MULTI-INSTANCE DETECTION AND CLASSIFICATION OF RED BLOOD CELLS IN MICROSCOPIC IMAGES FOR SICKLE CELL DISEASE TESTING

*Wei Qiu**[1,2], *Jiaming Guo**[1,2], *Xiang Li**[1], *Mengjia Xu*[2], *Mo Zhang*[2], *Ning Guo*[1], *Quanzheng Li*[1]

[1]Department of Radiology, Massachusetts General Hospital, Boston, MA.
[2]Peking University, Beijing, China.
*Joint first authors

ABSTRACT

Cell detection and cell type classification from biomedical images plays an important role for high-throughput imaging and various clinical application. While classification of single cell sample can be performed using standard computer vision and machine learning methods, analysis of multi-instance samples (region containing congregating cells) is more challenging, as separation of individual cells can be difficult (e.g. touching cells) or even impossible (e.g. overlapping cells). As multi-instance images are common in analyzing Red Blood Cell (RBC) for Sickle Cell Disease (SCD) diagnosis, we develop and implement a multi-instance cell detection and classification framework to address this challenge. The framework firstly trains a region proposal model based on Region-based Convolutional Network (RCNN) to obtain bounding-boxes of regions potentially containing single or multiple cells from input microscopic images, which are extracted as image patches. High-level image features are then calculated from image patches through a pre-trained Convolutional Neural Network (CNN) with ResNet-50 structure. Using these image features inputs, six networks are then trained to make multi-label prediction of whether a given patch contains cells belonging to a specific cell type. As the six networks are trained with image patches consisting of both individual cells and touching/overlapping cells, they can effectively recognize cell types that are presented in multi-instance image samples. Finally, for the purpose of SCD testing, we train another machine learning classifier to predict whether the given image patch contains abnormal cell type based on outputs from the six networks. Testing result of the proposed framework shows that it can achieve good performance in automatic cell detection and classification.

Index Terms— cell type, sickle cell disease, multi-instance classification

1. INTRODUCTION

Sickle cell disease (SCD) is a type of inherited red blood cell (RBC) disorder which can cause life-threatening complications. Automatic classification and diseased cell detection based on cell texture and morphological features has become a viable and important approach for SCD diagnosis, as manual inspection of RBC images is time and labor-consuming. More generally, automatic cell detection and cell type classification is a crucial step of high-throughput imaging as well as many other clinical applications. Towards the purpose of cell detection and classification, various solutions have been developed, such as CellProler [1], CellTrack [2] or Fiji [3]. Recent advancement of deep learning-based approaches have shown superior performance in extracting more discriminative image features with higher generalizability for various biomedical image analysis tasks including cell classification [4, 5], detection [6], segmentation/semantic segmentation [7, 8, 9] and counting [10].

While deep learning-based approaches have achieved good performance in classifying single cell patches [11, 12] [13], in practice a common challenge is the presence of multiple cells congregating together in one sample image patch. We formulate this challenge as the "multi-instance classification" problem, where it can be difficult (e.g. touching cells) or even impossible (e.g. overlapping cells) to fully separate individual instances out in those samples. As normal classifiers are trained for only dealing with single instance, those multi-instance samples have to be discarded [5] during training, and can cause incorrect classification results if such samples are presented in testing data. Among various multi-instance methods that have been previously developed, CapsNet [14] can analyze highly overlapping objects and has inspired many applications based on it. However, most of the current models developed using CapsNet are focusing on single-label classification problem [15, 16] [17], due to limitation in the original CapNets that it does not allow more than one instance of the same class to be presented in the image.

To address the challenge of multi-instance classification in biomedical image analysis, while at the same time aiming at improving the diagnostic accuracy and efficiency for SCD, we propose a cell detection and classification framework that can automatically extract image patches consisting of single or multiple cells, and perform multi-label classification as well as abnormal cell detection on the extracted image patches.

Manuscript 644 submitted to 2019 IEEE International Symposium on Biomedical Imaging (ISBI). Received October 16, 2018.

The framework firstly performs region proposal of the input full-scale microscopic image through a Region-based Convolutional Network (RCNN) implemented by Faster-RCNN [18] and extract image patches automatically. In the next step, the proposed framework uses Convolutional Neural Network (CNN) with network structure of ResNet-50 and pre-trained on ImageNet dataset [19] to extract high-level image features (i.e. outputs from the last convolution layer) from the image patches. Afterwards, six classification networks using the extracted image features as input are trained to classify whether the input image patch contains cell(s) belonging to a specific cell types or not. Similar scheme for multi-label classification has also been applied in previous works [20, 21] and [22]. For the purpose of SCD testing, we further apply Gradient Boosting classifier to determine whether the given image patch contains "abnormal" cell types or not, based on the outputs from six classification networks. The proposed framework is tested on microscopic RBC images from SCD patients, showing its capability of performing fully automatic cell detection, cell type classification and SCD testing.

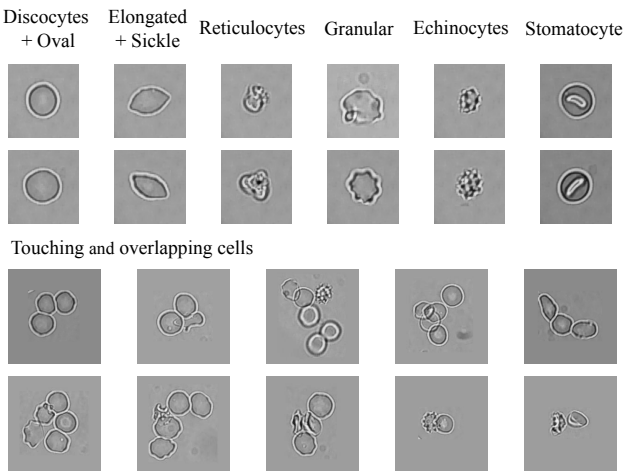

Fig. 1. Top panel: Sample image patches belonging to six cell types. Bottom panel: Sample images patches containing touching/overlapping cells.

2. MATERIALS AND METHODS

2.1. Data acquisition and bounding-box extraction

RBC microscopic images used in this work are collected from UPMC (University of Pittsburgh Medical Center). Raw data contains 313 images with size of 1920×1080. Details of data acquisition and description can be found in [5]. Data used in this work includes 1080 single-cell patches processed in [5], as well as 1389 multi-cell patches with touching/overlapping cells that are manually identified from raw images. According to the protocol in [5], we define six cell types for RBC, visualizations of the six cell types as well as samples of touching/overlapping cells can be found in Fig.1.

In order to automatically extract image patches from the full-scale 1920×1080 microscopic images, we utilize Faster-RCNN [18] model which has achieved state-of-the-art performances and high process speed for object detection and region proposal tasks. In the proposed framework, Faster-RCNN is trained on images with mask on regions that contain single or multiple cells. Thus it can accurately detect multi-instance image patches. Extracted image patches are then resized to 224×224 pixels in order to be used as input for later networks.

2.2. Multi-label classification with transfer learning

In order to perform effective multi-label cell classification in a supervised approach, one major challenge to be overcome is the lack of training samples, which is a common problem when applying deep learning methods for medical image analysis [23, 24, 25, 26]. To solve this, we develop a transfer learning scheme which utilizes ResNet-50 network [27] pre-trained on ImageNet [19] to extract high-level image features. Specifically, the pre-trained ResNet-50 is applied on all the available sample image patches (i.e. using them as testing input). Outputs from the last convolution layer of ResNet-50, which can be considered as high-level representation of the input image patches, are then stored and used for training the later cell type classification network. In this way, we transfer the massive information in the ImageNet database to this application through convolution operations, resulting in the extracted image features. These image features, formed as 2048-d vector for each input image patch, where 2048 is the number of convolution kernels in the last convolution layer of ResNet-50. The framework then trains six customized fully connected networks with one fully-connected layer and one softmax output layer. Each network performs binary classification for one cell type, where its input is the 2048-d feature vector, and output is the probability of whether a certain cell type is presented in the input image patch. Optimization of the classification networks is performed by Adam optimizer [28] with learning rate of 10^{-5}. The loss is measured by cross entropy with L2 regularization. Finally, outputs of the six networks are aggregated together into a 6-d vector, showing the probability for each of the six cell types. It should be noted that this output vector is not normalized (i.e. sum of probability is not 1), as we allow more than one cell types presented in the given image patch.

2.3. Binary classification for SCD Testing

As the ultimate goal of RBC image analysis for SCD testing is the detection of whether abnormal cells are presented in the given microscopic image, where "abnormal" is defined by five cell types: "Elongated and Sickle", "Reticulocytes", "Granular", "Echinocytes" and "Stomatocyte", we further construct a binary classifier using Gradient Boosting classi-

Manuscript 644 submitted to 2019 IEEE International Symposium on Biomedical Imaging (ISBI). Received October 16, 2018.

fier [29] to discriminate "normal" cells versus "abnormal" cells. Input of the Gradient Boosting classifier is the 6-d vector from the six classification networks, and the output is ground-truth knowledge of whether any abnormal cells are presented in the given image patch.

3. RESULTS

3.1. Classification performance of cell patches

To evaluate the performance of the proposed framework, we firstly test its cell type classification module (i.e. feature extraction and classification networks) on manually-identified cell patches through 5-fold cross validation. Classification accuracy for the six cell types, as measured by Area Under the Curve (AUC), are listed in the fist row in Table 1, marked by *Model A*. It can be seen that classification accuracy for all individual cell types are all above 0.9. Further, for a given image patch with arbitrary number of cells belonging to same or different cell types, the proposed model can simultaneously identify all the cell types at accuracy of 0.722. In comparison, if the proposed model is used to classify image patches containing only a single cell (second row in Table 1, marked by *Model A**), it can achieve overall classification accuracy of 0.932, which outperforms the accuracy reported in our previous work (0.893) [5].

In order to investigate whether the current classification module benefits from the extra multi-instance training samples, we further train a same set of six classification networks with only single-cell image patches. Its classification performance on the mixed dataset with both single and multi-cell patches is listed in the third row in Table 1, marked by *Model B*. Overall classification accuracy of *Model B* decreases dramatically comparing with *Model A* (0.649 versus 0.722). While it achieves higher accuracy for classifying "Oval+Disc" cell type (which contains the largest number of samples), for all the other five cell types its performance is lower.

Several sample cases where *Model A* (i.e. network used in the proposed framework) makes correct classification while *Model B* (single-cell network) fails are visualized in Figure 2. For image patch "sample 1" which contains cell types of "Oval" and "Stomatocyte", *Model A* correctly identify both cell types (with predicting probability of 0.999 and 0.615), while *Model B* classify the patch as only "Stomatocyte" (with predicting probability of 0.998). For image patch "sample 2" which contains cell types of "Oval" and "Echinocytes", *Model A* correctly identify both cell types (with predicting probability of 0.703 and 1), while *Model B* predicts no label for the patch (i.e. outputs from all six networks are lowered than the threshold). For image patch "sample 3" which contains cell types of "Oval" and "Granular", *Model A* correctly identify both cell types (with predicting probability of 1 and 0.542), while *Model B* classify the patch as only "Oval" (with predicting probability of 1). It can be found that for image patches containing multiple cell types, *Model B* will either predict only one label or no label at all, while *Model A* can identify all the correct cell types. The result shows that only by adding multi-cell data into the training samples, the classification network can learn how to accurately handle them.

(a) sample 1 (b) sample 2 (c) sample 3

Fig. 2. Visualizations of image patches containing multiple cell types. They are all correctly classified by *Model A* yet incorrectly classified by *Model B*.

Finally, we train the Gradient Boosting classifier from outputs of the six classification networks for patch-wise SCD testing. The proposed Gradient Boosting classifier achieves average accuracy of 85.1% through 5-fold cross validation, indicating that for a given image patch with arbitrary number of cells belonging to same or different cell types, the classifier can determine whether there is at least one abnormal cell at high accuracy.

3.2. Automatic analysis of full-scale microscopic image

By applying the Faster-RCNN module of the proposed framework on the input full-scale image, we can automatically obtain bounding-box of potential cells and the corresponding image patches for later classification analysis. Mean average precision(mAP) of Faster-RCNN on the experiment data is 0.899. A sample cell detection and classification result is shown in Fig.3. The sample result illustrates that our proposed framework is capable of performing fully automatic cell detection and classification from raw image input, achieving end-to-end image-based SCD testing, and readily usable in real practice.

4. CONCLUSIONS & DISCUSSION

In this work, we propose a deep learning-based framework to perform automatic cell detection and classification from RBC microscopic images. The framework is specifically designed to solve complex imaging scenario involving multi-instance classification problem, where cells in the input image can be touching or overlapping with each other and cannot be separated. Experimental results show that the classification networks utilizing transfer learning scheme can achieve better performance than baseline models and previous works, deal with more complex cell imaging conditions, and partially

Manuscript 644 submitted to 2019 IEEE International Symposium on Biomedical Imaging (ISBI). Received October 16, 2018.

Table 1. The ROC-AUC score of six classifiers using different experiment settings

Experiment	AUC						Accuracy
	Oval + Disc	Elon + Sick	Reti	Gran	Echi	Stom	
Mixed dataset train, mixed dataset test (*Model A*)	0.971	**0.943**	**0.967**	**0.933**	**0.985**	**0.908**	**0.722**
Mixed dataset train, single-cell dataset test (*Model A**)	0.995	0.994	1.000	0.998	0.999	0.998	0.932
Single-cell dataset train, mixed dataset test (*Model B*)	**0.974**	0.891	0.952	0.819	0.935	0.671	0.649

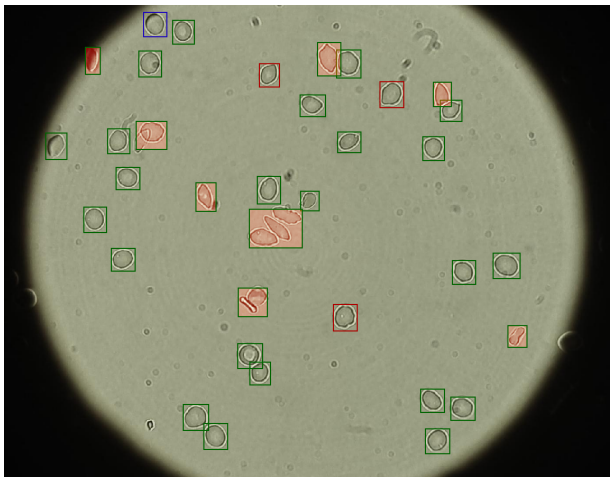

Fig. 3. Sample cell detection and classification result of our proposed framework. All cell regions (i.e. image patches) are surrounded in colored boxes. Patches that are missed by the detection network (Faster-RCNN) are colored in blue. Patches that are both successfully detected and classified are colored in green. Patches that are successfully detected yet mis-classified are colored in red. Regions with red mask highlight the presence of abnormal cell types correctly detected and classified by the proposed framework.

address the highly challenging multi-label, multi-instance classification problem. Testing results on full-scale raw microscopic image input show high robustness of the proposed framework and its potential usefulness in clinical practice.

As our proposed framework consists of multiple cascaded modules, while each module deals with one image analysis task individually, errors across different modules can accumulate which lowers the final SCD testing accuracy. Thus in the next step we will integrate the currently decoupled modules into an end-to-end framework.

5. REFERENCES

[1] Anne E. Carpenter, Thouis R. Jones, Michael R. Lamprecht, Colin Clarke, In Han Kang, Ola Friman, David A. Guertin, Joo Han Chang, Robert A. Lindquist, Jason Moffat, Polina Golland, and David M. Sabatini, "Cellprofiler: image analysis software for identifying and quantifying cell phenotypes," *Genome Biology*, vol. 7, no. 10, pp. R100, Oct 2006.

[2] Ahmet Sacan, Hakan Ferhatosmanoglu, and Huseyin Coskun, "Celltrack: an open-source software for cell tracking and motility analysis," *Bioinformatics*, vol. 24, no. 14, pp. 1647–1649, 2008.

[3] Johannes Schindelin, Ignacio Arganda-Carreras, Erwin Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Stephan Preibisch, Curtis Rueden, Stephan Saalfeld, Benjamin Schmid, et al., "Fiji: an open-source platform for biological-image analysis," *Nature methods*, vol. 9, no. 7, pp. 676, 2012.

[4] Syed Hamad Shirazi, Arif Iqbal Umar, Nuhman Ul Haq, Saeeda Naz, Muhammad Imran Razzak, and Ahmad Zaib, "Extreme learning machine based microscopic red blood cells classification," *Cluster Computing*, , no. Suppl, pp. 1–11, 2017.

[5] M. Xu, D. P. Papageorgiou, S. Z. Abidi, M Dao, H. Zhao, and G. E. Karniadakis, "A deep convolutional neural network for classification of red blood cells in sickle cell anemia," *Plos Computational Biology*, vol. 13, no. 10, pp. e1005746, 2017.

[6] Yao Xue and Nilanjan Ray, "Cell detection with deep convolutional neural network and compressed sensing," *CoRR*, vol. abs/1708.03307, 2017.

[7] YM Hirimutugoda and Gamini Wijayarathna, "Image analysis system for detection of red cell disorders using artificial neural networks," *Sri Lanka Journal of Bio-Medical Informatics*, vol. 1, no. 1, 2010.

[8] H. A. Elsalamony, "Healthy and unhealthy red blood cell detection in human blood smears using neural networks," *Micron*, vol. 83, pp. 32–41, 2016.

[9] Mo Zhang, Xiang Li, Mengjia Xu, and Quanzheng Li, "Rbc semantic segmentation for sickle cell disease based on deformable u-net," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, Alejandro F. Frangi, Julia A. Schnabel, Christos Davatzikos, Carlos Alberola-López, and Gabor Fichtinger, Eds., Cham, 2018, pp. 695–702, Springer International Publishing.

Manuscript 644 submitted to 2019 IEEE International Symposium on Biomedical Imaging (ISBI). Received October 16, 2018.

[10] Weidi Xie, J. Alison Noble, and Andrew Zisserman, "Microscopy cell counting and detection with fully convolutional regression networks," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 6, no. 3, pp. 283–292, 2018.

[11] Khin Win, Somsak Choomchuay, Kazuhiko Hamamoto, and Manasanan Raveesunthornkiat, "Detection and classification of overlapping cell nuclei in cytology effusion images using a double-strategy random forest," *Applied Sciences*, vol. 8, no. 9, pp. 1608, Sep 2018.

[12] Gaobo Liang, Huichao Hong, Weifang Xie, and Lixin Zheng, "Combining convolutional neural network with recursive neural network for blood cell image classification," *IEEE Access*, vol. 6, pp. 36188–36197, 2018.

[13] Zhimin Gao, Lei Wang, Luping Zhou, and Jianjia Zhang, "Hep-2 cell image classification with deep convolutional neural networks," *IEEE journal of biomedical and health informatics*, vol. 21, no. 2, pp. 416–428, 2017.

[14] Sara Sabour, Nicholas Frosst, and Geoffrey E Hinton, "Dynamic routing between capsules," in *Advances in Neural Information Processing Systems 30*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds., pp. 3856–3866. Curran Associates, Inc., 2017.

[15] R. Chen, M. A. Jalal, L. Mihaylova, and R. K. Moore, "Learning capsules for vehicle logo recognition," in *2018 21st International Conference on Information Fusion (FUSION)*, July 2018, pp. 565–572.

[16] Dan Rosa de Jesus, Julian Cuevas, Wilson Rivera, and Silvia Crivelli, "Capsule networks for protein structure classification and prediction," *CoRR*, vol. abs/1808.07475, 2018.

[17] Tomas Iesmantas and Robertas Alzbutas, "Convolutional capsule network for classification of breast cancer histology images," in *Image Analysis and Recognition*, Aurélio Campilho, Fakhri Karray, and Bart ter Haar Romeny, Eds., Cham, 2018, pp. 853–860, Springer International Publishing.

[18] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in *Advances in Neural Information Processing Systems 28*, C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, Eds., pp. 91–99. Curran Associates, Inc., 2015.

[19] J. Deng, W. Dong, R. Socher, L. Li, Kai Li, and Li Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255.

[20] Jianqing Zhu, Shengcai Liao, Zhen Lei, and Stan Z Li, "Multi-label convolutional neural network based pedestrian attribute classification," *Image and Vision Computing*, vol. 58, pp. 224–229, 2017.

[21] Nikolaos Sarafianos, Theodoros Giannakopoulos, Christophoros Nikou, and Ioannis A. Kakadiaris, "Curriculum learning for multi-task classification of visual attributes," *CoRR*, vol. abs/1708.08728, 2017.

[22] Nicolas Coudray, Paolo Santiago Ocampo, Theodore Sakellaropoulos, Navneet Narula, Matija Snuderl, David Fenyö, Andre L Moreira, Narges Razavian, and Aristotelis Tsirigos, "Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning," *Nature medicine*, p. 1, 2018.

[23] Shin Hoo-Chang, Holger R Roth, Mingchen Gao, Le Lu, Ziyue Xu, Isabella Nogues, Jianhua Yao, Daniel Mollura, and Ronald M Summers, "Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics and transfer learning," *IEEE transactions on medical imaging*, vol. 35, no. 5, pp. 1285, 2016.

[24] Ha Tran Hong Phan, Ashnil Kumar, Jinman Kim, and Dagan Feng, "Transfer learning of a convolutional neural network for hep-2 cell image classification," in *IEEE International Symposium on Biomedical Imaging*, 2016, pp. 1208–1211.

[25] Phillip M Cheng and Harshawn S Malhi, "Transfer learning with convolutional neural networks for classification of abdominal ultrasound images," *Journal of digital imaging*, vol. 30, no. 2, pp. 234–243, 2017.

[26] Sulaiman Vesal, Nishant Ravikumar, AmirAbbas Davari, Stephan Ellmann, and Andreas Maier, "Classification of breast cancer histology images using transfer learning," in *Image Analysis and Recognition*, Aurélio Campilho, Fakhri Karray, and Bart ter Haar Romeny, Eds., Cham, 2018, pp. 812–819, Springer International Publishing.

[27] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[28] Diederik P. Kingma and Jimmy Ba, "Adam: A method for stochastic optimization," *CoRR*, vol. abs/1412.6980, 2014.

[29] Jerome H. Friedman, "Greedy function approximation: A gradient boosting machine," *The Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001.