MalariaNet: A Computationally Efficient Convolutional Neural Network Architecture for Automated Malaria Detection

Rohan Bhansali
Senior at Loudoun Academy of Science
Director at ConnectAI
Ashburn, United States

Rahul Kumar
Senior at Loudoun Academy of Science
Director at the Aluna Research Group
Ashburn, United States

Abstract— **Despite much progress in detection and treatment, malaria remains one of the most prevalent diseases on earth, both in terms of incidence and death rate. Multiple studies have shown that early detection of malaria is paramount to preventing fatal outcomes; however, current testing methods have notable issues involving cost and accessibility. As a result, deep learning algorithms have been developed for malaria detection and have achieved state of the art results in rapid diagnosis; however, it has been noted that the computational expense of running elaborate models makes deep learning based detection methods inaccessible in remote areas of the world. We develop a computationally efficient, relatively shallow neural network architecture that can diagnose malaria from cell images obtained from thin blood smear slides. Specifically, our algorithm, dubbed MalariaNet, is a 7-layer convolutional neural network trained using the Adaptive Moment Estimation algorithm on the open source NIH malaria dataset, containing 27,588 images of parasitized and uninfected cells. We report that MalariaNet achieves an accuracy of 0.968, F1 score of 0.955, precision of 0.946, and recall of 0.974. We hope that our computationally considerate model inspires more research in producing accessible artificial intelligence solutions for disease detection tasks.**

Keywords—Deep learning, neural network, malaria, cells

I. INTRODUCTION

Malaria is a mosquito-borne infectious disease that is caused by single-celled, parasitic microorganisms of the Plasmodium genus and is typically spread via the bite of infected female Anopheles mosquitoes, as the parasites from the saliva of the mosquito are transferred to the blood of the affected human [5]. Subsequently, the parasites transfer to the liver where they are able to efficiently develop and reproduce. The disease manifests itself as fever, tiredness, vomiting, and headaches; in severe cases, it can also result in yellow skin, seizures, coma, and even death. These symptoms typically present themselves ten to fifteen days after infection, but recurrence can occur months later. Although medicinal research has made strides in malaria prevention, detection, and alleviation, malaria remains the leading cause of death in the world and at this time, there exists no effective vaccination. Most cases occur in tropical and subtropical regions, particularly Sub-Saharan Africa, Asia, and Latin America. Overall, malaria was estimated to affect 228 million people and cause an estimated 405,000 deaths in 2018, with 94% of these fatalities occurring in Africa [4]. Consequently, the disease is estimated to cost Africa over $12 billion due to healthcare costs, decreased workforce, and negative effects on tourism.
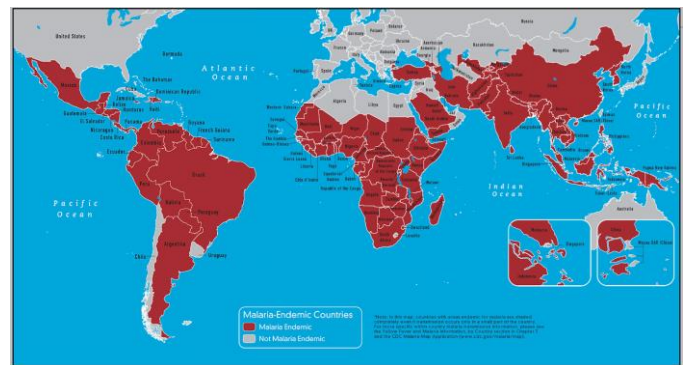

Fig. 1. Map of malaria-endemic countries [4]

The most widely used methods for malaria diagnosis fall into two categories: direct and indirect [17]. Direct methods of detection rely upon confirming a diagnosis based upon the discovery of parasitic bodies or parts of parasitic bodies. Indirect methods involve the detection of antibodies known to be relevant agents in the diagnosis of malaria. The advantages and disadvantages of the primary direct and indirect malaria diagnostic test methods were recorded by Talapko, et al. and are re-expressed in Table 1 and Table 2.

TABLE I.

Direct Methods	Advantages	Disadvantages
Microscopic analysis	Quick and inexpensive	Requires high-level experience and training
Rapid diagnostic tests	Quick and simple to perform	Low sensitivity and expensive
Molecular tests	Highly sensitive and accurate	Expensive and time-consuming in a large fraction of cases

TABLE II.

Indirect Methods	Advantages	Disadvantages
Indirect Immunofluorescence	High specificity and sensitivity	Time-consuming to perform and subjective evaluation of results
ELISA	Correct determination of type, highly specific	Time-consuming to perform and expensive

As seen in both Table 1 and Table 2, the current methods of malaria diagnosis all feature a combination of issues in price, accessibility, evaluation of results, or low levels of sensitivity.

According to the CDC, microscopic analysis of potentially infected cells remains the gold standard in malaria diagnosis [4]. However, they note that the quality and accuracy of the test is highly dependent on the experience of the laboratorian conducting the diagnostic. As a result, malaria detection through microscopic analysis is inaccessible to millions around the world since there are limited numbers of experienced laboratorians in many African and Asian countries. For these reasons, despite being one of the cheapest and quickest methods of malarial diagnosis, microscopic analysis is still unattainable for the majority of at-risk individuals across the world.

A. Previous Work

Numerous studies in the past have explored the use of artificial intelligence based techniques for malaria related tasks. For example, one group utilized feature extractors implemented through pretrained deep convolutional neural network models in order to classify parasitized and non-parasitized malaria cells from blood smears [13]. Another group developed a five step image analysis system that drew upon various machine learning systems for optimized malaria detection from blood smears. Other groups have drawn upon powerful deep learning architectures for malaria detection. Specifically, models such as ResNet-50, DenseNet-121, Xception, VGG-16, and other elaborate architectures have been used for malaria classification, and have achieved state of the art classification accuracies [6, 9]. Finally, a recent study utilized a deep convolutional neural network in conjunction with cyclical stochastic gradient descent and automated learning rate optimization for malaria detection [11].

Outside of classification, other work has been done in using deep learning methodologies for segmentation tasks. Anggraini et al. utilized segmentation models to automatically separate images of parasitized malaria cells from their backgrounds. Another group used unsupervised methods to automatically cluster malaria cells using blood smear images [1]. Other researchers utilized spatial uncertainty sampling to reduce domain shift bias in segmentation mask generation for parasitized cells [16].

Overall, it can be concluded that although machine learning techniques have shown a high propensity for malaria related tasks, limitations exist. Specifically, the aforementioned studies almost exclusively employ complex, deep neural network architectures that demand high levels of computational power for training. To that end, we note an inherent lack in algorithms with the ability to provide state of the art diagnostic metrics while maintaining ideal computational efficiency. Since malaria classification from blood smear images is a feature rich task, our research looks to explore the implementation of a concise, relatively shallow, convolutional neural network architecture for accurate detection.

II. DATA

A. Dataset Acquisition

The dataset used in this study was provided by the National Institute of Health (NIH) and contained 27,558 images with equal instances of parasitized and non-parasitized red blood cells. The cells were stained with Giemsa, mimicking the procedure that would be undertaken when using microscopic analysis to diagnose malaria. A sample of the dataset is shown in Fig. 2 and Fig. 3.
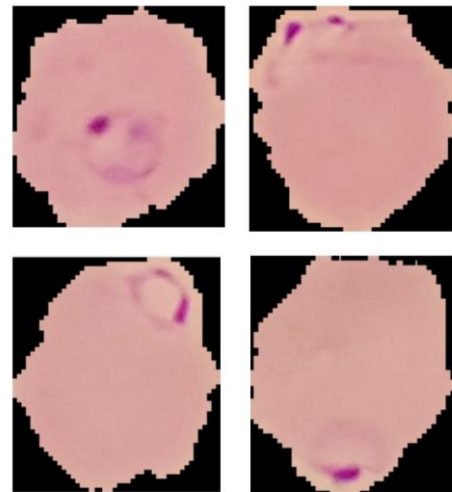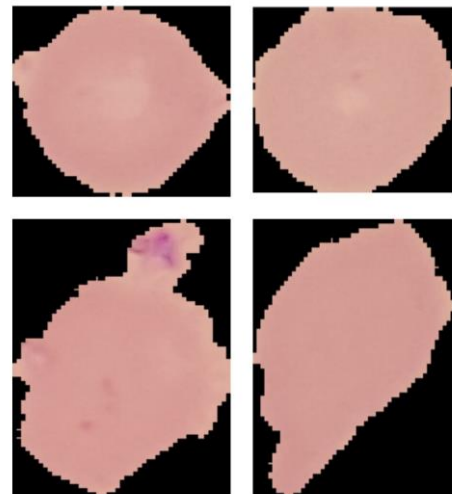

Fig. 2. Parasitized red blood cell smears

Fig. 3. Uninfected red blood cell smears

B. Image Preprocessing

The first step was to process the images and split them into train and test subsets, thereby making the dataset ready for usage. The images were processed via the nearest-neighbor interpolation technique, in which a group of pixels automatically assume the magnitude of the pixel closest to it. It is the simplest method of multivariate interpolation, which also makes it the most computationally efficient. The images were downscaled to a standard size of 100x100 pixels, which was necessary as the images were previously of variable size and scale. Applying nearest-neighbors interpolation minimized image distortion and retained information better than less sophisticated methods such as resizing or cropping. We then placed 80% of the images in a training dataset, while the remaining 20% were used for testing.

III. MALARIANET

The malaria detection task is a binary classification problem, where the input is an image of a cell X and the output is a binary label $y \in \{0, 1\}$ indicating the absence or presence of malaria, respectively.

To accomplish this task, a seven-layer CNN was utilized that we dubbed MalariaNet. This model contained two convolution layers, two pooling layers, a flattening layer, and two fully connected dense layers. The dense layers alleviate the vanishing gradient problem, strengthen feature propagation, encourage feature reuse, and substantially reduce the number of parameters. The network structure can be seen in MalariaNet Layer Map. Within our model, we utilized Adaptive Moment Estimation (Adam), an adaptive learning rate optimization algorithm. Our activation functions for the intermediate layers was the rectifier function, with the final layer using a sigmoid nonlinearity. A batch size of ten was used over 25 epochs, after each of which the model output accuracy metrics and loss, given by the binary cross-entropy loss function: $L(X, y) = -w + \cdot y \log p(Y = 1|X) - w \cdot (1 - y) \log p(Y = 0|X)$, where $p(Y = i|X)$ is the probability that the network assigns to the label i, $w+ = |N|/(|| + ||)$, and $w- = |P|/(|P| + |N|)$ with $|P|$ and $|N|$ the number of parasitized and uninfected cells in the training set respectively.

In order to prevent the networks from overfitting, early stopping was performed by saving the network after every epoch and choosing the saved network with the lowest loss on the tuning set. Overall, 2,177,185 parameters were trained and optimized for this task. The architecture of the network can be seen in Table 3.

TABLE III.

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 98, 98, 32)	896
Max_pooling2d_1 (MaxPooling2)	(None, 49, 49, 32)	0
conv2d_2 (Conv2D)	(None, 47, 47, 32)	9248
max_pooling2d_2 (MaxPooling2)	(None, 23, 23, 32)	0
flatten_1 (Flatten)	(None, 16928)	0
Dense_1 (Dense)	(None, 128)	2166912
Dense_2 (Dense)	(None, 1)	129

IV. RESULTS

The progression of MalariaNet training can be seen in Fig. 4, 5, and 6. We present loss, accuracy, F1 score, precision, and recall by each epoch.
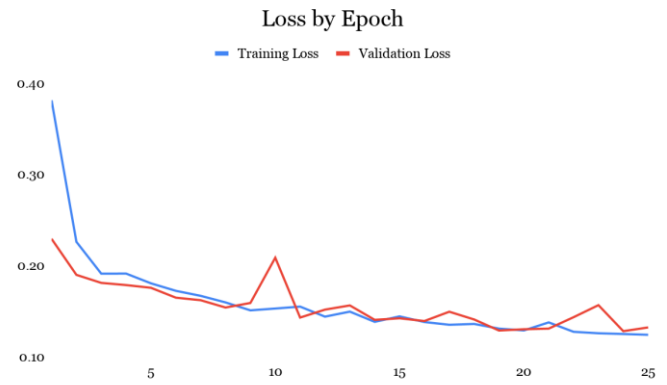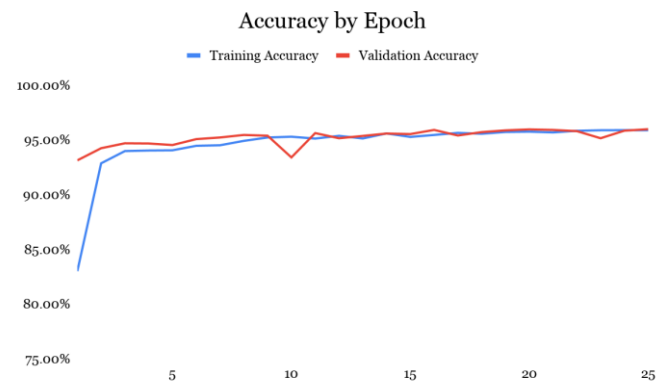

Fig. 4. Loss by epoch
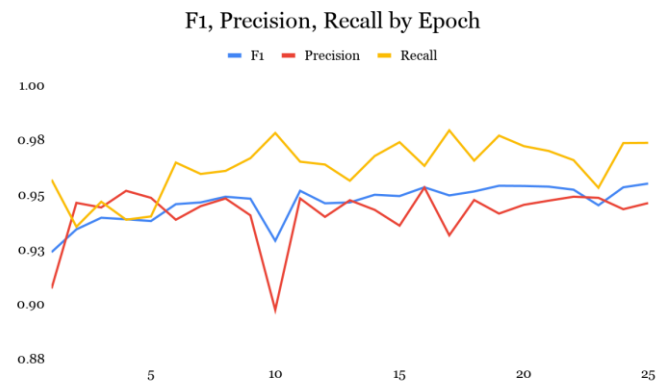

Fig. 5. Accuracy by epoch

Fig. 6. F1, precision, recall by epoch

MalariaNet achieved relatively high performance in all metrics after training was complete, as can be seen in Table 4.

TABLE IV.

Accuracy	F1 Score	Specificity	Sensitivity
0.970	0.955	0.946	0.974

V. CONCLUSION

Malaria afflicts hundreds of thousands of people around the world and is among the leading causes of death in developing countries [5]. The disease is primarily diagnosed through the technique of microscopic analysis, which necessitates trained laboratories and operational facilities that

are often unavailable. Seeking to alleviate this predicament, we developed a compact and efficient convolutional neural network architecture, MalariaNet. We then trained MalariaNet on a publicly available dataset containing approximately 27,000 parasitized and uninfected red blood cell images. Our model attained high performance metrics across the board, including a 96.8% accuracy and 0.955 F1 score. MalariaNet's intriguing diagnostic capabilities provide the potential for its implementation in remote locales while concurrently promoting subsequent exploration regarding the ability of relatively shallow deep learning architectures in boosting accuracy and efficiency. We hope that our model will be used to further the development of efficient algorithms for disease detection.

REFERENCES

[1] Abdul-Nasir, A. S., Mashor, M. Y., & Mohamed, Z. (2017). Clustering approach for unsupervised segmentation of malarial Plasmodium vivax parasite. https://doi.org/10.1063/1.5005337

[2] Analysis of research and development priorities for ... (n.d.). Retrieved from https://www.who.int/research-observatory/analyses/malaria_rd_priorities_working_paper.pdf?ua=1

[3] Anggraini, D., Nugroho, A. S., Pratama, C., Rozi, I. E., Iskandar, A. A., and Hartono, R. N. "Automated status identification of microscopic images obtained from malaria thin blood smears," in Proceedings of the 2011 International Conference on Electrical Engineering and Informatics, Bandung, Indonesia, July 2011.

[4] CDC - Parasites - Malaria. (2020, April 1). Retrieved from https://www.cdc.gov/parasites/malaria/index.html

[5] Garrido-Cardenas, J. A., Cebrián-Carmona, J., González-Cerón, L., Manzano-Agugliaro, F., & Mesa-Valle, C. (2019, May 31). Analysis of Global Research on Malaria and Plasmodium vivax. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6603864/.

[6] Gopakumar, G. P., Swetha, M., Siva, G. S., and Sai Subrahmanyam, G. R. K. "Convolutional neural network-based malaria diagnosis from focus stack of blood smear images acquired using custom-built slide scanner," Journal of Biophotonics, vol. 11, no. 3, 2018.

[7] Improving Malaria Parasite Detection from Red Blood Cell ... (n.d.). Retrieved from https://www.researchgate.net/publication/334669002_Improving_Malaria_Parasite_Detection_from_Red_Blood_Cell_using_Deep_Convolutional_Neural_Networks

[8] Kaewkamnerd, S., Uthaipibull, C., Intarapanich, A., Pannarut, M., Chaotheing, S., and Tongsima, S. "An automatic device for detection and classification of malaria parasite species in thick blood film," BMC Bioinformatics, vol. 13, Supplement 17, p. S18, 2012

[9] Liang, Z., Powell, A., Ersoy, I., et al., "CNN-based image analysis for malaria diagnosis," in 2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 493–496, Shenzhen, China, December 2016.

[10] Malaria Datasets | National Library of Medicine. (n.d.). Retrieved from https://lhncbc.nlm.nih.gov/publication/pub9932

[11] Masud, M., Alhumyani, H., Alshamrani, S. S., Cheikhrouhou, O., Ibrahim, S., Muhammad, G., … Shorfuzzaman, M. (2020, July 8). Leveraging Deep Learning Techniques for Malaria Parasite Detection Using Mobile Application. https://www.hindawi.com/journals/wcmc/2020/8895429/. learning and deep learning architectures for malaria detection on cell images. Retrieved from https://www.spiedigitallibrary.org/conference-proceedings-of-spie/11139/2524681/Performance-analysis-of-machine-learning-and-deep-learning-architectures-for/10.1117/12.2524681.short?SSO=1

[12] Narayanan, B. N., Ali, R., & Hardie, R. C. (2019, September 6). Performance analysis of machine learning and deep learning architectures for malaria detection on cell images. Search the world's largest collection of optics and photonics applied research. https://www.spiedigitallibrary.org/conference-proceedings-of-spie/11139/2524681/Performance-analysis-of-machine-learning-and-deep-learning-architectures-for/10.1117/12.2524681.short?SSO=1.

[13] Rahman, A., Zunair, H., Rahman, M., Yuki, J., Biswas, S., Alam, M., Mahdy, M. (2019, July 23). Improving Malaria Parasite Detection from Red Blood Cell using Deep Convolutional Neural Networks. Retrieved December 19, 2020, from https://arxiv.org/abs/1907.10418

[14] Rajaraman, S., Antani, K., Poostchi, M., et al., "Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images," PeerJ, vol. 6, article e4568, 2018.

[15] Shi, G. (2019, January 30). Detecting malaria using deep learning. Retrieved from https://towardsdatascience.com/detecting-malaria-using-deep-learning-fd4fdcee1f5a

[16] Soleimany, A. P., Suresh, H., Ortiz, J. J. G., Shanmugam, D., Gural, N., Guttag, J., & Bhatia, S. N. (2019, November 30). Image segmentation of liver stage malaria infection with spatial uncertainty sampling. arXiv.org. https://arxiv.org/abs/1912.00262.

[17] Talapko, J., Škrlec, I., Alebić, T., Jukić, M., & Včev, A. (2019, June 21). Malaria: The Past and the Present. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6617065/