

Design, Development,
Analysis and Performance
Evaluation of Malaria
Parasite Detection
Deep Learning
Models.

Table of Contents

Literature Review	5
Data Collection and Processing	6
Model Building	7
Model Evaluation and Comparison.....	14
Discussion and Conclusion	16
<u>References</u>	17

Introduction

Malaria is a parasitic infection caused by plasmodium- a specie of the protozoan parasite and transmitted through the bite of an infected female Anopheles mosquito (WHO, 2016). It is prevalent in the tropical and subtropical regions of the world and is the singular most common cause of death in these regions. To quantify this, malaria is responsible for over 400,000 deaths and 200 million infections globally (Poostchi, et al., 2018). To combat and subsequently eradicate these parasitic diseases, so much is being invested in developing new tools, however, these tools are not all accessible to national malaria programs (Landier, et al., 2016).

Despite the efforts of researchers and governments from around the world the burden of malaria on global health has remained on the rise. There are now cases of resistance to common and affordable malaria treatments, and urbanizing trends like immigration and population growth in most regions have also led to the re-introduction of malaria to places where they have hitherto been eradicated or not in existence (Erdman & Kain, 2008).

The main challenge facing the reduction in malaria mortality rate and successful treatment of the disease is the accurate diagnosis of the disease which is currently done by examining blood samples for parasitaemia (WHO, 2016). Consequently, fast and accurate diagnostics are therefore imperative malaria control strategies. One of the state-of-the-art methods for fast and reliable detection of malaria parasitic infection is through the use of a deep learning model known as Convolutional Neural Networks(CNN) which has a high accuracy rate in classifying infected and uninfected blood samples and is easily reproducible.

This paper looks at the design, development, analysis, and performance of CNN Deep models and makes a comparison with other models built on some existing pre-trained models such as Xception, VGG19, and ResNet50 to determine the best model for classifying parasitic malaria images.

Literature Review

Due to the limitations inherent in the use of manual methods for the detection of malaria-infected cells, efforts have been made toward overcoming this challenge through automatic means, earlier studies, however, focused on the automatic quantification of parasitic cells rather than automatic detection. The MalariaCount software automatically analyzes images of Giemsa-stained Red Blood Cell (RBC) smears and generates accurate parasitaemia values that are comparative to results obtained from manual counts (Sio, et al., 2007), the use of unsupervised methods in quantifying malaria parasitaemia by splitting of infected and uninfected cells obtained from blood smear images (Halim, et al., 2006).

Many recent studies deployed deep learning models in malaria detection as they are easily reproducible and outperform machine learning models (Rajaraman, et al., 2018). A 16-layer Convolutional Neural Network was used to distinguish infected from uninfected cells in a thin blood smear, after applying ten-fold cross-validation (Liang, et al., 2016). This model was able to achieve an average accuracy of 97.4%. Gopakumar, et al., (2018) applied a focus stack of blood smear images to a custom-built convolutional neural network using 2 level segmentation strategy to resolve the cell counting limitation. This approach enhanced the rate of infection detection, recording sensitivity, and specificity of 97% and 98% respectively.

Malaria being a global health burden, the estimated global investment target required to achieve a minimum of 90% reduction in infection and death rates from 2016 to 2030 stands at \$22.8billion (Patouillard, et al., 2017). To meet the targeted reduction in mortality rate, ground-breaking diagnostic methods have to be developed through research. Current and future trends in medical image processing look at the application of different deep learning techniques for early detection and image classification (Razzak, et al., 2018), the same approach has been adopted in this report for analyzing cell images for the detection of malaria.

Data Collection and Processing

The data used for this project consists of 27,560 cell images of different sizes in Portable Network Graphics (PNG) format divided equally into parasitized (13,780) and uninfected (13,780) cell images and downloaded from the Kaggle website, an online community for data scientists([Malaria Cell Images Dataset | Kaggle](#)).

The original images were collected using a Malaria screener, a smartphone application developed by researchers at the Lister Hill National Centre for Biomedical Communication alongside the National Institutes of Health (NIH)'s National Institute of Allergy and Infectious Diseases and Mahidol-Oxford University. (National Library of Medicine, n.d.)

Data Preprocessing

Data preprocessing started with uploading of cell images to the work environment, followed by reading, resizing and labeling of the images as (parasitized =0, uninfected =1) using necessary relevant libraries. All images were resized to 200x200pixels to make it easy for the model to learn from a standardized image size, treated as colored-RGB, and split into training and test data in the ratio of 7:3.

Model Building.

Convolution Neural Network(CNN)

CNN is a deep learning model that cannot only process image data but performs image recognition tasks with high levels of accuracy. The network consists mainly of three distinct layers; convolution, pooling, and the fully connected layer. Feature extraction of the input image data is handled by the convolving and pooling layers while the classification of the extracted feature is performed by the fully connected layer (Yamashita, et al., 2018)

CNN was used as the base model for this project. Firstly, a customized sequential CNN model was built from scratch and trained using the preprocessed malaria data for image classification and prediction followed by leveraging the capabilities of some pre-trained models (Xception, VGG19, and ResNet50) using transfer learning techniques for classification and prediction.

1) Custom CNN Model

The custom-built CNN model has 3 convolution layers, 3 pooling layers, and 2 fully connected layers. The input layer consists of 32 filters and a 3x3 kernel with a 200x200pixels RGB image fed to it as input data.

The second and third convolution layers have 64 filters, 3x3 kernel, and 128 filters, 3x3 kernels respectively. A Rectified Linear Unit (RELU) activation function $f(x)=\max(0, x)$, was applied to the outputs of each convolution layer, followed by a (2x2, 2) max pooling and a 20% dropout.

The first fully connected layer has 512 neurons with a 50% dropout. The output from the first fully connected layer is fed to a second fully connected layer having 2 neurons and a sigmoid classifier. The model was compiled by applying categorical –cross-entropy as loss function and Adam as optimizer. The model performance is detailed below.



Figure 1. model performance on the train and test data in terms of accuracy and loss.

Column1	Precision	Recall	F1-Score	Support
Uninfected	0.98	0.88	0.93	4146
Parasitized	0.89	0.98	0.94	4122
Weighted Average	0.94	0.93	0.93	8268

Figure 2. performance measures of the custom-built module with an **accuracy of 0.93**

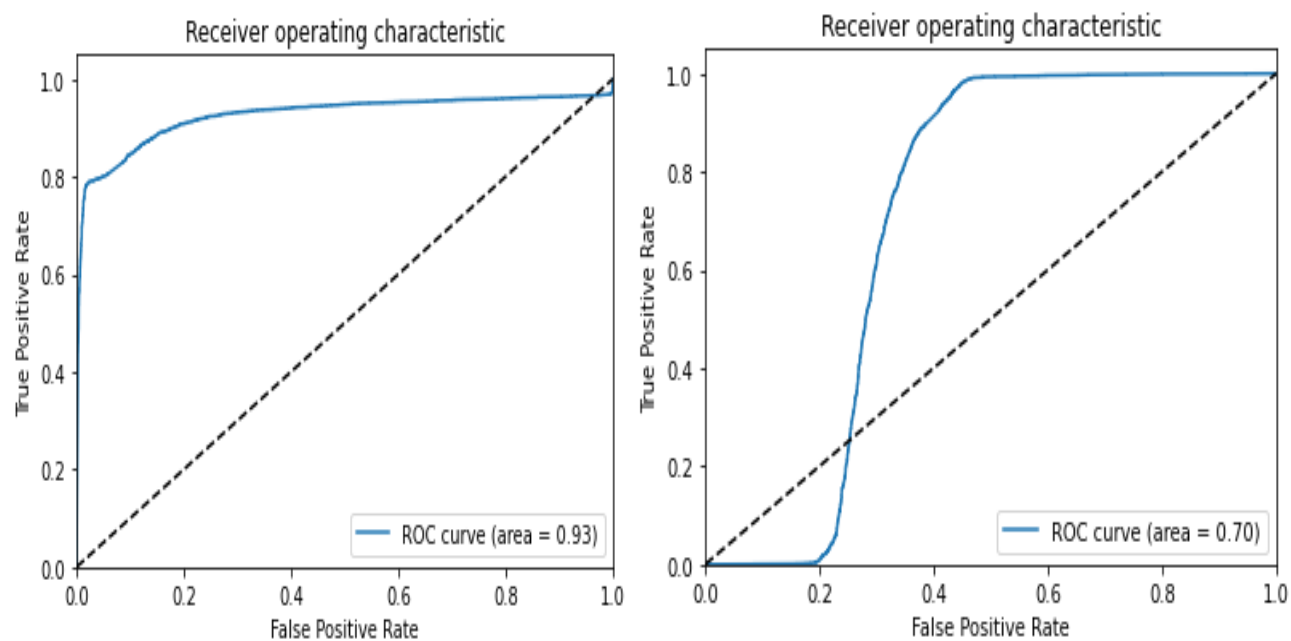


Figure 3. AUC -Receiver operating characteristic of *the custom-built model*

2) *Pre-trained Models*

The technique of transfer learning was explored; this was done by leveraging the extractive capabilities of some of the best existing pre-trained models through fine-tuning the models by replacing the input and last layers of the model and training the new model using our data.

The top and bottom layers of the pre-trained models were removed, with the top layer replaced with an input shape of (200, 200, 3) and the bottom layer replaced with the dense layer with only 2 neurons. The pre-trained weights of the layers of the base –model were also retained by setting base model training = false.

3) *Xception*

The Xception model was put forward by Francois Chollet and developed by Google, it is a convolution neural network with depthwise separable convolutions with 71 deep layers (Chollet, 2017) and is capable of classifying 1000 object categories.

The performance of the new model with Xception as the base model is detailed below with overall accuracy of 0.82

	Precision	Recall	F1-Score	Support
Uninfected	0.8	0.85	0.83	4146
Parasitized	0.84	0.78	0.81	4122
Weighted Average	0.82	0.82	0.82	8268

Figure 4. Performance measures of the Xception base model

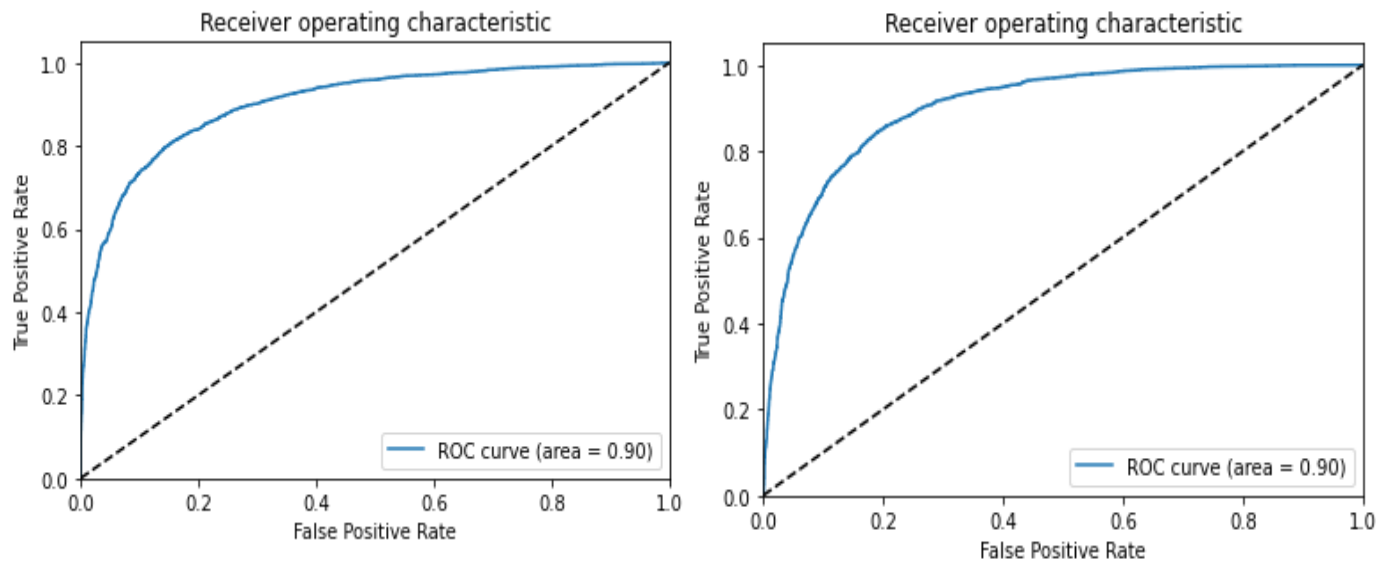


Figure 5. AUC -Receiver operating characteristic of the Xception base model

4) VGG19

The VGG19 was developed by the Visual Geometry Group at Oxford University in 2014, it is made up of 16 convolution layers, 3 fully connected layers, 5 max-pooling, and a single softmax layer (Simonyan & Zisserman, 2014).

The performance of the new model with VGG19 as the base model is detailed below with an accuracy of 0.94

	Precision	Recall	F1-Score	Support
Uninfected	0.92	0.96	0.94	4146
Parasitized	0.96	0.91	0.93	4122
Weighted Average	0.94	0.94	0.94	8268

Figure 6. Performance measures of the VGG19 base model

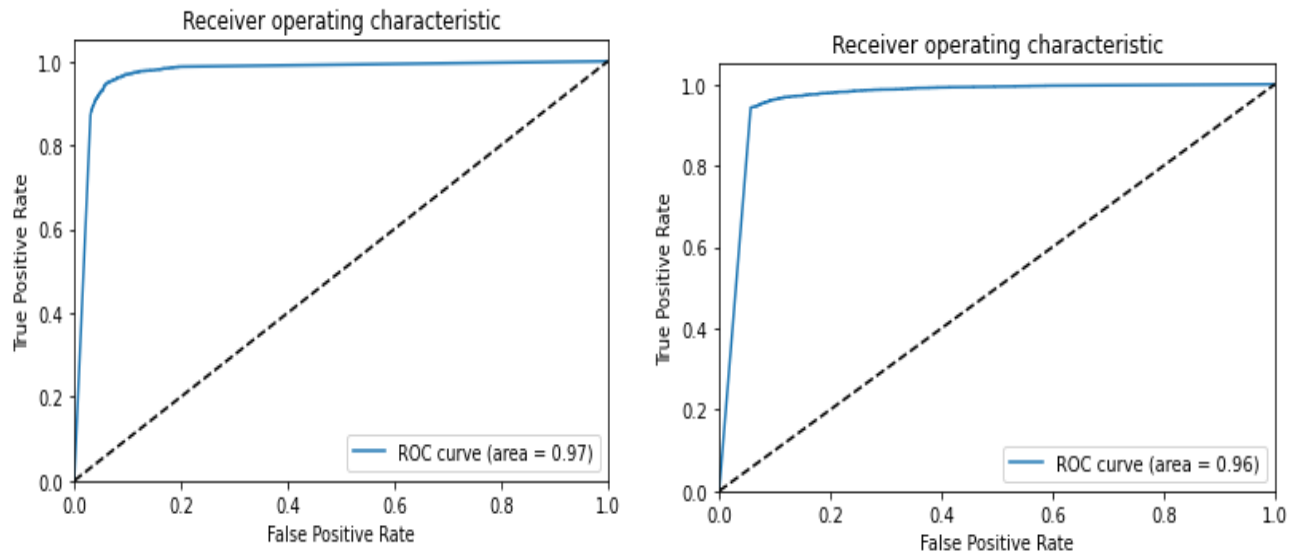


Figure 7. AUC -Receiver operating characteristic of the VGG19 base model

5) ResNet50

The Residual Network50 (ResNet50) is a CNN that has 50 deep layers and was proposed in a research paper by He, et al., (2016). The model earned the top position in the ILSVRC and COCO competition in 2015 (He, et al., 2016)

The summary and performance of the new model with ResNet50 as the base model are detailed below. The model achieved an accuracy of 0.94

Column1	Precision	Recall	F1-Score	Support
Uninfected	0.94	0.94	0.94	4146
Parasitized	0.94	0.94	0.94	4122
Weighted Average	0.94	0.94	0.94	8268

Figure 8. Performance measures of the ResNet50 base model

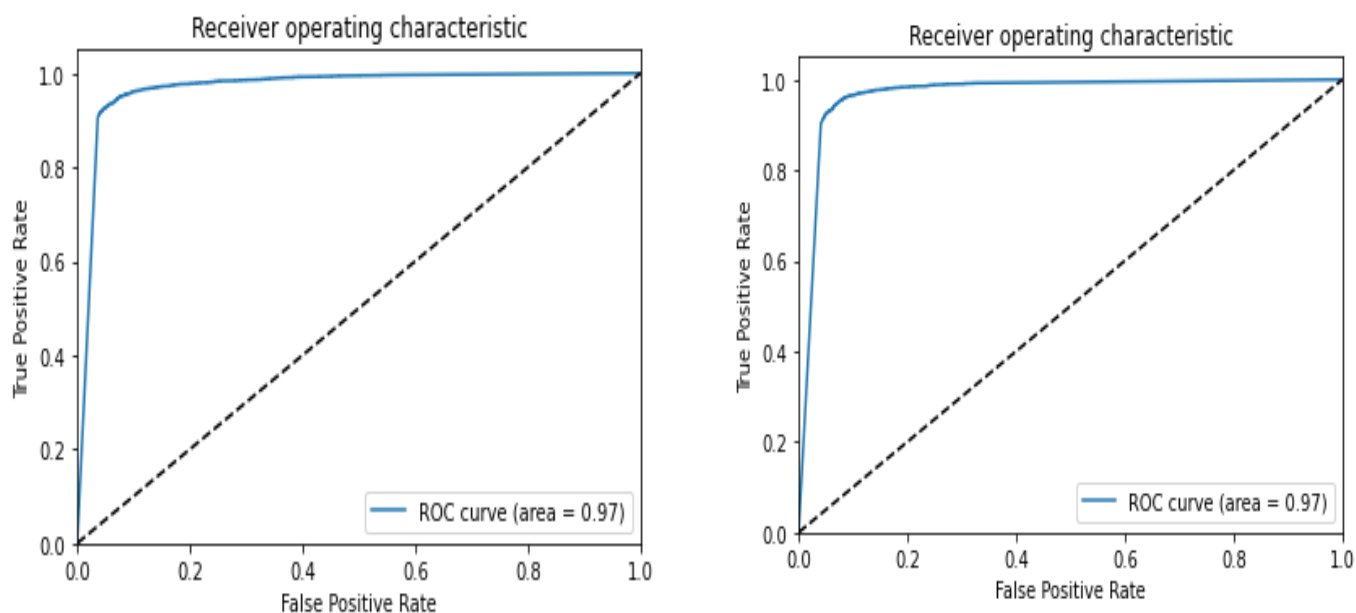


Figure 9. AUC -Receiver operating characteristic of the ResNet50 base model

Model Evaluation and Comparison

This project aims to build a deep learning model that will accurately classify malaria cell images, to achieve this, four CNN models were built, the first model was custom-built, and the others leveraged on pre-trained models. Here the performance metrics obtained for each of the models are analyzed to determine the best model for the project.

Models	Precision	Recall	F1-Score	AUC	Accuracy
Custom	0.9350	0.9300	0.9350	0.8150	0.9300
Xception	0.8200	0.8150	0.8200	0.9000	0.8200
VGG19	0.9400	0.9350	0.9350	0.9650	0.9400
ResNet50	0.9400	0.9400	0.9400	0.9700	0.9400

Figure 10. Performance measures comparison for the four models

Models	Class	Precision	Recall	F1-Score	AUC	Accuracy
Custom	Uninfected	0.9800	0.8800	0.9300	0.9300	-
	Parasitized	0.8900	0.9800	0.9400	0.7000	-
Macro Avg.		0.9350	0.9300	0.9350	0.8150	0.9300
Xception	Uninfected	0.8000	0.8500	0.8300	0.9000	-
	Parasitized	0.8400	0.7800	0.8100	0.9000	-
Macro Avg.		0.8200	0.8150	0.8200	0.9000	0.8200
VGG19	Uninfected	0.9200	0.9600	0.9400	0.9600	-
	Parasitized	0.9600	0.9100	0.9300	0.9700	-
Macro Avg.		0.9400	0.9350	0.9350	0.9650	0.9400
ResNet50	Uninfected	0.9400	0.9400	0.9400	0.9700	-
	Parasitized	0.9400	0.9400	0.9400	0.9700	-
Macro Avg.		0.9400	0.9400	0.9400	0.9700	0.9400

Figure 11. Performance measures for parasitized and uninfected classes

From the performance table above, it is clear that the VGG19 and ResNet50 models performed better across all the parameters. The custom-built model also performed well with an accuracy of 93% however, it recorded a very low precision for the parasitized class, an indication that the model could not significantly identify or predict the positive class correctly. This is also reflected by the very low value of the Area Under the Curve (AUC) of the Receiver Operating Characteristics Curve.

The VGG19 model had a better precision for predicting parasitized cells compared to uninfected cells and recorded a good average across all parameters. However, accuracy is not the sole determiner of how well a model is performing, for a project such as this, where the risk of error is high, the desire is also to have a model with a very low False Negative rate. It is therefore imperative that look at all the metrics before selecting the preferred model.

To conclude, the ResNet50 model outperformed the other models in predicting both parasitized and uninfected cells and recorded the best accuracy and AUC-ROC values.

Discussion and Conclusion

This project seeks to develop a deep learning model that will automatically and accurately diagnose an infected RBC. To achieve this, the malaria dataset from NLM was used to train the models. First, a custom CNN model was designed with 3-convolution, 3 max-pooling, and 2 fully connected layers with batch normalization introduced to standardize the inputs to each subsequent layer. A dropout of 20% was added to each layer and 50% to the first fully connected layer to reduce overfitting and improve accuracy. A final accuracy of 0.93 was achieved after many adjustments to the parameters. Three other models were built leveraging on capabilities of already existing award-winning models (Xception, VGG19, and ResNet50). The performance of all the models was evaluated, collected, and analyzed side by side to determine the model with the best performance. While the custom-built model on the average performed well, the precision for predicting infected cells was low, and also the AUC value. The Xception model though a well-regarded model in the field only had an accuracy of .82, the lowest. The VGG19 model performed comparatively well across all metrics there is a need for a good balance across all metrics, for this project, the sensitivity(recall) metric is very important. The desire is to adopt a model with the lowest false negatives, which translates to a higher recall value. The ResNet50 model proved to be the most efficient across all metrics and was adopted as the best for this project

References

- Chollet, F., 2017. *Xception: Deep learning with depthwise separable convolutions*. s.l., s.n., pp. 1251-1258.
- Erdman, L. K. & Kain, K. C., 2008. Molecular diagnostic & surveillance tools for global malaria control. *Travel medicine and infectious disease*, 6(1-2), pp. 82-99.
- Gopakumar, G. et al., 2018. Convolutional neural network-based malaria diagnosis from focus stack of blood smear images acquired using custom-built slide scanner. *Journal of biophotonics*, 11(3), p. e201700003.
- Halim, S. et al., 2006. *Estimating Malaria Parasitaemia from Blood Smear Images*. s.l., s.n., pp. 1-6.
- Hawkes, M., Katsuva, J. P. & Masumbuko, C. K., 2009. Use and limitations of malaria rapid diagnostic testing by community health workers in war-torn Democratic Republic of Congo. *Malaria Journal*, 8(1), p. 308.
- He, K., Zhang, X., Ren, S. & Sun, J., 2016. *Deep residual learning for image recognition*. s.l., s.n., pp. 770-778.
- Hommelsheim, C. et al., 2014. PCR amplification of repetitive DNA: a limitation to genome editing technologies and many other applications. *Scientific Reports*, Volume 4, p. 5052.
- Hung, J. & Carpenter, A., 2017. *Applying Faster R-CNN for Object Detection on Malaria Image*. s.l., s.n., pp. 808-813.
- Landier, J. et al., 2016. The role of early detection and treatment in malaria elimination. *Malaria Journal*, 15(1), pp. 1-8.
- Liang, Z. et al., 2016. *CNN-based image analysis for malaria diagnosis*. s.l., s.n., pp. 493-496.
- National Library of Medicine, n.d. *National Library of Medicine*. [Online]
Available at: <https://lhncbc.nlm.nih.gov/LHC-downloads/downloads.html#malaria-datasets>
[Accessed 21 May 2022].
- Patouillard, E. et al., 2017. Global investment targets for malaria control and elimination between 2016 and 2030. *BMJ Global Health*, 2(2), p. e000176.
- Poostchi, M. et al., 2018. Image analysis and machine learning for detecting malaria. *Translational Research*, Volume 194, pp. 36-55.
- Rajaraman, S. et al., 2018. Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images. *PeerJ*, Volume 6, p. e4568.
- Razzak, M. I., Naz, S. & Zaib, A., 2018. Deep learning for medical image processing: Overview, challenges and the future. *Classification in BioApps*, pp. 323-350.
- Ross, N. E., Pritchard, C. J., Rubin, D. M. & Dusé, A. G., 2006. Automated image processing method for the diagnosis and classification of malaria on thin blood smears. *Medical and Biological Engineering and Computing*, 44(5), pp. 427-436.
- Simonyan, K. & Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Sio, S. W. S. et al., 2007. MalariaCount: An image analysis-based program for the accurate determination of parasitemia. *Journal of Microbiological Methods*, 68(1), pp. 11-18.
- WHO, 2016. *Malaria microscopy quality assurance manual-version 2*. s.l.:World Health Organization.

Yamashita, R., Nishio, M., Do, R. K. G. & Togashi, K., 2018. Convolutional neural networks: an overview and application in radiology. *Insights into Imaging*, 9(4), pp. 611-629.