PROPOSAL

As a standard practice in clinical medicine, microscopic examination of peripheral blood plays an important role in the field of diagnosis and control of major diseases. It is uniquely capable of discerning clinically relevant morphologic features of hematopoietic cells, including abnormal white blood cells (WBCs, also known as leukocytes) in lymphoma, leukemia, dysplasia and other diseases. The gold-standard morphologic profiling of blood cells relies heavily on manual smear processing techniques and visual inspection. Blood smear preparation and interpretation are thought to be negatively affected by observer bias and also involve labor-intensive processes that require highly skilled technologists. As such, there has been considerable interest in developing systems for automated classification of digital images of peripheral blood smears with high sensitivity and specificity. We do not expect a popular demand for image based automated peripheral blood and bone marrow smear examination device in near future. The reasons are

1) Data obtained from Flowcytometry based blood cell analyzers preclude the necessity of morphological examination in about 80% of cases and also provide valuable morphological clues in vast majority of the remaining cases.
2) Microscopic examination for malaria parasites is being routinely done on thick smear using a special stain. This method is highly sensitive and takes less time. Moreover, rapid strip based malaria antigen tests are highly sensitive and specific, affordable, and are routinely being used in remote areas.

However, there are serious lacunae in the existing system of blood smear examination. The most important issue is identification of abnormal cells that has immense impact on the diagnosis, management and outcome of patients. Early identification of these cells can avoid delay in the diagnosis of potentially fatal diseases and completely reverse the outcome. Unfortunately, due to lack of proper training and lack of access to expert opinion in most part of the world, these abnormalities are overlooked and under reported. Another important issue is the critical analysis of red blood cell morphology, which not only provide important clues for identification of the cause of anemia and avoid unnecessary expensive investigations, but also can provide clues for the diagnosis of many non-hematological systemic diseases.

In recent years researchers have made efforts to automate the process using machine learning methods. There are three broad types of approaches:

1) Traditional feature engineering involving image segmentation, feature extraction, feature selection and finally classification. Some of the well-known classification techniques.
2) Object classification approach involving patch-generation followed by (convolutional neural network) CNN-based classification.
3) Object detection approach involving CNN-based architecture that will identify the location of objects in the image and will also predict their classes.

Proposed research work

To develop web based classification system for abnormal cells in peripheral blood smear by analyzing microphotographs taken by cell-phone camera with an intent to differentiate between normal and abnormal, and specific or group classification if abnormal cell found.

Data collection.

In order to perform this study, we will establish the peripheral leukocyte micro-images database that is labeled with ground truth boxes by three experts. All annotated color images will have same size, and include 11 categories of leukocytes, i.e., blast (Blast), promyelocyte (PRO), myelocyte (MYE), metamyelocyte (MET), band neutrophil (bNEU), segmented neutrophil (sNEU), lymphocyte (LYM), monocyte (MO), eosinophil (EO), basophil (BA) and reactive lymphocyte (rLYM). We will also try to include four non-WBC types, i.e., nucleated red blood cell (NRBC), giant platelet (gPLT), smudge (SM) cell and artefact (Artefact).

Cell detection method.

The first part of the automatic cell detection system is to address a user-friendly annotation tool so that Doctors/Technicians/trained students can do the annotation job easily. Labelling of the cells of interest has to be automatic and this requires implementation of a segmentation algorithm. Figure 1 shows four sample input images (input to the proposed system). Input images are actually of same size. Figure 2 shows automatic detection of the cells of interest. Cells are to be marked either by (i) best fit regular shape, i.e., circle/ellipse or (ii) identification of irregular elastic boundary [?? Research Q: which one should we follow?]

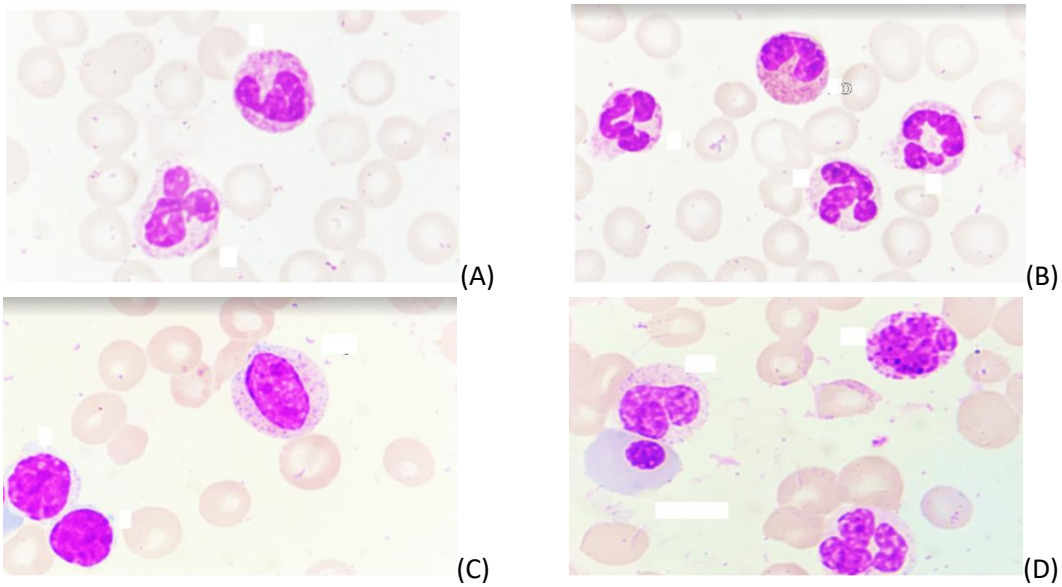

Figure 1: Input Images

Once the ROIs [region of interest] are marked on an image, option must be giving so that manual correction is possible. The system may make two types of errors and suitable manual

correction technique must be provided to correct system errors: (i) The system may completely miss detection of a cell: in such cases, someone can tap on the ROI so that system gets input to know where the cell is and then it automatically draws the boundary (ii) The system has detected the cell but made mistake in marking the boundary properly: annotating App should have facility so that someone can alter the marked boundary by changing it using fingertip.

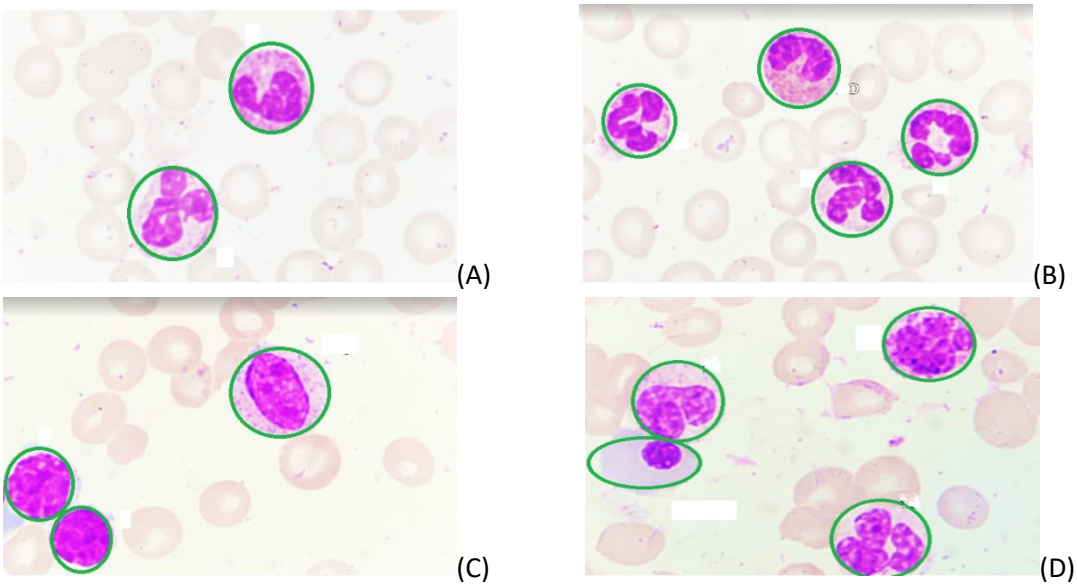

Figure 2: Labelling of cells and their boundaries.

After the right cell is detected and its boundary is marked, a check-list is popped up so that the annotator checks the right radio button to annotate the cell type (one out of 15 types as identified by Dr. Banerjee). Figure 3 illustrates this facility. For an image all this information is stored [boundary information and corresponding cell type] for future use.

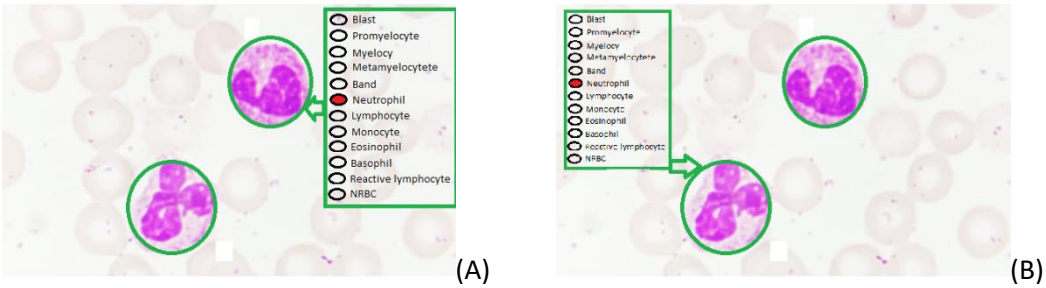

Figure 3: Annotation of Cell type

Caution: If system often fails to detect cells and their boundaries, annotation task will involve more human effort for labelling cell boundaries too. I am hopeful that with today's learning systems automatic cell segmentation wouldn't be bad. So we can be hopeful.

Cell type annotation:

Phase-I: Initially, cells will be labelled (through radio button facility) by annotators (Figure 3). Once the system gets enough labelled data (let's say, 50 instances per cell type totaling 750 labelled instances), it will learn how to tag cells automatically. Note that for collecting 750 labelled instances, we may not need 750 many images as an image may contain several instances of same or different cells. Maybe, we need around 400 images to achieve this annotation task.

Phase-II: At this point onward, the annotators need not mark the cell type manually but verify the cell type tagged by the system. Hence, the annotating App will provide facility where an annotator can alter the label of a cell if the system has made mistake in labelling. Let's assume Phase-II would verify (and thereby collecting) annotation of another 750 new instances. During this phase, we can have some idea about the machine accuracy for cell type detection. Also, during both phases, we will have idea about the system's accuracy in detecting cells and their boundaries. If the system performs poorly in this section, we can fix it during Phase-I so that Phase-II does not suffer from this problem.

Classification method to be used during Phase-II: Since Phase-I would give us 50 samples per cell type; we would implement a CNN based classification to get preliminary results. From my experience I hope current deep systems won't make us disappointed.

After we finish Phase-II, we will have annotation of more 750 instances (roughly 50 instances per classes) and with this we can retrain the system to improve its accuracy. This retraining can be done in batches during Phase-II. Also, we can consider incremental training during Phase-II and some master students can burn their fingers in trying such alternative approaches. But finally, we will have a system trained on 1500 instances (approx. 100 samples/cell type).

System evaluation: As specificity and sensitivity are two standard metrics for evaluating such medical image processing systems, we will follow these metrics to report the system accuracy.

Machine learning methods

In many medical application domains, the UNET [UNET] architecture yields the state-of-the-art segmentation results. This architecture comprises a contracting path to capture context and a symmetric expanding path that enables localization. Several variants of the UNET have been proposed in the literature that have been found to deliver improved performance. The UNet++ [UNET++] architecture has nested and dense skip connections, potentially making it more capable of capturing fine-grained details by gradually enriching the high-resolution feature maps from the encoder network, and then fusing it with the feature maps from the decoder network. The work also evaluated the UNet++ architecture on the nuclei segmentation task in microscopic images. [ATTENTION-U-NET] implemented attention gate in a standard U-Net architecture that allows the attention coefficient to be more specific to local regions. Medical images can be ambiguous and a group of annotators can produce

a set of diverse but plausible segmentations. The recently proposed Probabilistic UNET [P-UNET] architecture, a generative segmentation model combining UNET and variational autoencoder, can assist in clinical decision making by generating multiple plausible segmentation hypotheses to be used for possible diagnoses and suggest actions for resolving ambiguities. [MULTI-SCALE] proposes a different type of architecture where a multi-scale strategy combines semantic information at different levels and self-attention modules to progressively aggregate relevant contextual features. Some of the other promising medical image segmentation architectures includes Recurrent Residual CNN based on UNet [R2UNET] (this architecture has been used for nuclei segmentation in [NUCLEI-R2UNET]), Attention based Semi-supervised Deep Networks [ASD-NET].

Some research works have already attempted to address the problem of WBC differential based on images using ML methods. For example, in [PERIPHERAL-LUKOCYTE] used object detection techniques, namely Single Shot Multibox Detector (SSD) [SSD] and You Only Look Once (YOLOv3) [YOLOv3], for 11 types of peripheral leukocyte recognition. A few of the other well-known deep learning based object detection architectures include Faster R-CNN [FASTER-R-CNN], Mask R-CNN [MASK-R-CNN], Region-based Fully Convolutional Networks (R-FCN) [R-FCN]:. Some of these frameworks could be adapted for the purpose of leukocyte detection. The initial objective of this project would be to develop standard models for segmentation, object detection and classification based on medical image datasets on specific problems, e.g. leukocyte recognition, assessment of dysplasia [DYSPLASIA].

The developed models need to be robust to challenges arising from real-world medical datasets. One of the key issues pertaining to the performance of such models in arbitrary medical environments is the inter-laboratory variability in the data -- the models are often “over-trained” to perform well on a dataset originating from some particular lab and may not have broad applicability required of a clinical algorithm [TRANSLATIONAL-AI]. This issue may be addressed to certain extent by training the model on data from multiple laboratories (or at least data representative of such variability). But the models themselves need to be inherently robust such that they account for the uncertainty in the model parameters as well as the noise in the data. Most deep learning algorithms cannot provide reliable estimates of uncertainty. New methods approximating Bayesian inference in deep networks can make predictions as well as produce estimates of uncertainty [BAYESIAN-DEEP]. Recent works combining Bayesian methods with deep networks have shown that methods based on stochastic regularization techniques can capture meaningful uncertainty, and at the same time scale well to high-dimensional data. Methods in the framework of Bayesian deep learning have already been used for classification [BAYESIAN-DEEP], segmentation [SEG-UN], object detection [OBJ-DETECT-UN] – these are the key tasks in this particular project. If for a certain image the model is able to generate high level of uncertainty along with the erroneous prediction, then the system can recommend human intervention, and can avoid danger to human lives. A Bayesian deep learning model for semantic segmentation will generate predictions for each pixel, and also provide pixel-wise uncertainty estimates [SEG-UN]. In the context of object detection, the uncertainty should be indicative of detection accuracy, where a higher spatial or label uncertainty is correlated with inaccuracy in localisation or semantic classification [OBJ-DETECT-1].

REFERENCES:

[U-NET]: Ronneberger et. al. "U-Net: Convolutional Networks for Biomedical Image Segmentation," MICCAI 2015.

[UNET++]: Zhou et. al. "UNet++: A nested U-Net Architecture for Medical Image Segmentation," Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Springer, Cham, 2018. 3-11

[ATTENTION-U-NET]: Schlemper et. al. "Attention gated networks: Learning to leverage salient regions in medical images," Medical Image Analysis, 53: 197-207, 2019.

[P-UNET]: Kohl et. al. "A Probabilistic U-Net for segmentation of ambiguous images," NIPS 2018.

[MULTI-SCALE]: Sinha, A., and Jose D. "Multi-scale self-guided attention for medical image segmentation." IEEE Journal of Biomedical and Health Informatics 2020.

[R2UNET]: Md Zahangir, A., et al. "Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation." 2018.

[NUCLEI-R2UNET]: Md Zahangir, A. et al. "Nuclei segmentation with recurrent residual convolutional neural networks based U-Net (R2U-Net)." *NAECON 2018-IEEE National Aerospace and Electronics Conference*. IEEE, 2018.

[ASD-NET]: Dong, N. et. al. "ASDNet: Attention based semi-supervised deep networks for medical image segmentation," International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2018.

[PERIPHERAL-LUKOCYTE]: Wang, Q., et al. "Deep learning approach to peripheral leukocyte recognition." *Plos one* 14.6 (2019)

[YOLOv3]: Redmon, J., and Ali F. "Yolov3: An incremental improvement." *arXiv preprint arXiv:1804.02767* (2018).

[SSD]: Liu, W., et al. "Ssd: Single shot multibox detector." *European conference on computer vision*. Springer, Cham, 2016.

[FASTER-R-CNN]: Ren, S., et al. "Faster r-cnn: Towards real-time object detection with region proposal networks." *IEEE transactions on pattern analysis and machine intelligence* 39.6 (2016): 1137-1149.

[MASK-R-CNN]: He, K., et al. "Mask r-cnn." *Proceedings of the IEEE international conference on computer vision*. 2017.

[R-FCN]: Dai J, Li Y, He K, Sun J. RFCN: object detection via region based fully convolutional networks. NIPS. 2016:379–387.

[DYSPLASIA]: Mori, Jinichi, et al. "Assessment of dysplasia in bone marrow smear with convolutional neural network." *Scientific reports* 10.1 (2020): 1-8.

[TRANSLATIONAL-AI]: Serag, A., et al. "Translational AI and deep learning in diagnostic pathology." *Frontiers in Medicine* 6 (2019).

[BAYESIAN-DEEP]: Gal, Yarin, and Zoubin Ghahramani. "Dropout as a bayesian approximation: Representing model uncertainty in deep learning." *international conference on machine learning*. 2016

[OBJ-DETECT-UN]: Miller, Dimity, et al. "Evaluating merging strategies for sampling-based uncertainty techniques in object detection." *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019.

[SEG-UN]: Mukhoti, Jishnu, and Yarin Gal. "Evaluating bayesian deep learning methods for semantic segmentation." *arXiv preprint arXiv:1811.12709* (2018).

[OBJ-DETECT-UN-1]: Miller, Dimity, et al. "Benchmarking Sampling-based Probabilistic Object Detectors." *CVPR Workshops*. Vol. 3. 2019.