i,j+1}\} & \text{if } NFF_{iter} \text{ is even} \\ \{l_{i,j} \quad l_{i-1,j} \quad l_{i,j-1}\} & \text{if } NFF_{iter} \text{ is odd} \end{cases}$$

$$l_{i,j} \leftarrow \underset{(x_p,y_p) \in s_p}{\arg\min}\ D((i,j),(x_p,y_p))$$

//Random Search
for $(i,j) \in A_{mask}$ **do**
for $k \leftarrow 1 : \log_\alpha w$ **do**

$$u_k \leftarrow l_{ij} + w\alpha^k R_k$$

$$l_{i,j} = \underset{(u_x,u_y) \in u}{\arg\min}\ D((i,j),(u_x,u_y))$$

$w = \frac{d_c}{\alpha^r}$, where d_c is the diameter of the largest typical cells in the image and we evaluate it to be 40 experimentally, α is a ratio between search windows size and we set it to 2 for the best accuracy, r is the index of layer in pyramid built. The random search list u is described in Equation 7, where l is the nearest neighboring map and R is a list of random number in [-1,1] × [-1,1]. The propagation and random search tasks are interleaved in each pyramid layer.

$$u_k = l_{i,j} + w\alpha^k R_k \tag{7}$$

The constrained searching area is described in Fig.5. We have the speedup rates tested on three 2000×2000 sub-image, with 15%, 30% and 45% masked region respectively, along with their downsampled results at multiple scales. The acceleration results are presented in Fig.6. It should also be noticed that the pattern and color around, as well as the number of boundary pixels of the masked region rate, are also important parameters to the acceleration rates.

IV. Results

To evaluate the performance of the proposed methods, we use our own collected private dataset for evaluation.

A. Dataset

The dataset is collected from Shanxi Tumor Hospital with ethics approval. All the sensitive information of patients has been eliminated. With only ten 2000×2000 sub-images annotated for detection and segmentation, we perform the tests on the real-life clinical data of 49 positive liquid-based cytology (LBC) microscope images we collected from hospitals, each at the magnification rate of 40× and around 56,000×56,000 pixels. Cropped to patches of 2000×2000 or 500×500, 100 sub-images are selected and balanced in subtypes for verification.

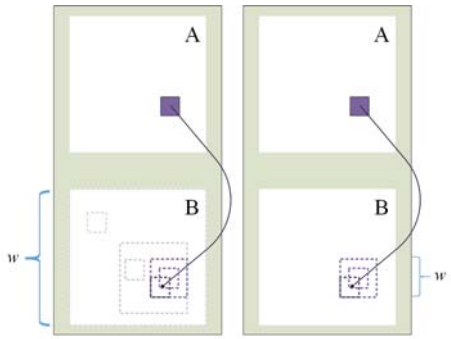

Fig. 5. Random search window size of original image area (left) and optimized narrowed area (right).

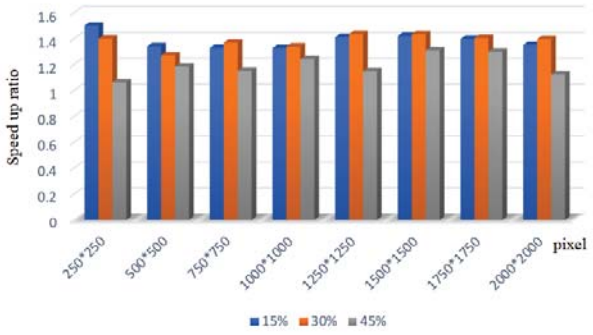

Fig. 6. Performance comparison of the optimized PatchMatch algorithm with original algorithm.

B. Estimation of Inflammation Coverage

We profile the inflammation in term of the area proportion. Evaluated on the 100 images including the original images in Fig.7, we summarize the result in Table I, where Inf is the percentage of inflammatory cell in the whole image including background, Cell is the percentage of all cells in the whole image, and Inf/Cell is the proportion of all normal and abnormal tissues or cells. Inf GT is the ground truth of inflammatory cell pixels in the image, Inf Detected is the pixels correctly detected by our method, Cell GT is the ground truth of pixels of all cells in the image and Accuracy is the percentage of inflammatory cell pixels correctly detected by our method. The proposed architecture shows a high accuracy in the successful detection of inflammatory cells experimentally.

C. Elimination of Inflammatory Cells

We show part of the processing results in Fig.7 and Fig.8. Their detailed performances and overall performance of 100 sub-images are described in Table II. GT refers to the ground truth of the total number of inflammatory cells in each sub-image from a to h and average value among all 100 sub-images, NN TP/FP represents the true/false positives detected by neural network, and HC TP/FP represents that

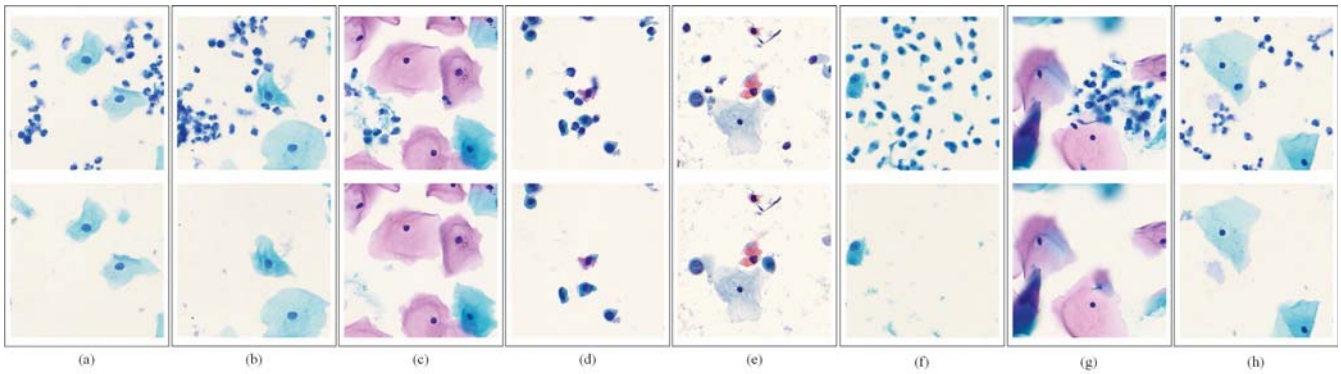

Fig. 7. A comparison before (top) and after (bottom) processing with our proposed model for detection, segmentation and image inpainting for inflammatory cells in cervical cell microscope image.

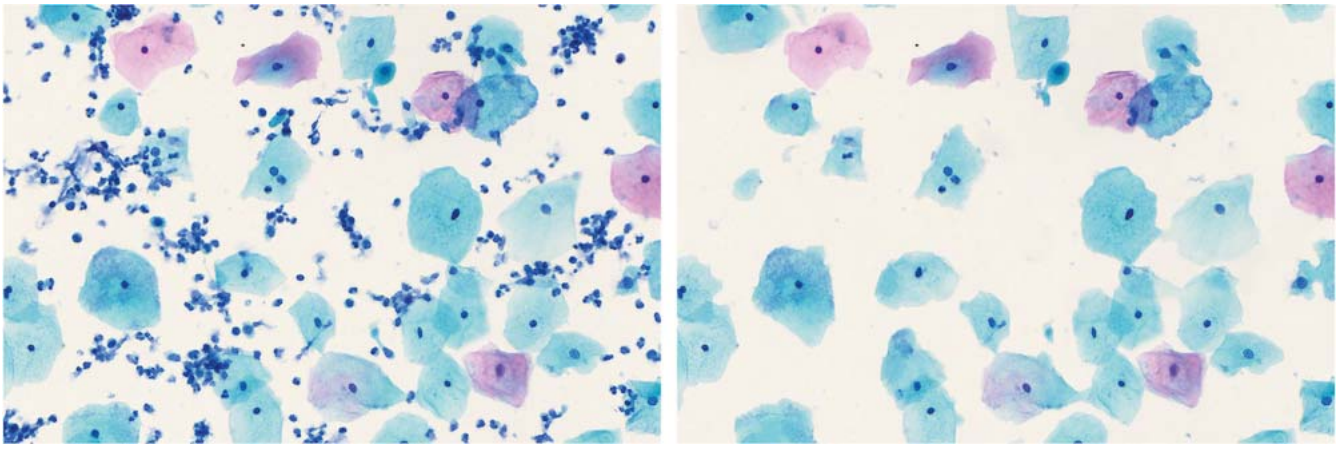

Fig. 8. A comparison before (left) and after (right) processing with our proposed model for detection, segmentation and image inpainting for inflammatory cells in cervical cell microscope image

Fig. 9. An example of unsuccessful detection and removal in our proposed method.

TABLE I
PERFORMANCE OF CELL DETECTION (PIXELS OUT OF 250000 PIXELS PER IMAGE)

ID	Inf	Cell	Inf/Cell	Inf GT	Inf Detected	Cell GT	Accuracy
a	10.32%	19.35%	53.3%	25803	25791	48372	99.9%
b	12.01%	26.95%	44.6%	30041	29970	67389	99.7%
c	2.69%	67.91%	3.9%	6731	6704	169774	99.5%
d	3.92%	8.13%	48.2%	9791	9791	20331	100%
e	1.57%	21.19%	7.4%	3901	3880	52970	99.4%
f	17.04%	20.44%	83.4%	42613	42588	51104	99.9%
g	8.32%	56.75%	14.7%	20783	20549	141874	98.8%
h	9.26%	38.84%	23.8%	23159	23083	97097	99.6%
100 images	8.18%	33.45%	24.5%	20457	20141	83654	98.4%

after handcrafted features applied. Here by true positives, we refer to the inflammatory cells instead of abnormal cell in cancer detection. Some low-quality image have clustering inflammatory cells overlapping on other cells. We noticed that some inflammatory cells are left partially on the image in some cases, shown in Fig. 9, and it is also tough for cytotechnologist to discriminate some nuclei from inflammatory cells in this sample. Verified by cytotechnologist, an average of 95.0% inflammatory cells are detected at pixel level at the first stage, and about an improvement to 96.0% at the second stage, and they are well inpainted at the third stage without causing unexpected contents. While inflammatory cells are often filtered by chemical methods to get satisfied cytology samples, our proposed method provide the possibility to filter them with digital methods efficiently.

TABLE II
THE PERFORMANCE OF ELIMINATION FROM DIGITAL CYTOLOGY

ID	GT	NN TP	NN FP	HC TP	HC FP
a	31	31	1	31	0
b	42	38	1	42	0
c	10	8	1	10	0
d	10	10	0	10	0
e	6	5	0	6	0
f	55	51	1	55	0
g	28	23	0	27	0
h	33	32	2	32	0
100 images	28.28	25.03	0.70	26.77	0.02

V. CONCLUSION AND FURTHER WORK

In this paper, we propose a systematic approach to evaluate inflammation coverage and remove the inflammatory cells from digital cytology microscopic images, based on primary real-life clinical dataset analysis. We design a two-stage framework for the whole framework. In the first step, to accurately estimate the inflammation, we design an end-to-end residual neural network for cell detection and segmentation. In the next step, for an elimination of the inflammatory cells that often taken as false positive or false negative in the cervical cell subtype analysis, we employ morphological operations for boundary correction and a biased PatchMatch algorithm to fill the vacancies with reasonable background patterns. As we use a very small annotated samples for training and large dataset for inference, a higher proportion of inflammatory cells are expected to be detected and removed by further annotated training data. As our final goal is to assist cancer diagnosis, we have also developed a classification neural network. This classification task uses the results of the nuclei locations from the first stage of our proposed framework. We experimentally find the enhanced aggregation help to reduce false positives in abnormal cell or cancer detection.

REFERENCES

[1] J. Jantzen, J. Norup, G. Dounias, and B. Bjerregaard, "Pap-smear benchmark data for pattern classification," *Nature Inspired Smart Information Systems (NiSIS)*, 01 2005.

[2] L. Zhang, H. Kong, C. T. Chin, S. Liu, X. Fan, and T. Wang, "Automation-assisted cervical cancer screening in manual liquid-based cytology with hematoxylin and eosin staining." *Cytometry. Part A : the journal of the International Society for Analytical Cytology*, vol. 85 3, pp. 214–30, 2014.

[3] P. E. Castle, S. L. Hillier, L. K. Rabe, A. Hildesheim, R. Herrero, M. C. Bratti, M. E. Sherman, R. D. Burk, A. C. Rodriguez, M. Alfaro, M. L. Hutchinson, J. Morales, and M. Schiffman, "An association of cervical inflammation with high-grade cervical neoplasia in women infected with oncogenic human papillomavirus (hpv)," *Cancer Epidemiology and Prevention Biomarkers*, vol. 10, no. 10, pp. 1021–1027, 2001. [Online]. Available: https://cebp.aacrjournals.org/content/10/10/1021

[4] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105. [Online]. Available: http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf

[5] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," *IEEE Conference Publication*, 2014.

[6] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[7] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition." *CoRR*, vol. abs/1512.03385, 2015. [Online]. Available: http://dblp.uni-trier.de/db/journals/corr/corr1512.htmlHeZRS15

[8] S. Ren, K. He, R. B. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," *CoRR*, vol. abs/1506.01497, 2015. [Online]. Available: http://arxiv.org/abs/1506.01497

[9] N. Dhungel, G. Carneiro, and A. P. Bradley, "The automated learning of deep features for breast mass classification from mammograms," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*, S. Ourselin, L. Joskowicz, M. R. Sabuncu, G. Unal, and W. Wells, Eds. Cham: Springer International Publishing, 2016, pp. 106–114.

[10] M. Moradi, Y. Gur, H. Wang, P. Prasanna, and T. Syeda-Mahmood, "A hybrid learning approach for semantic labeling of cardiac ct slices and recognition of body position," in *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, April 2016, pp. 1418–1421.

[11] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," *Lecture Notes in Computer Science*, pp. 234–241, May 2015.

[12] Y. Liu, K. Gadepalli, M. Norouzi, G. E. Dahl, T. Kohlberger, A. Boyko, S. Venugopalan, A. Timofeev, P. Q. Nelson, G. S. Corrado, J. D. Hipp, L. Peng, and M. C. Stumpe, "Detecting cancer metastases on gigapixel pathology images." *CoRR*, vol. abs/1703.02442, 2017. [Online]. Available: http://dblp.uni-trier.de/db/journals/corr/corr1703.htmlLiuGNDKBVTNCHPS17

[13] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, pp. 115–, Jan. 2017. [Online]. Available: http://dx.doi.org/10.1038/nature21056

[14] J. Ke, Z. Jiang, C. Liu, T. Bednarz, A. Sowmya, and X. Liang, "Selective detection and segmentation of cervical cells," in *Proceedings of the 2019 11th International Conference on Bioinformatics and Biomedical Technology*, ser. ICBBT'19. New York, NY, USA: ACM, 2019, pp. 55–61. [Online]. Available: http://doi.acm.org/10.1145/3340074.3340081

[15] M. Ali, A. Sarwar, V. Sharma, and J. Suri, "Artificial neural network based screening of cervical cancer using a hierarchical modular neural network architecture (hmnna) and novel benchmark uterine cervix cancer database," *Neural Computing Applications*, no. 4, pp. 1–15, 2017.

[16] L. Zhang, M. Sonka, L. Lu, R. M. Summers, and J. Yao, "Combining fully convolutional networks and graph-based approach for automated segmentation of cervical cell nuclei," in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, April 2017, pp. 406–409.

[17] N. B. Byju, V. K. Sujathan, P. Malm, and R. R. Kumar, "A fast and reliable approach to cell nuclei segmentation in pap stained cervical smears," *Csi Transactions on Ict*, vol. 1, no. 4, pp. 309–315, 2013.

[18] Y. Song, L. Zhang, S. Chen, D. Ni, B. Y. Lei, and T. Wang, "Accurate segmentation of cervical cytoplasm and nuclei based on multiscale convolutional network and graph partitioning." *IEEE Trans. Biomed. Engineering*, vol. 62, no. 10, pp. 2421–2433, 2015. [Online]. Available: http://dblp.uni-trier.de/db/journals/tbe/tbe62.htmlSongZCNLW15

[19] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation." *CoRR*, vol. abs/1411.4038, 2014. [Online]. Available: http://dblp.uni-trier.de/db/journals/corr/corr1411.htmlLongSD14

[20] P. Soille, *Morphological Image Analysis: Principles and Applications*. Emerald Group Publishing Limited, 2003.

[21] C. Barnes, E. Shechtman, A. Finkelstein, and D. B. Goldman, "PatchMatch: A randomized correspondence algorithm for structural image editing," *ACM Transactions on Graphics (Proc. SIGGRAPH)*, vol. 28, no. 3, Aug. 2009.

[22] D. Simakov, Y. Caspi, E. Shechtman, and M. Irani, "Summarizing visual data using bidirectional similarity," in *2008 IEEE Conference on Computer Vision and Pattern Recognition*, June 2008, pp. 1–8.