Selective Detection and Segmentation of Cervical Cells

Jing Ke
Shanghai Jiao Tong University, Shanghai 200240, China [1]
University of New South Wales, Sydney, NSW 2052, Australia [2]
kejing@sjtu.edu.cn

Zhaoming Jiang, Changchang Liu
Shanghai Jiao Tong University
Shanghai 200240, China
jiangzhaoming@sjtu.edu.cn

Tomasz Bednarz
University of New South Wales
Sydney, NSW 2052, Australia
t.bednarz@unsw.edu.au

Arcot Sowmya
University of New South Wales
Sydney, NSW 2052, Australia
a.sowmya@unsw.edu.au

Xiaoyao Liang
Shanghai Jiao Tong University
Shanghai 200240, China
liang-xy@cs.sjtu.edu.cn

ABSTRACT

Accurate detection and segmentation of cervical cells is often considered as a critical prerequisite of the prediction of dysplasia or cancer either by a pap smear or the lately developed liquid-based cytology (LBC). The computer-aided detection in microscope images can relieve the pathologists from strenuous manual labors with higher accuracy and efficiency. In the segmentation tasks of real-life clinical data, one challenging issue is the mis-identification of other cells, such as inflammatory cells, with similar appearance of nuclei in shape, size and texture. With a large distribution in the whole slide, even overlap up to 50% to 75% percentage of normal or abnormal cells, these cells are usually detected and segmented as nuclei. In this paper, compared with the typical three-catalogue segmentation methods of nuclei, cytoplasm and background proposed in the literature, we provide a discrimination between inflammatory cells and nuclei by adding a new catalogue. We present two novel convolutional neural networks (CNN), a deeply fine-tuned model and a trained from scratch model. The models enable us to sensitively detect and remove background noises such as mucus or red blood cells. We also profile a detailed performance comparison between these two methods, with the advantages of either network presented. The experiments are based on the sufficient clinical dataset we collected, and the results show the effectiveness of proposed approaches in selective cell detection and segmentation.

CCS Concepts

Computing methodologies → Neural networks;
Applied computing → Imaging;
Computing methodologies → Image segmentation

Keywords

Cervical cancer; Nuclei detection and segmentation; Inflammatory cells; Convolutional neural network; Deep learning

1. INTRODUCTION

ICBBT'19, May 29–31, 2019, Stockholm, Sweden

ACM ISBN 978-1-4503-6231-3/19/05…$15.00
DOI:https://doi.org/10.1145/3340074.3340081

In recent years, the number of deaths caused by cervical cancer has decreased significantly thanks to the regular Pap tests, where the early stage abnormal cells can be found by liquid-based cytology (LBC) or Pap smear test [4]. LBC is an advance in cervical cancer screening with advantages in sample quality, reproducibility, sensitivity, and specificity [4] compared with Pap smear test. Computer-Aided Detection/Diagnosis (CAD/CADx) in microscope images can relieve pathologists from strenuous manual labors, with relatively low cost, high efficiency and less misdiagnosis.

As early as mid-1950s, methods were discussed for automatic screening of normal from abnormal smears of exfoliated cells, aiming at increasing the productivity of mass screening for cervical cancer. At that time, cytology literature described that, the cancer cells distinguish themselves from the normal cells, with the features such as enlarged or more deeply stained nuclei, and a smaller cytoplasm-nuclear ratio. Since Antoni van Leeuwenhoek's pioneering steps in the 1970s towards exploiting microscopic imaging at the cellular level, it has been recognized that the central problem, often regarded as the cornerstone of image analysis in biomedical studies, is the image segmentation [13, 19].

In traditional model-based methods for segmentation, variation in the shape of the cells was studied to classify cancerous cells for their highly heterogeneous shapes and size features [18]. One method proposed was the level-set segmentation of cervical cells in [6]. A database generated by synthetically overlapping images of free-lying cervical cells was used. An edge-based approach using customized Laplacian of Gaussian (LoG) filter was designed to segment cell nuclei in bright-field microscope images of Pap smear [2]. For the advantage of finding boundaries, a graph-search based technique was used in abnormal nucleus segmentation in Herlev [8] and HEMLBC dataset [11]. In recent years, as object recognition and detection methods [5, 7, 8] have been significantly improved by deep learning, some learning-based models have been studied extensively in segmentation. An unsupervised approach is proposed which involves multi-scale hierarchical segmentation algorithm and a binary classifier for the separation of nuclei from cytoplasm with the cell region [3]. A cascade sparse regression chain model (CSRCM) is proposed, training a regressor instead of a classifier to return object locations and entire boundaries of clustered objects simultaneously, where the detection and segmentation problem was formulated in a unified regression problem [16]. More generally, for biomedical image segmentation, a popular U-net [14] is proposed based on the concept of Fully Convolutional Networks (FCN) [10].

However, nowadays the success of cell segmentation is still out of reach today due to the complexity of histopathological images such as proximate cells, background clutters, poor contrast, as

well as large variability in shape and size, stains, and cell densities. We have noticed that in real-life cervical cell images, inflammatory cells, including subtypes of neutrophils, eosinophils and lymphocytes, may have a very large distribution in the whole slide and often overlap on other normal or abnormal cells in pap smear of LBC, shown in Figure 1. Nevertheless, the methods proposed in the literature mainly approach to an accurate segmentation between nuclei and cytoplasm or clear hypothetical boundaries in overlapped cells, while inflammatory cells were either absent from the scenarios [23] or segmented as nuclei [2,3,16,17]. With quite similar appearance in shape, size and texture of nuclei but little cytoplasm around, inflammatory cells are often recognized as nuclei, and this mis-identification may cause potential problems in the later cervical dysplasia or cancer cell recognition.

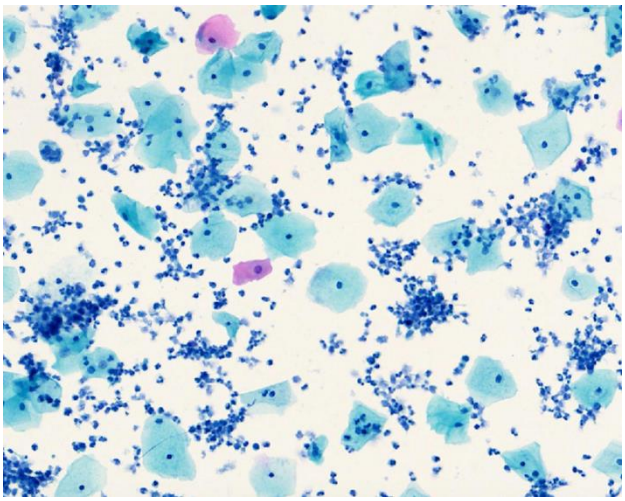

Figure 1. A massive number of inflammatory cells spread in digital cervical cell images

Although widely spread in the image, inflammatory cells are not included in widely used public dataset like Herlev or HEMLBC 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.

In this paper, to address the issue presented, we propose two convolutional neural networks (CNN) based architecture for reliable detection of nuclei. Our main results are summarized as follows: 1) We separate cytoplasm, nuclei, inflammatory cells and background into four categories, marked by different color, while mucus, red blood cells and other noises are ignored in boundary detection. If the Pap smear showed that the inflammation is severe, a detailed examination is often required [4], so we catalogue them to another type in the cell detection. To the best of our knowledge, this is the first approach which is capable of distinguishing nuclei from other cells similar in shape or size such as inflammatory cells in cervical cell segmentation. 2) Without hand-crafted features, the we propose a hybrid architecture of both standard and dilated convolution employing U-Net as a baseline and a modified VGG16 based framework with an extension of transfer learning [1, 20] for the segmentation and classification task. 3) We conducted a detailed comparison to evaluate these two approaches and demonstrate the advantages of either network based on the sample data and experimental tests. 4) With only a very small set of annotated samples of the target catalogues, we perform the training and test from the cervical LBC clinical data we collected in the recent 18 months, among which 90% from positive samples and 10% from negative samples, covering all typical subtypes of cervical cancer. To the best of our knowledge, these cervical cell microscopy images we use for detection and segmentation have the most complex appearances compared with related works.

2. METHOD

We proposed two end-to-end training architectures to jointly detect and segment nuclei, cytoplasm, inflammatory cells and background. We employ U-Net as the decoder and the encoder part of either architecture is modified. The VGG16 based architecture is fine-tuned with our annotated samples in transfer learning and the HSDC architecture is trained with only our data learning from scratch. The encoder accepts a random size input map, and the predicted segmentation map has a lower resolution compared with input. Given the large size of histology image, the input size is dependent on the performance achieved on GPU.

2.1 Encoding Module based on the VGG16

In this method, initial weights are generated by the VGG16 network that was pre-trained on a large dataset ImageNet. VGG16 is a well-established CNN classifier for image recognition tasks and has been utilized in many applications. This fast transition of initial weight effectively adapts the features to our histological images. Its advantage of preventing overfitting issues is well-suited in the limited labeled data in fine-tuning. The encoder employs part of VGG16, containing 13 convolution layers with 3x3 kernels and 4 max-pooling layers, but replaces the last pooling layer with a dropout layer with a rate of 0.5, shown in Figure 2.

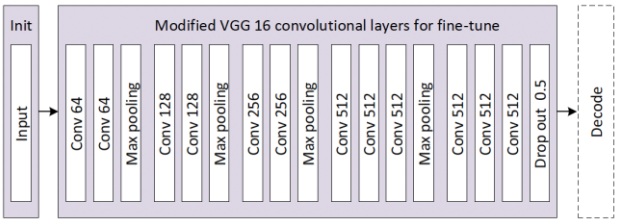

Figure 2. Encoder layers employed from VGG16

Generally, in the 5-level encoder, higher layers extract general features, while lower levels extract more complex and specific features [20]. As the pre-trained image is performed on very large dataset while the target dataset is quite small and dissimilar, we retain the weight in the upper three levels for general features, replace the last fully connected layer with the four-class segmentation prediction, and fine tune the last two levels for specific features. In the first step during the transfer learning, the first four levels are frozen, and the last level is learnt from scratch on our labeled cervical cell dataset. When the training starts to reach convergence, the second lowest level is unfrozen for training together with the last level. In this fine-tune, L2-normalization regulation is applied on trained layers, and the penalty factor is determined up to the best performance in benchmark test.

2.2 Encoding Module with HSDC Block

In this proposed architecture, a Hybrid Standard and Dilated Convolution (HSDC) block is designed and used in each level in the descending part. In standard convolution, a larger filter can collect more feature information, yet at a cost of introducing more parameters to learn and occupying more GPU memory. Dilated convolution is a method of increasing receptive view by

expanding the size of convolution kernel, where some pixels are skipped by input and filled with value 0 [11, 21]. In this way, we aggregate multi-scale contextual information without losing resolution or analyzing rescaled images. For a kernel of size (i.e. length on a dimension) $S_{original}$ and a dilation rate $R_{dilation}$, the dilated kernel size is $S_{dilated} = (S_{original} - 1) \times R_{dilation} + 1$. In this case, the dilated convolution runs faster compared with a standard convolution with the same kernel size. However, in a dilated kernel, as part of the pixels are not effective, feature information will be lost especially for the pixels around the kernel center, which leads to a loss of accuracy in pixel-level dense prediction. To address this issue, we add standard convolution layers for output to different channels to retain the capability in feature extraction, which is named the Hybrid Standard and Dilated Convolution (HSDC) block in Figure 3.

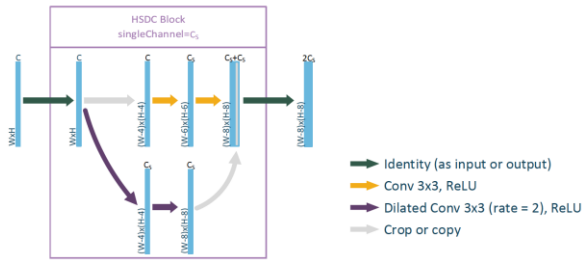

Figure 3. Hybrid Standard and Dilated Convolution (HSDC) block. Two standard and two dilated convolutions output a number of channels C_S independently

One HSDC block is composed of two standard and two dilated convolutions. For F as input with N channels, G as output of $2*C_S$ channels, each channel c is mapped to an output of $G(c)$. In each channel, K is kernel radius, h and w are the coordination and n is the stride in dilated convolution, the output value G is either performed by convolution or dilated convolution, described as:

$$G(h,w,c) = \begin{cases} \sum_{c=1}^{N} \sum_{i,j=-s}^{s} F(h+i, w+j, c) * K(i,j,c), & 1 \le c \le C_S \\ \sum_{c=1}^{N} \sum_{i,j=-s}^{s} F(h+n*i, w+n*j, c) * K(i,j,c), & C_S \le c \le 2C_S \end{cases} \quad (1)$$

The network employing HSDC blocks as encoder is shown in Figure 4. Consequently, fewer convolution layers involved in this network, it achieves a larger receptive field than the VGG16 based network. Before the commencement of the training phase, weights in each convolutional layer are initialized by values randomly sampled from a normal distribution with a zero mean and small standard deviation, where full training is performed on our target dataset. The same as VGG16 based architecture, a dropout rate of 0.5 is used in the last two layers to prevent over-fitting and the L2-normalization is applied.

2.3 Data Pre-processing and Augmentation

As the cervical cell images made from LBC is much better than Pap smear, we need not perform any pre-processing for an improved contrast between cells and the background. For segmentation annotation, we classify materials into four catalogues of cytoplasm, nuclei, inflammatory cells and background with manually labeled masks confirmed by pathologists. We do not give specific labels to normal or abnormal cells, neither the grading nor classification of subtypes of cancers.

The training dataset contains 100 labeled sub-images (either annotated or confirmed by pathologists), 90% with presence of abnormal cells, covering all subtypes of cervical cells we collected, and their segmentation mask is used as ground truth for evaluation.

Data augmentation improves the accuracy of output from convolutional networks and reduces overfitting [10]. Elastic deformation is used in our experiment as a major method for data augmentation in training for the slides made from pap smear or LBC in cervical cancer screening. We also rotate the sampled images, as it is crucial to the success of the image recognition of convolutional networks [15]. The most common methods of flipping and shifting are applied to increase data samples as well. But rescaling or stretching methods are not applicable as the size and shape are considered as a crucial factor in cell recognition. The feature of abnormality or normality should not be changed in data augmentation for microscopy image, which is quite different from data augmentation for general object detection applications.

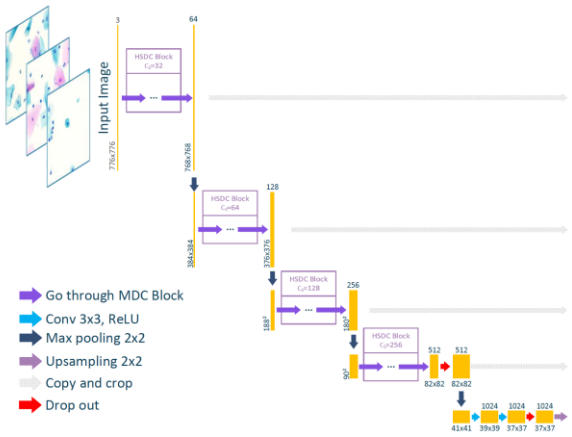

Figure 4. Encoder with HSDC block for a larger receptive field, applicable to any image input size

3. RESULTS AND DISCUSSION

We collected the cervical dataset mainly from 6 hospitals in the past 18 months with sensitive information removed. The subtype classification and ground truth of segmentation was either annotated or confirmed by pathologists. All slides were made by LBC technique stained with H&E. In total we collected 49 positive samples, approved by hospitals, all presented at a high resolution of 40x magnification and around 56,000x56,000 pixels in one microscopy image. We cropped 100 typical images for training and 45 images for test, covering many kinds of complex appearances of all subtypes of cervical cells collected. The size of training images is around 900x900.

We do not apply the algorithms to public dataset like Herlev or HEMLBC dataset, since they contain only classified single cell with clear appearance, where the complexities of touching cells, widely spread inflammatory cells, blood cells in broken or clusters are not considered. As there is no previous work in inflammatory cell detection and segmentation, we present a detailed performance comparison between the two models for transfer learning and full learning respectively.

3.1 Experiment Results

Generally, the VGG16 based network has a better performance in object recognition even when the objects are absent in the training data compared with the HSDC model, as its weight is acquired from pre-training on ImageNet for transfer learning, especially for general feature extractions. The HSDC is trained from scratch on our labeled cervical cell training dataset. But it is also prone to produce tiny fragments in segmentation, and sometimes misidentifies one nucleus as multiple nuclei, incurring false positive (FP) in segmentation result. The HSDC encoder uses a larger receptive field in convolution operation, thus it usually produces a smooth boundary between nuclei and cytoplasm, or between cytoplasm and background, and approaches better to the ground truth in microscopy images. Applied on the clinical data, the advantages and weaknesses in classification and segmentation tasks are compared and presented in Figure 5 to Figure 7 between the proposed two networks. In these figures, (a) represents the cropped original cervical cell images for network input maps, (b) and (c) show the output segmentation results from the proposed VGG16 based network and HSDC network respectively, and their segmentation mask results are presented at the lower row. Figure 5 and Figure 6 show test results with the presence of superficial squamous, dysplasia and inflammatory cells in the clipped clinical cervical images. Both networks perform well at detection and segmentation of nuclei and cytoplasm, as well as recognition of isolated inflammatory cells at the background or overlapped ones on other cells.

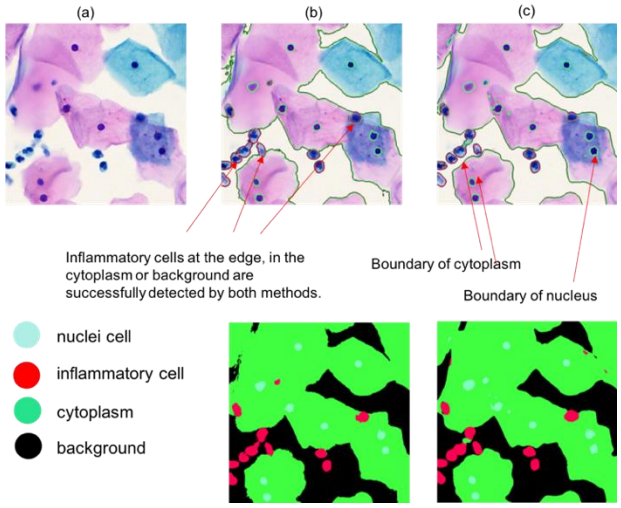

Figure 5. Most of the scenario in a normal slide contain both inflammatory cells and superficial squamous cells, where inflammatory cells at the edge, in the cytoplasm or in the background can be successfully detected by both methods

Designed to classify a variety kind of objects, the VGG16 based network has a better location and recognition of tiny nuclei, but prone to mistakenly segmenting one nucleus into several and resulted in over segmentations. In the contrast, the HSDC network performs better at recognition of a single nucleus even with inhomogeneity in the H&E stain as it contains more contextual information at each layer. Figure 7 shows the result with a subtype of carcinoma cancer that is not labeled or masked in the training data. The result demonstrates that, without right annotation in training images, the VGG16 based network makes better predictions as the initial weights are acquired from pre-training on ImageNet from a large dataset, but the HSDC network fails to detect and segment such materials and hence treats them as noises hence not included in detection and segmentation. Consequently, the performance of the HSDC network will probably get improved with more annotated samples involved, but we notice it is also more sensitive at recognizing a certain subtype of cells as leant from scratch. The results of segmentation mask also clearly show that the overlapped or broken blood cells are ignored for segmentation in both methods.

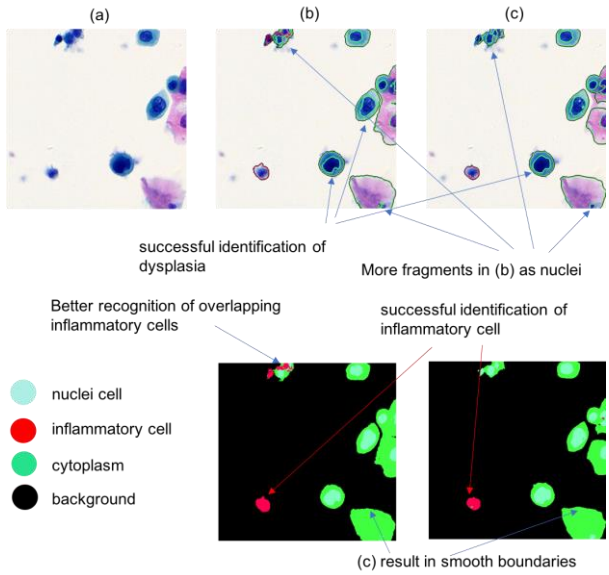

Figure 6. Detection and segmentation results with dysplasia and inflammatory cells in the scenario. Generally, the VGG16 based network has a better location and recognition of tiny nuclei, but prone to view one nucleus as several. The HSDC network performs better in output smooth boundaries and identification of a single nucleus even with stain variation

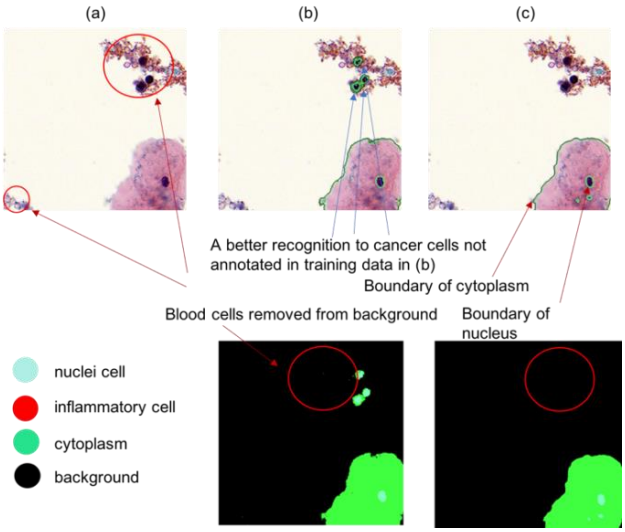

Figure 7. Segmentation results in cancer cells not annotated in training data. The VGG16 based network performs better at feature extraction as the initial weights are acquired from pre-training. The HSDC network cannot recognize and segment such kind of nuclei or cytoplasm

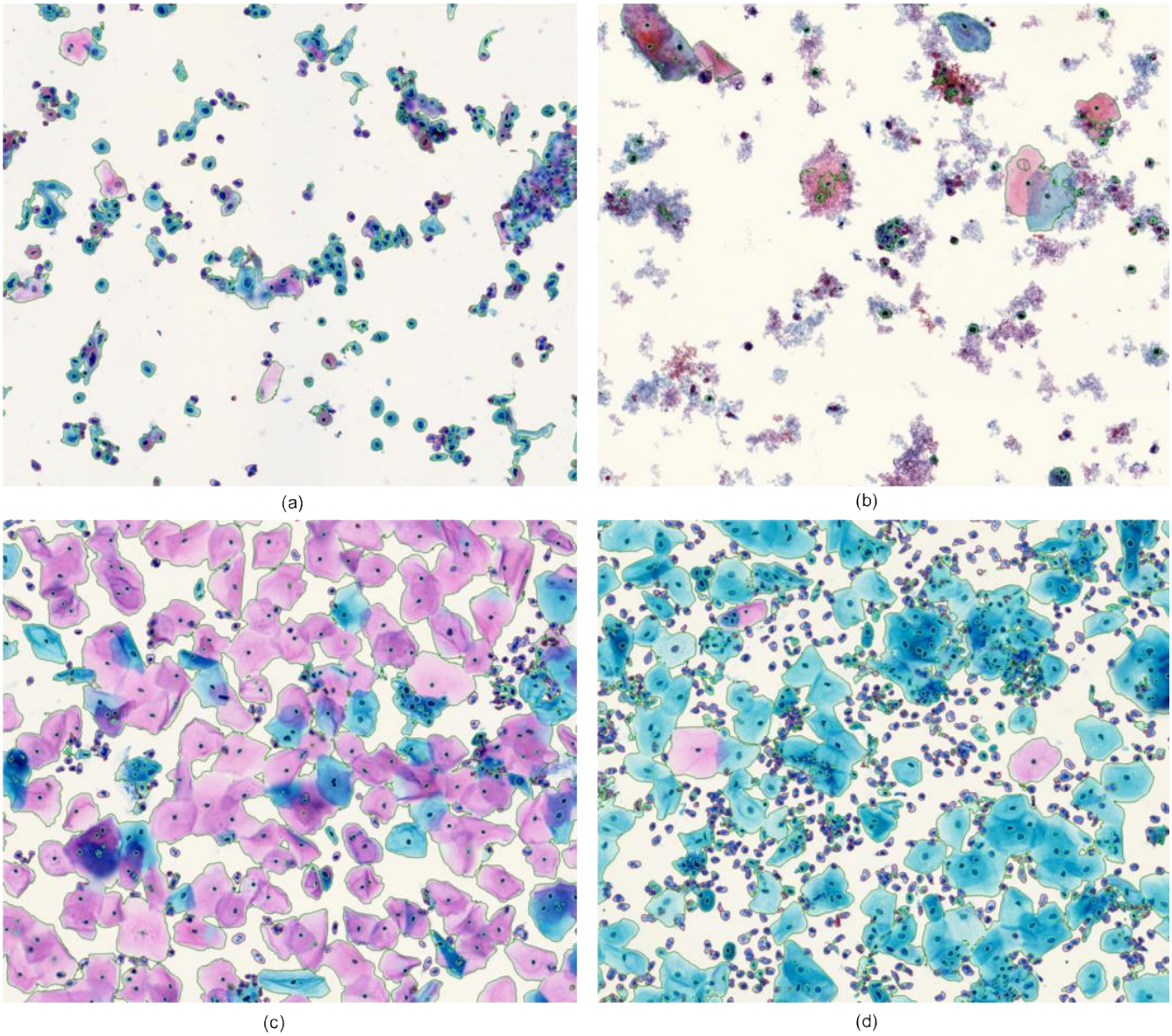

Figure 8. Abnormal cervical cells. (a) Carcinoma cells (b) Carcinoma cells with red blood cells (blood cells are detected rather than catalogued to cytoplasm or nuclei) (c) Low-grade squamous cells contained while most cells are normal (d) Superficial squamous cells overlapped by a massive number of inflammatory cells. We give a discrimination between nuclei outlined with blue color and inflammatory cells outlined with red color

We select some typical and challenging clinical microscopy images and present the detection and segmentation results in Figure 8, all with abnormal cervical cells that in (a) carcinoma cells are presented with several normal cells; (b) shows carcinoma cells with red blood cells around or overlapped, and blood cells are detected rather than catalogued to cytoplasm or nuclei; (c) shows several low-grade squamous cells while most cells are normal; (d) shows superficial squamous cells overlapped by a massive number of inflammatory cells, and these inflammatory cells are often treated as nuclei in previous segmentation tasks.

The accurate location and segmentation results are the essential premise of successful cancer cell recognition and diagnosis. We also notice that the boundary location and nuclei recognition for normal cells are much better than abnormal cells, as there is strong contrast between nucleus and cytoplasm in the appearance of normal cells. As a result, the precision is prone to decline when more positive samples involved in the experimental tests.

3.2 Performance

We evaluate the performance of our methods by several criteria for our special functions at cervical cell recognition and segmentation. Table 1 and Table 2 shows recall rate, precision rate, and F1 score of the two networks for segmentation results of four catalogues of background (Bak), cytoplasm (Cyt), nuclei (Nuc) and inflammatory cell (Inf). For a catalogue *i* in segmentation, we get recall:

$$R_i = \frac{TP_i}{TP_i + FN_i} \tag{2}$$

and precision:

$$P_i = \frac{TP_i}{TP_i + FP_i} \quad (3)$$

for F1 score calculation:

$$F1_i = \frac{2 \times P_i \times R_i}{P_i + R_i} = \frac{2 \times TP_i}{2 \times TP_i + FP_i + FN_i} \quad (4)$$

We can see that, with a very small number of samples trained, both networks have good performance at the discrimination between nuclei and inflammatory cells, which is one of the most important functions of the networks. The main error around 10% is caused by pixel-wise discrimination between nuclei and cytoplasm. They are caused either by deviant location in blurring boundaries or misidentification of a dark region in cytoplasm as a nucleus due to inhomogeneous stain in complex scenarios. We also noticed that many errors in pixel-wise classification are caused by inhomogeneity in H&E stain that even pathologists have difficulties in locating exactly accurate cytoplasm boundaries during annotation for manually labeled ground truth. This may add to the loss in recall and precision. Generally, the mis-identification between cytoplasm and nuclei at boundaries does not affect much on a successful detection of a nucleus or an inflammatory cell. For both methods, we achieved a good performance of only round 3% error in the location and identification between inflammatory cells and nuclei.

Table 1. Recall and precision of the proposed models

	Network	Ground Truth	Segmentation results			
			Bak	Cyt	Nuc	Inf
recall	VGG	Nuc	0.00221	0.113	**0.854**	0.0304
		Inf	0.0998	0.106	0.0360	**0.758**
	HSDC	Nuc	7.91E-05	0.118	**0.872**	0.00984
		Inf	0.100	0.180	0.0649	**0.655**
precision	VGG	Nuc	0.0100	0.0904	**0.865**	0.0351
		Inf	0.152	0.0550	0.0317	**0.761**
	HSDC	Nuc	0.00221	0.123	**0.816**	0.0585
		Inf	0.138	0.0236	0.0127	**0.826**

Table 2. F1 scores on each catalogue

Network	Class				Average
	Bak	Cyt	Nuc	Inf	
VGG	0.980	0.960	0.859	0.759	0.890
HSDC	0.981	0.961	0.843	0.730	0.879

The experimental tests are conducted on high-resolution LBC-based histopathology images, each up to 56,000 x 56,000 pixels at the magnification rate of 40. And in the benchmark test, the parameter setting of input maps size 1976x1976 and predicated map size 1700x1700 achieves the peak performance. The partitioned segmentation is performed within 14 minutes of a whole slide with both perimeter and mask output. The experiments are all carried out on NVIDIA GTX 1080Ti.

4. ACKNOWLEDGMENTS

This research is primarily supported by Tongwei Biotechnology Company. Great thanks to their funding, medical dataset and pathologists.

5. CONCLUSION

Based on the complex scenario challenges in real-life cervical cell clinical data, we propose a fine-tuned VGG16 based module and a learning-from-scratch HSDC module for the novel four-catalogue segmentation task. Both networks show good performance at detection, recognition and segmentation of cytoplasm, nuclei, inflammatory cells and background in microscopy image. The mucus, red blood cells and other noises are detected in the background and not involved in boundary detection and segmentation. We have a comparison between these two methods in detection and segmentation of challenging appearance of cells in a variety of complexities. We use only a small number of annotated samples for training, and the images we select for training and test have covered all the subtypes of cervical cells from the large dataset we collected, among which 90% from abnormal cells with complex shapes. We also analyzed the possibility of improving our results with a larger annotated dataset in training, especially with the HSDC based architecture. In the future work, we will have a further optimization on both networks we have proposed and give suggestions over which network to choose based on the provided annotations and further requirements.

6. REFERENCES

[1] Bengio, Y., Deep learning of representations for unsupervised and transfer learning. In *Proceedings of ICML Workshop on Unsupervised and Transfer Learning*, 17-36, 2012.

[2] Byju, N., Sujathan, V.K., Malm, P., and Kumar, R.R., A fast and reliable approach to cell nuclei segmentation in PAP stained cervical smears. *CSI transactions on ICT 1*, 4, 309-315, 2013. DOI= http://dx.doi.org/10.1007/s40012-013-0028-y.

[3] Gen çtav, A., Aksoy, S., and Önder, S., Unsupervised segmentation and classification of cervical cell images. *Pattern recognition 45*, 12, 4151-4168, 2012. DOI= http://dx.doi.org/10.1016/j.patcog.2012.05.006.

[4] Gibb, R. K. and Martens, M.G., The impact of liquid-based cytology in decreasing the incidence of cervical cancer. *Reviews in Obstetrics and Gynecology 4*, Suppl 1, S2, 2011.

[5] Girshick, R., Donahue, J., Darrell, T., and Malik, J., Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 580-587, 2014. DOI= http://dx.doi.org/10.1109/cvpr.2014.81.

[6] He, K., Zhang, X., Ren, S., and Sun, J., Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770-778, 2016. DOI= http://dx.doi.org/10.1109/cvpr.2016.90.

[7] Hu, R., Doll ár, P., He, K., Darrell, T., and Girshick, R., Learning to segment everything. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 4233-4241, 2018. DOI= http://dx.doi.org/10.1109/cvpr.2018.00445.

[8] Jantzen, J., Norup, J., Dounias, G., and Bjerregaard, B., Pap-smear benchmark data for pattern classification. *Nature inspired Smart Information Systems (NiSIS 2005)*, 1-9, 2005.

[9] Krizhevsky, A., Sutskever, I., and Hinton, G.E., Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, 1097-1105, 2012. DOI= http://dx.doi.org/10.1145/3065386.

[10] Long, J., Shelhamer, E., and Darrell, T., Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 3431-3440, 2015. DOI= http://dx.doi.org/10.1109/cvpr.2015.7298965.

[11] Lu, Z., Carneiro, G., and Bradley, A.P., An improved joint optimization of multiple level set functions for the segmentation of overlapping cervical cells. *IEEE Transactions on Image Processing 24*, 4, 1261-1272, 2015. DOI= http://dx.doi.org/10.1109/tip.2015.2389619.

[12] Mehta, S., Rastegari, M., Caspi, A., Shapiro, L., and Hajishirzi, H., Espnet: Efficient spatial pyramid of dilated convolutions for semantic segmentation. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 552-568, 2018. DOI= http://dx.doi.org/10.1007/978-3-030-01249-6_34.

[13] Meijering, E., Cell segmentation: 50 years down the road [life sciences]. *IEEE Signal Processing Magazine 29*, 5, 140-145, 2012. DOI= http://dx.doi.org/10.1109/msp.2012.2204190.

[14] Ronneberger, O., Fischer, P., and Brox, T., U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* Springer, 234-241, 2015. DOI= http://dx.doi.org/10.1007/978-3-319-24574-4_28.

[15] Shin, H.-C., Roth, H.R., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D., and Summers, R.M., Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Transactions on Medical Imaging 35*, 5, 1285-1298, 2016. DOI= http://dx.doi.org/10.1109/tmi.2016.2528162.

[16] Song, J., Xiao, L., and Lian, Z., Contour-Seed Pairs Learning-Based Framework for Simultaneously Detecting and Segmenting Various Overlapping Cells/Nuclei in Microscopy Images. *IEEE Transactions on Image Processing 27*, 12, 5759-5774, 2018. DOI= http://dx.doi.org/10.1109/tip.2018.2857001.

[17] Song, Y., Tan, E.-L., Jiang, X., Cheng, J.-Z., Ni, D., Chen, S., Lei, B., and Wang, T., Accurate cervical cell segmentation from overlapping clumps in pap smear images. *IEEE Transactions on Medical Imaging 36*, 1, 288-300, 2017. DOI= http://dx.doi.org/10.1109/tmi.2016.2606380.

[18] Wienert, S., Heim, D., Saeger, K., Stenzinger, A., Beil, M., Hufnagl, P., Dietel, M., Denkert, C., and Klauschen, F., Detection and segmentation of cell nuclei in virtual microscopy images: a minimum-model approach. *Scientific Reports 2*, 503, 2012. DOI= http://dx.doi.org/10.1038/srep00503.

[19] Xue, Y. and Ray, N., Cell Detection in Microscopy Images with Deep Convolutional Neural Network and Compressed Sensing, 2017. *arXiv preprint arXiv:1708.03307*.

[20] Yosinski, J., Clune, J., Bengio, Y., and Lipson, H., How transferable are features in deep neural networks? In *Advances in neural information processing systems*, 3320-3328, 2014.

[21] Yu, F. and Koltun, V., Multi-scale context aggregation by dilated convolutions, 2015. *arXiv preprint arXiv:1511.07122*.

[22] Zhang, L., Kong, H., Liu, S., Wang, T., Chen, S., and Sonka, M., Graph-based segmentation of abnormal nuclei in cervical cytology. *Computerized Medical Imaging and Graphics 56*, 38-48, 2017. DOI= http://dx.doi.org/10.1016/j.compmedimag.2017.01.002.

[23] Zhang, L., Lu, L., Nogues, I., Summers, R.M., Liu, S., and Yao, J., DeepPap: deep convolutional networks for cervical cell classification. *IEEE journal of biomedical and health informatics 21*, 6, 1633-1643, 2017. DOI= http://dx.doi.org/10.1109/jbhi.2017.2705583.