DEEP LEARNING TOWARDS EFFICIENT MALARIA DATASET CREATION

Martha Shaka[1], Frederick Apina[1], Nyamos Waigama[1], Halidi Maneno[1], Said H Said[1], Emilian Ngatunga[1], Said Mmaka[1], Simon Chaula[1], Imani Sulutya[1], and Merikiadi Mashaka[2].*
1 UDOM 2 BENJAMIN MKAPA HOSPTAL
**marthashaka@gmail.com*

INTRODUCTION

- Malaria remains one of the major threats to public health and economic development in Africa [1].

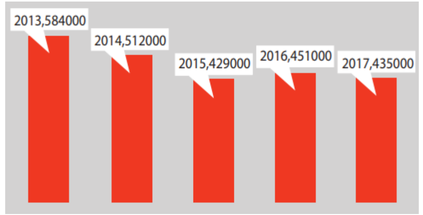

Figure 1: Showing the death caused by Malaria [1]

- Prompt diagnosis and treatment of malaria reduces deaths.
- Microscope is the golden standard of malaria diagnosis and poses the following challenges:-
  - ✓Requires expert, consumes more time, expensive and error prone process.
- Several AI techniques have been adopted to solve these challenges [2].
- However, in the context of developing countries, there is a shortage of data for research and the development of such tools.
- This work pushes deep learning towards efficient malaria dataset creation in solving challenges in malaria diagnosis.

DATASET CREATION

No of patients	No of parasites	No white blood cells
20 +ve cases	47512	6376
20 -ve cases	-	6683

Table 1: Summary of the dataset and its distribution of parasites and white blood cell count.

- Data distribution ratio for model development

✓ **70:15:15**

METHODOLOGY

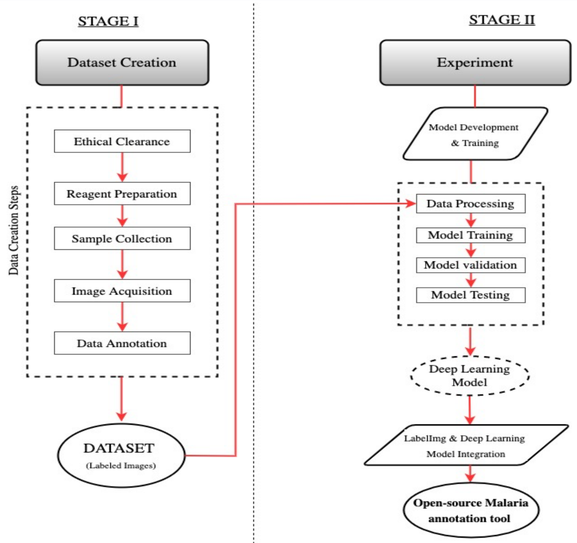

Figure 2: Methodology flow chart to summarize the two major stages of this work

- A team of three lab technologist collected and annotated the images all together.
- The images were annotated using the LabelImg annotation tool.
- We used Detectron2 architecture[3] and transfer learning to train our model.

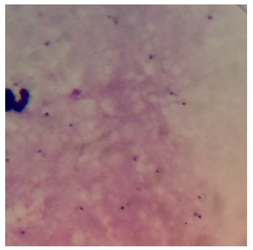

Figure 3: Captured image

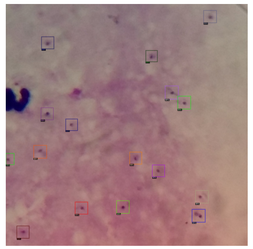

Figure 4: Annotated image

EXPERIMENTAL RESULTS

- We calculated the average precision (AP) for each class individually across all of the **IoU thresholds** (0.5 and increasing to 0.95 in steps of .05)
- it took one third less time to detect and count the plasmodium compared to having a person perform manual counting

Backbone	AP on Plasmodium	AP_{50} on Plasmodium	AP_{75} on Plasmodium	AP on WBC	AP_{50} on WBC	AP_{75} on WBC
R50-FPN	71.76	83.76	68.45	89.88	93.55	87.01
R101-FPN	79.06	88.03	70.33	92.20	95.01	90.67

Table 2: showing the results of the models with the metric is broken out by object class

CONCLUSION

- the proposed annotation tool can create the labelled dataset seven times faster than lab technologists.
- The future work will include using the tool to create more malaria dataset to improve the deep learning model and eventually integrate a mobile application as diagnostic tool.

REFERENCES

[1] WHO malaria 2018 report. Retrieved on 1st March 2019 from https://apps.who.int/iris/bitstream/handle/10665/275867/9789241565653-eng.pdf?ua=1

[2] J.A. Quinn, R. Nakasi, P.K. Mugagga, P. Byanyima, W. Lubega, A. Andama. Deep Convolutional Neural Networks for Microscopy-Based Point of Care Diagnostics. Proceedings of the International Conference on Machine Learning for Health Care, Journal of Machine Learning Research W&C track, Volume 56, 2016

[3] Wu Y, Massa F, Kirillov A, Wan-Yen Lo & Ross Girshick (2019). Detectron2: A PyTorch-based modular object detection library. Retrieved on Jan 2019 from https://ai.facebook.com/blog/-detectron2-a-pytorch-based-modular-object-detection-library-/

ACKNOWLEDGEMENT