Cascading YOLO: Automated Malaria Parasite Detection for *Plasmodium Vivax* in Thin Blood Smears

Feng Yang[a], Nicolas Quizon[a], Hang Yu[a], Kamolrat Silamut[b], Richard J Maude[b,c,d], Stefan Jaeger[a], Sameer Antani[a]

[a]National Library of Medicine, National Institutes of Health, Bethesda, MD 20894, USA; [b]Mahidol-Oxford Tropical Medicine Research Unit, Bangkok, Thailand; [c]Centre for Tropical Medicine and Global Health, Nuffield Dept of Medicine, University of Oxford, Oxford, UK; [d]Harvard TH Chan School of Public Health, Harvard University, Boston, USA

ABSTRACT

Malaria, caused by *Plasmodium* parasites, continues to be a major burden on global health. *Plasmodium falciparum* (*P. falciparum*) and *Plasmodium vivax* (*P. vivax*) pose the greatest health threat among the five malaria species. Microscopy examination is considered as the gold standard for malaria diagnosis, but it requires a significant amount of time and expertise. In particular, the detection of *P. vivax* is more difficult due to the lower parasitemia levels as compared to *P. falciparum*. In this work, we develop a rapid and robust diagnosis system for the automated detection of *P. vivax* parasites using a cascaded YOLO model. This system consists of a YOLOv2 model and a classifier for hard-negative mining. Results from 2567 thin blood smear images of 171 patients show the cascaded YOLO model improves the mean average precision about 8% compared to the conventional YOLOv2 model.

Keywords: Computer-Aided Diagnosis, Deep Learning, YOLO, *Plasmodium vivax*, Malaria

1. INTRODUCTION

Malaria continues to be a major burden on global health, causing approximately half a million fatalities each year[1]. Malaria in humans is caused by parasites of five *Plasmodium* species, of which *Plasmodium falciparum* (*P. falciparum*) and *Plasmodium vivax* (*P. vivax*) pose the greatest health threat. Microscopy examination of peripheral blood smears is considered as the gold standard for malaria diagnosis[2]. However, it requires a significant amount of time and expertise. The objective of this work is to develop an automated diagnostic system for *P. vivax* malaria. Accurate automated detection of *P. vivax* is made difficult due to the lower parasitemia levels typically observed in these infections[3], as compared to *P. falciparum*. This requires processing of large numbers of images and/or imaging of regions of thin blood smears that are dense with cells and challenging to segment. Examination of thick blood smears allows more parasites to be seen in fewer images but image analysis and speciation are far more difficult due to the high cell density.

In recent years, some approaches have been reported for *P. falciparum* parasite detection in both thin and thick blood smears, and reviews of the published literatures may be found in [4–9]. However, only a few works are reported on the identification/detection of *P. vivax*. Yunda et al.[10] apply a gradient-technique-based segmentation in thick blood smear images to extract background from objects, and then extract wavelet-based features of the objects followed by a neural network classifier for parasite species identification. Evaluation on 112 images shows that the neural network classifier achieved a specificity of 71.6% for *P. vivax* identification. Ghosh et al.[11] first crop thin blood smear images into 1000×800 pixels, and then apply a fuzzy divergence based segmentation method on *P. vivax* infected region(s). Based on the pre-segmentation objects, the authors extract four textural features (average intensity, skewness, uniformity and entropy) and a feature called fractal calculation and apply Support Vector Machine (SVM) and Bayesian classifiers for *P. vivax* identification[12]. Their classification model on cells from 100 normal patients and 100 infected patients achieves an accuracy of 95% for SVM classifier and 98% for Bayesian classifier respectively. Gitonga et al.[13] apply an Artificial Neural Network (ANN) on segmented cells to classify the parasite species in thin blood smears images. The authors report that their ANN classifier achieves an accuracy of 96.2% in recognizing *P. vivax* from other *plasmodium* species on 205 infected images. Maysanjaya et al.[14] first manually crop parasite infected regions of 250×250 pixels from acquired thin blood smear images, and then segment *P. vivax* parasites by using color channel combination and Otsu method. Evaluation on 30 infected region images shows the method achieves an accuracy of 93.33%. Penas et al.[15] first perform a parasite segmentation through connected components analysis, and then apply a Convolutional Neural Network (CNN) for parasite

species classification in thin blood smears. Evaluation on 363 images shows that the classifier has achieved an accuracy of 87.9% for parasite species classification.

We give in Table 1 an overview of the existing detection approaches for *P. vivax* parasites in thin and thick blood smears. In summary, the existing methods either use traditional image processing techniques for *P. vivax* segmentation or work directly on manually cropped *P. vivax* infected region images. Therefore, most of the quantitative experimental evaluations are performed for the classification stage instead of the segmentation stage. In addition, the evaluations are based on cell-level; that is, the input sample is a single cell image and the evaluation is typically a cell classification accuracy. However, the ultimate goal for malaria patient diagnosis is to detect and classify all cells (both parasites and false positives) for a patient. A satisfying cell-level classification performance does not necessarily assure good performance to patient level.

Our main contribution in this paper is two-fold: First, we propose a cascaded YOLO model consisting of a YOLOv2 model and an AlexNet classifier for hard-negative mining to reduce the false positive errors of the parasite detection. The automated detection model achieves a mean average precision of 79% for *P. vivax* detection on 171 patients. To the best of our knowledge, we are the first to detect *P. vivax* parasites using deep learning methods. Second, this is the first work to evaluate the performance of *P. vivax* detection on patient-level. We organize the rest of the paper as follows: Section 2 presents the details of our method. Section 3 introduces the dataset and the experimental results. In Section 4, we discuss our results and conclude the work.

Table 1. Existing approaches applied to *P. vivax* detection in thin and thick blood smears.

Authors	Methods	Images	Patients	Thick/Thin smear	Results
Yunda et al., 2012[10]	Perform a segmentation to extract background from objects, and then extract wavelet-based features of objects for the following classification of different parasite species.	112	-	Thick	71.6% specificity for classification
Ghosh et al., 2013[11,12]	Apply a fuzzy divergence based segmentation method on *P. vivax* infected region(s), then extract four textural features and a fractal calculation feature and apply SVM/Bayesian classifier to identify *P. vivax* parasites from normal erythrocytes.	-	200	Thin	95% accuracy for SVM classifier and 98% accuracy for Bayesian classifier
Gitonga et al., 2014[13]	Apply an ANN on segmented cells for parasite species classification.	205	-	Thin	96.2% accuracy for ANN classifier
Maysanjaya et al., 2016[14]	Apply Otsu threshold to the combination of color channels (red and saturation channels) to segment P. vivax.	30	-	Thin	93% accuracy on manually cropped ROI images
Penas et al., 2017[15]	Perform a parasite segmentation through connected components analysis and apply a CNN model for parasite species classification.	363	-	Thin	87.9% accuracy for classification

2. METHODS

2.1 YOLO model

The You Only Look Once (YOLO)[16] model was introduced to create a one-step processing for simultaneous object detection and classification. It uses a single CNN operating directly on an image and outputting bounding box coordinates and class probabilities. It is fast enough for real-time object detection. The idea of YOLO differs from two-stage object detection models, such as RCNN[17], Fast-RCNN[18], Faster-RCNN[19], in that bounding box prediction and class prediction are performed simultaneously. An input image is first divided into a $S \times S$ grid, and then B bounding boxes are predicted in each grid cell, along with confidence scores for those boxes. Simultaneously, C conditional class probabilities are also

predicted. These probabilities are conditioned on the grid cells containing an object, and only one set of class probabilities is predicted per grid cell.

The YOLO model has been improved in different versions. YOLOv2[20] makes a large improvement in the reduction of localization errors and increase of mean Average Precision (mAP) [21,22]. YOLOv3[23] improves the model for large-scale and multi-label object detection. In our work, we choose YOLOv2 instead of YOLOv3 since YOLOv2 achieves better mAP on our dataset.

2.2 Cascaded YOLO model

We developed a rapid and robust diagnosis system for the automated detection of *P. vivax* parasites using a cascaded YOLO model. This system consists of a YOLOv2 model and a transferred AlexNet classifier for hard-negative mining. We illustrate the flowchart of our cascaded YOLO model in Fig. 1(a). YOLOv2 model is used to get a rough detection of *P. vivax* parasites that may include false positives, while the transferred AlexNet classifier is used to reduce false positives.
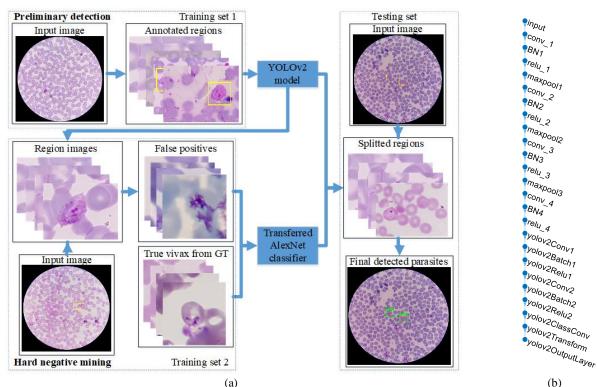

Figure 1. (a) Flowchart of the proposed cascaded YOLO model. GT indicates ground-truth. (b) Network structure of the YOLOv2 model in (a). Note that conv indicates a convolutional layer, BN indicates a batch normalization layer, and relu represents a rectified linear unit (ReLU).

For training, we first split images of 4032×3024×3 pixels into regions of 672×504×3 pixels. Manual ground-truth annotations are also mapped to corresponding region space. Based on the regions along with annotations (Training set 1 in Fig. 1(a)), we train a YOLOv2 model, whose network structure is listed in Fig. 1(b). The trained YOLOv2 model is then applied to new region images (Training set 2 in Fig. 1(a)), and bounding boxes are predicted for parasites. By comparing the bounding boxes with ground-truth annotations, we can get the false positives generated by the YOLOv2 model. These false positives are considered as hard-negatives since they are too similar to *P. vivax*-infected regions to be detected correctly. In order to reduce false positive errors, we train an AlexNet classifier with those hard-negatives and annotated parasites from ground truth. For testing, we also split blood smear images into regions of 672×504×3 pixels, which are then screened for parasites using our cascaded model of YOLOv2 and AlexNet classifier. The detected parasite coordinates are finally re-projected to the original image space for visualization and evaluation.

3. EXPERIMENTS AND RESULTS

3.1 Data acquisition

We photographed Giemsa-stained thin blood smears from 171 *P. vivax* infected patients in Bangkok, Thailand, using a smartphone camera that we attached to the eyepiece of a microscope. We captured 2567 images with 100x magnification in RGB color space with a 3024×4032 pixel resolution. An expert blood smear reader manually annotated each image at the Mahidol-Oxford Tropical Medicine Research Unit (MORU), Bangkok, Thailand. We de-identified all images and their annotations, and archived them at the National Library of Medicine (NIH IRB#12972).

3.2 Data Partitioning

We split a dataset of 171 patients into ten folds at patient-level. Each fold includes 17 *P. vivax* infected patients. The system performance is evaluated with ten-fold cross validation. For each run, the YOLOv2 model is first trained on region images of 119 patients, then an AlexNet classifier for hard-negative mining is trained on 34 patients, and finally the cascaded YOLO model is evaluated on the 17 patients of the test fold.

3.3 Evaluation Metric

To evaluate the system performance, we calculate the average precision (AP)[21,22], which summarizes the shape of the precision-recall curve and which is defined as the mean precision at a set of eleven equally spaced recall levels $r \in [0, 0.1, \dots, 1]$:

$$AP = \frac{1}{11}\Sigma_{r\in[0,0.1,\dots,1]} p_{interp}(r). \tag{1}$$

To compute $p_{interp}(r)$, the precision at each recall level r is interpolated by taking the maximum precision measured for our method for which the corresponding recall exceeds r:

$$p_{interp}(r) = \max\big(p(\bar{r})\big), \bar{r}:\bar{r} \geq r, \tag{2}$$

where $p(\bar{r})$ is the measured precision at recall $\bar{r}$.

3.4 Experimental results

We perform experiments on thin blood smear images from 171 patients with ten-fold cross validation. The Precision-Recall curve for the cascaded YOLO model with ten-fold cross validation is shown in Fig. 2. Mean average precision (mAP) values of the cascaded YOLO model and of the conventional YOLOv2 model are listed in Table 2. We observe that the cascaded YOLO model improves the mAP value by about 8% compared to the conventional YOLOv2 model. Fig. 3 shows a practical example of *P. vivax* parasite detection using the conventional YOLOv2 model and our proposed cascaded YOLO model. The number of false positives generated by YOLOv2 has been effectively reduced by cascading the AlexNet classifier trained on the mined hard-negatives.
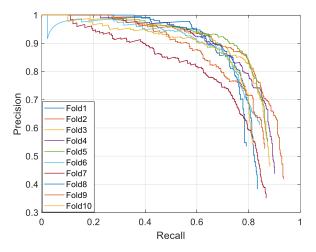

Figure 2. Precision-Recall curve for the cascaded YOLO model with ten-fold cross validation.

Table 2. Mean average precision for the cascaded YOLO model and the plain YOLOv2 model.

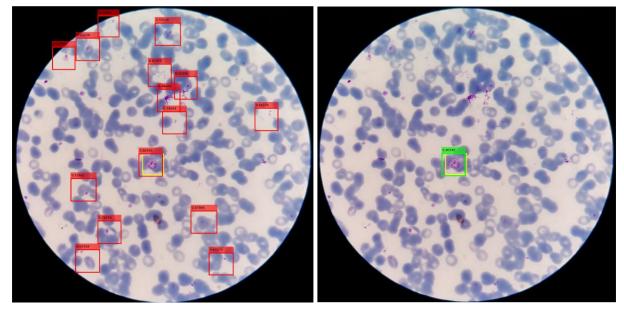| Model	mean Average Precision (mAP)	Std of Average Precision
YOLOv2	71.34%	3.38%
Cascaded YOLO	79.22%	4.15%

Figure 3. An example of *P. vivax* parasite detection using the conventional YOLOv2 model (left) and the cascaded YOLO model (right). Yellow bounding boxes are the ground truth. Red boxes indicate *P. vivax* parasites detected by the conventional YOLOv2 model, including false positives. Green boxes indicate the detected *P. vivax* parasites detected by our cascaded YOLO model.

4. DISCUSSION AND CONCLUSION

In this work, we develop an automated system for *P. vivax* detection based on a cascaded model of YOLOv2 and AlexNet classifier for hard-negative mining. In Section 3, we show that, by cascading a hard-negative mining classifier, the mean average precision on patient-level has been improved by about 8%.

In our cascaded YOLO model, YOLOv2 model achieves better results than YOLOv3 model. This may be due to two factors. First, YOLOv3 both extracts features in multi-scales and predicts bounding boxes in different scales, and thus is better for detection of objects in large scales. However, in our case, the size of *P. vivax* parasites stays in a very narrow range (about 200×200×3 pixels). Second, YOLOv3 improves the average precision for multi-label object detection, which is not our case.

In our experiments, we split images of 4032×3024×3 pixels into regions of 672×504×3 pixels for training the YOLOv2 model. We have also compared the performance of splitting images into regions of 504×378×3 pixels and of 288×216×3 pixels. We observe that YOLOv2 model achieves best detection results when the object size is about 10-15% of the input image. In this case, detailed features can still be preserved after down-sampling process with the network.

Our system can be used in resource-limited regions without the need for specific malaria expertise. It also provides a reliable and standardized interpretation of thin blood smears. Future work will improve the average precision by integrating the model in one network structure and refining the loss function.

ACKNOWLEDGEMENT

REFERENCES

[1]   World Health Organisation., [World malaria report 2018], World Health Organization (2018).

[2]   World Health Organization., [Guidelines For The Treatment of Malaria, Third edition], World Health Organization (2015).

[3]   Moreira, C. M., Abo-Shehada, M., Price, R. N. and Drakeley, C. J., "A systematic review of sub-microscopic Plasmodium vivax infection," Malar. J. **14**(1), 1–10 (2015).

[4]   Poostchi, M., Silamut, K., Maude, R. J., Jaeger, S. and Thoma, G., "Image analysis and machine learning for detecting malaria," Transl. Res. **194**, 36–55 (2018).

[5]   Pattanaik, P. A. and Swarnkar, T., "Comparative analysis of morphological techniques for malaria detection," Int. J. Healthc. Inf. Syst. Informatics **13**(4), 1–17 (2018).

[6]   Rosado, L., M. Correia da Costa, J., Elias, D. and S. Cardoso, J., "A Review of Automatic Malaria Parasites Detection and Segmentation in Microscopic Images," Anti-Infective Agents **14**(1), 11–22 (2016).

[7]   Liang, Z., Powell, A., Ersoy, I., Poostchi, M., Silamut, K., Palaniappan, K., Guo, P., Hossain, M. A., Sameer, A., Maude, R. J., Huang, J. X., Jaeger, S. and Thoma, G., "CNN-based image analysis for malaria diagnosis," Proc. - 2016 IEEE Int. Conf. Bioinforma. Biomed. (BIBM 2016), 493–496 (2016).

[8]   Poostchi, M., Ersoy, I., McMenamin, K., Gordon, E., Palaniappan, N., Pierce, S., Maude, R. J., Bansal, A., Srinivasan, P., Miller, L., Palaniappan, K., Thoma, G. and Jaeger, S., "Malaria parasite detection and cell counting for human and mouse using thin blood smear microscopy," J. Med. Imaging **5**(4), 1 (2018).

[9]   Yang, F., Poostchi, M., Yu, H., Zhou, Z., Silamut, K., Yu, J., Maude, R. J., Jaeger, S. and Antani, S., "Deep Learning for Smartphone-based Malaria Parasite Detection in Thick Blood Smears," IEEE J. Biomed. Heal. Informatics, in press. (2019).

[10]   Yunda, L., Alarcón, A. and Millán, J., "Automated Image Analysis Method for p-vivax Malaria Parasite Detection in Thick Film Blood Images," Sist. y Telemática **10**(20), 9 (2012).

[11]   Ghosh, M., Das, D., Chakraborty, C. and Ray, A. K., "Plasmodium vivax segmentation using modified fuzzy divergence," Proc. 2011 Int. Conf. Image Inf. Process. (ICIIP 2011) (2011).

[12]   Ghosh, M., Das, D., Chakraborty, C. and Ray, A. K., "Quantitative characterisation of Plasmodium vivax in infected erythrocytes: a textural approach," Int. J. Artif. Intell. Soft Comput. **3**(3), 203 (2013).

[13]   Gitonga, L., Maitethia Memeu, D., Amiga Kaduki, K., Allen Christopher Kale, M. and Samson Muriuki, N., "Determination of Plasmodium Parasite Life Stages and Species in Images of Thin Blood Smears Using Artificial Neural Network," Open J. Clin. Diagnostics **4**(2), 78–88 (2014).

[14]   Maysanjaya, I. M. D., Nugroho, H. A., Setiawan, N. A. and Murhandarwati, E. E. H., "Segmentation of Plasmodium vivax phase on digital microscopic images of thin blood films using colour channel combination and Otsu method," Proc. - AIP Conf. **1755** (2016).

[15]   Penas, K. E. D., Rivera, P. T. and Naval, P. C., "Malaria Parasite Detection and Species Identification on Thin Blood Smears Using a Convolutional Neural Network," Proc. - 2017 IEEE 2nd Int. Conf. Connect. Heal. Appl. Syst. Eng. Technol. CHASE 2017, 1–6 (2017).

[16]   Redmon, J., Divvala, S., Girshick, R. and Farhadi, A., "You only look once: Unified, real-time object detection," Proc. - IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., 779–788 (2016).

[17]   Girshick, R., Donahue, J., Darrell, T. and Malik, J., "Rich feature hierarchies for accurate object detection and semantic segmentation," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., 580–587 (2014).

[18]   Girshick, R., "Fast R-CNN," Proc. IEEE Int. Conf. Comput. Vis., 1440–1448 (2015).

[19]   Ren, S., He, K., Girshick, R. and Sun, J., "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," IEEE Trans. Pattern Anal. Mach. Intell. **39**(6), 1137–1149 (2017).

[20]   Redmon, J. and Farhadi, A., "YOLO9000: Better, faster, stronger," Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognit., 6517–6525 (2017).

[21]   Everingham, M., Van Gool, L., Williams, C. K. I., Winn, J. and Zisserman, A., "The pascal visual object classes (VOC) challenge," Int. J. Comput. Vis. **88**(2), 303–338 (2010).

[22]   Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P. and Zitnick, C. L., "Microsoft COCO: Common objects in context," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) **LNCS 8693**(PART 5), 740–755 (2014).

[23]   Redmon, J. and Farhadi, A., "YOLO v.3" (2018).