**Clustering-Based Dual Deep Learning Architecture for Detecting Red Blood Cells in Malaria Diagnostic Smears** (IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS, VOL. 25, NO. 5, MAY 2021)

With the use of a dual deep learning architecture, we are able to detect red blood cells in thin blood smears, RBCNet (name of architecture) uses cell clustering instead of region proposals. Cell detection accuracy was higher than 97%. Most malaria screening algorithms begin with finding foreground masks using thresholding methods (Otsu’s method, kmeans, adaptive histogram thresholding, Zack thresholding), the following are different techniques used to separate touching cells; distance transform, watershed, morphological operations and active contours, these are arguably the main actors in the difficult process of cell detection. These methods can work quite well however encounter problems or fail when they encounter challenging conditions due to the fact that they depend on certain criteria such as intensity or energy. These methods can also work quite well when used in combination with smaller data sets however with large scale data and large-scale imaging (higher resolutions) they begin to produce inconsistent results and deviate more.

The motivation behind the development of the dual deep learning architecture is to be able to reduce the computationally expensive cost of object detection networks on large thin smear microscopy. This is done to achieve rapid fully automated malaria diagnosis.

Methods such as Watershed or Active contour suffer from higher Standard Deviation by up to 6 times more than the standard deviation produced by the method proposed by this paper. The projects pipeline works consistently well for all regions of the images, whereas other methods such as Active contour suffer from under or over-segmentation, which happens in regions of images with challenging light conditions or low contrast. This is why the U-Net architecture proves to be very influential in production higher quality results regardless of different conditions.

The training process consists of including a U-Net architecture with connected component labeling for extracting foreground masks from the cell images. The following stage uses a faster R-CNN architecture for refining cell clusters into individual cells with accurate boundaries. The U-Net architecture basically guides the detection process in the inference stage to provide the CNN with robust candidates, this creates a solution to images with large amounts of cells and dense clusters of cells.

The data set consists of thin blood smear images from human patients, these slides were photographed by a **smartphone camera.**

Conclusively the U-Net architecture is most influential in providing a robust foreground mask for CNNs to train from and therefore with the combination of U-Net and Faster R-CNN, a robust solution for detecting red blood cells in thin blood smears is achieved and therefore outperforms tradition cell detection methods, object detection deep learning methods and object detection deep learning methods. This pipeline is compatible with large scale data processing and inconsistent data imaging as it is able to produce quality results under challenging conditions.

**Invasion of Red Blood Cells by Malaria Parasites** (Alan F. Cowman1, \* and Brendan S. Crabb1, \* 1The Walter and Eliza Hall Institute of Medical Research, Melbourne 3050, Australia \*Contact: cowman@wehi.edu.au (A.F.C.); crabb@wehi.edu.au (B.S.C.) DOI 10.1016/j.cell.2006.02.006)

In humans, the most severe form of malaria is caused by Plasmodium falciparum, and at least one-third of the world’s population is at risk of infection, with over 300 million people developing clinical disease each year and at least 2 million deaths (Snow et al., 2005). Other Plasmodium spp. including P. vivax infect humans and cause considerable morbidity in those populations where it is endemic. The overall toll caused by malaria has changed little since the 1950s, when the first mass initiatives based on elimination of the mosquito vector and antimalarial prophylaxis were enacted. While there have been major reductions in mortality and morbidity in some areas, such as South Asia, malaria remains a major pediatric killer in many parts of sub-Saharan Africa, which bears the greatest burden of disease.

Malaria has plagued humans for many thousands of years, and, despite attempts to eradicate it through concerted programs of antimalarial drug treatment as well as vector control, the causative agent has adapted to these challenges. Major efforts are underway to develop new antimalarial drugs and a vaccine.

**Estimating malaria parasitemia from blood smear images** (Halim et al, Proceedings of the International Conference on Control, Automation, Robotics and Vision, Jan 2006)

Estimating parasitemia from blood smears accounting for uncertain imaging conditions, the method consist of using a multi-stage estimation process which is based on pattern matching with parameter optimization and cross-validation against expected biological characteristics. With the use of a training-based method they were able to achieve 92% and 95% precision and recall rates.

By using the following optimization approach the samples are able to be evaluated before the investigation of the infection status is carried out.

Segmentation techniques used by previous technology: thresholding, morphology and transformations and dedicated data description models.

The following are the problems related to techniques used in previous technology. In regard to thresholding, this does not account for uncertainty of the samples provided and therefore correct thresholds cannot be set in advance. In regard to applying morphological operations, the method suffers due to different cell shape variations and therefore cannot be correctly segmented. For model-based approaches, with the use of Bayes classifiers or regularized centroid transform, it is one of the most effective methods however suffers from the problem that if the data derivates from the model, it limits the applicability of the classifier. In summary these techniques are dependent for static pre-defined parameters, and they suffer if the data is not optimal.

In order for classification to be successful, the technique of compensation of image variability was used to achieve a reliable performance and achieves the actual classification of infected cells to be more stable. This allows us to take inspiration from this process in order to increase the accuracy of our predictions and have a more optimized classifier.

**An automatic device for detection and classification of malaria parasite species in thick blood film** (Kaewkamnerd et al, BMC Bioinformatics, Volume 13, 2012)

Accurate classification of parasite species suffers from inherent technical limitation of human inconsistency, with the development of image processing tools aimed at detection and classification of parasites in thick blood film this problem can be addressed as the limitation of detecting and processing parasites in thin blood film would be overcome as most image processing tools focus on thin blood film.

With the use of tools aimed at thick blood film, the classification of the Pv (P. vivax, causes less serious malaria) parasite was 75% and the accuracy of the Pf (P. falciparum, causes severe malaria) parasite was 90%. The accuracy of the prototype in regard to detecting parasite-positive and parasite-negative blood film, the rate was measured at 95% and 68.5% accuracy respectively.

Conclusively this method can be used for both detection and classification of malaria on thick blood film, the analysis program could accurately classify parasite species, into Pf or Pv, and into positive or negative malaria samples.

**Automated image processing method for the diagnosis and classification of malaria on thin blood smears** (N. E. Ross and C. J. Pritchard and D. M. Rubin and A. G. Dus\'e, Medical and Biological Engineering and Computing, Volume 44, Number 5, 2006)

The use of thin blood smears is advantageous as there is no loss of parasite during staining, the red blood cells are fixed, imaging artifacts are uncommon and most importantly, the parasite morphology is not distorted. The examination of the slides was carried out under oil-immersion, and under a light microscope with the magnification of 1000x.

The method itself examines the red blood cell and determines if malaria is present or not, and if present it determines the species of malaria (P. falciparum, P. vivax, P. ovale, P. malarie). The system has sensitivity as sometimes low number of parasites can be found which does not provide a good result. The algorithm is an image classification problem, therefore takes the form of standard pattern recognition and classification. It consists of four stages: image acquisition, pre-processing, feature generation and classification. A morphological method is used to identify the malaria in the stained blood slides.

Pre-processing removes unwanted effects like noise from the image and resolution is scaled down by a factor of four to speed up the system, the image is filtered by a 5x5 median filter followed by a morphological area closing filter using a disk shape structure with the radius of 6 pixels.

Granulometry is being used to find out the size distribution of the cells which allows us to find out the amount of cells in the image, and an average pattern spectrum is used to determine the sample magnitude and standard deviation of the cell radii.

For image segmentation, thresholding is primarily being used (rather than edge detection, and watershed method) and with the use of an image histogram the threshold selection can be used. Two threshold levels need to be determined from the histogram; one for the red blood cells and one for the parasites. A common problem experienced with global threshold selection is that the thresholding is not sensitive enough. The solution is to find local threshold levels in the red blood cells taking the minimum threshold in the histogram of the red blood cells. Therefore, both local and global thresholds are used.

For classification, a tree classifier with two nodes using BFF neural networks determines whether or not a cell is infected, and if so, the species of the malaria. Potential parasites are segmented from the image with a sensitivity of 92% and a PPV of 40%. The classifier is able to positively identify malaria parasites with a sensitivity of 85% and a PPV of 81%.