MOSQUITO-NET: A deep learning based CADx system for malaria diagnosis along with model interpretation using GradCam and class activation maps

Aayush Kumar | Sanat B. Singh | Suresh Chandra Satapathy | Minakhi Rout

Abstract
Malaria is considered one of the deadliest diseases in today's world, which causes thousands of deaths per year. The parasites responsible for malaria are scientifically known as *Plasmodium*, which infects the red blood cells in human beings. Diagnosis of malaria requires identification and manual counting of parasitized cells in microscopic blood smears by medical practitioners. Its diagnostic accuracy is primarily affected by extensive scale screening due to the unavailability of resources. State of the art Computer-Aided Diagnostic techniques based on deep learning algorithms such as CNNs, which perform an end to end feature extraction and classification, have widely contributed to various image recognition tasks. In this paper, we evaluate the performance of Mosquito-Net, a custom made convnet to classify the infected and uninfected blood smears for malaria diagnosis. The CADx system can be deployed on IoT and mobile devices due to its fewer parameters and computation power, making it wildly preferable for diagnosis in remote and rural areas that lack medical facilities. Statistical analysis demonstrates that the proposed model achieves greater accuracy than the previous SOTA architectures for malaria diagnosis despite being 10 times lighter in parameters and inference time. Mosquito-Net achieves an AUC of 99.009% and an F-1 score of 96.7% on the validation set.

KEYWORDS
AUC, CNNs, computer-aided diagnostic techniques, F-1 score, malaria, Mosquito-Net

1 | INTRODUCTION

Machine Learning algorithms for the healthcare domain have shown their potential to aid healthcare professionals and reduce medical diagnosis errors. Previously, various machine learning models employed for diagnosing malaria have used handcrafted feature extraction methods. These methods are optimized for a particular dataset having a specific probability distribution and suitable for handling particular variance in the distribution. In the current scenario, with the availability of many datasets and computational resources, Deep learning algorithms have overcome the previous shortcomings by automatically identifying the specific attributes from the structured data to carry out the specified tasks. A particular class of neural networks, Convolutional Neural Networks (CNNs), has shown significant and influential results in image classification, recognition, localization, and segmentation tasks for the medical domain.

Computer-Aided Diagnostic (CADx) tools can reduce the burden of a medical practitioner by reducing the manual labor required for disease classification with Deep Learning techniques/algorithms. One survey article gives insights on different manual and automated diagnostic methods...
for malaria detection (CADx tools) and the potential of neural networks and Deep Learning (Shang, 2016) for the task of malaria detection from blood-smeared microscopic images.

Malaria is usually spread and transmitted by female Anopheles mosquitoes. Plasmodium parasites are mainly responsible for infecting humans by attacking their RBCs (red blood cells) and cause deadly life-threatening conditions (World Health Organization, 2019). P. vivax was responsible for 75% of cases in the WHO Region of the Americas and 47% cases in India alone. About 67% of the estimated death counts were of children under 5 years of age (World Health Organization, 2020). The WHO recommends that all suspected malaria cases should be tested using parasite-based diagnostic testing. Either microscopy (Examination of Blood Smears via microscope) or RDT - Rapid Diagnosis Testing (Antigen Testing – less cost-effective in an endemic region like sub-Saharan Africa). Microscopy has been regarded as one of the most reliable ways of detection. In the process, the blood sample is put under a slide, and a dyeing agent is used for staining and highlighting the malaria parasites in RBCs. A trained clinician then examines the slide under a microscope, who manually counts the number of infected RBCs. According to the WHO protocols, up to 5000 cells should be counted manually by the clinician (World Health Organization: WHO, 2020). The clinician needs to be well trained and competent enough to identify malaria parasites. Human error becomes a prominent factor in manual diagnosis – accuracy impacted by factors like observer variability, extensive scale screening, etc. (Mitiku et al., 2003).

Generally, due to a limited amount of medical annotated data, a technique known as transfer learning is used to train the Deep Learning model for the target applications. Generally, pre-trained CNN models trained on large datasets like ImageNet (Deng et al., 2009; LeCun et al., 2015) are adjusted according to the requirements to carry out inference tasks. Such methods have proven effective but have a couple of drawbacks. Firstly, the pre-trained models start over-fitting on the training data, and the results are biased if the amount of training data is less. Secondly, the models’ dimensions are generally quite large in terms of the number of parameters making them computationally expensive to carry out inference tasks. Due to the stated reasons, their deployment on the edge and mobile devices for production is a difficult task.

In this paper, we propose Mosquito-Net, a Deep Learning-based CADx system for the task of malaria diagnosis using blood-smeared microscopic images. The proposed Mosquito-Net achieves a validation accuracy of 96.6% despite being about 10x times lighter than the previously mentioned state of the art architectures. The rest of the paper’s organization is described as follows: Section 2 contains the previous literature works. Mosquito-Net is depicted and described in detail in Section 3. The implementation results, comparison, and analysis of results are discussed in Section 4. Finally, the concluding remarks have been made in Section 5.

## 2 | RELATED WORKS

Many studies have been carried out, highlighting the considerable amount of work done for malaria diagnosis tasks using computer-aided diagnostic systems. Previous works (Dong et al., 2017) compare support vector machines (SVMs) and CNN’s performance in the classification of infected and normal healthy cells. Collection of images of segmented RBCs were spliced in 3 tier format – Train, Validation, and Test sets (randomly). While the SVM classifier performed well with around 92% accuracy, CNN’s emerged victorious with a significant difference with 95% accuracy. Liang et al. (2016) evaluated and compared the performances of pre-trained and custom CNN models in the classification of infected and normal healthy cell images by performing cross-validation studies (Figure 1).
Gopakumar et al. (2017) created a custom convnet to detect parasites’ presence over the slide images; it outperformed the SVM classifier significantly, which had 91.81% MCC with an MCC score of 98.77%. Rajaraman et al. (2018) used pre-trained CNNs over annotated clinical image datasets to extract features from the best working layers. They validated their performance in classifying infected (parasitized) and healthy normal cells statistically.

The ResNet-50 achieved the highest accuracy of 95.90%, among others. Even though we are advancing in machine learning, having a portable device for resource-constrained settings is required to carry out inference tasks for malaria diagnosis. The most accessible device present with one and all in the current time is smartphones, as their computation power has grown significantly over the past few years. While the results of the mentioned studies were excellent and promising, they were bulky in terms of parameters and required significant computation power for deployment on edge devices like smartphones.

3 | METHODOLOGY

We propose Mosquito-Net, a lightweight CADx system using CNNs, as they help in reducing the manual feature engineering and expertise that would have been required for traditional ML algorithms. However, they suffer from high variance and are often prone to overfitting, where they perform well on the training set but deteriorate on the validation set. We use a dropout of probability 0.2 and various Data augmentation techniques to overcome the overfitting problem. The augmentation methods applied consists of random flipping of images, resizing, colour jitter, translation, etc. The proposed workflow of the CADx system is depicted in Figure 2. We were able to create a model that provides promising performance and, at the same time, requires way less storage for deployment on edge devices like smartphones.

The details about the dataset and pre-processing steps followed are discussed in Sections 3.1 and 3.2. The architecture is described in Section 3.3, and Training details are discussed in Section 3.4. Finally, we also propose a web-based platform, which is outlined in Section 3.5.

3.1 | Dataset

The malaria dataset from Rajaraman et al. (2018) is used to train and test the proposed approach. The microscopic blood images and annotations were obtained from the National Library of Medicine, NLM (IRB#12972), which were de-identified. The images were captured using a mobile application and a built-in camera that was connected to a microscope. The dataset was prepared by collecting Giemsa stained thin blood smear slides from infected and healthy patients at Chittagong Medical College Hospital, Bangladesh. An expert slide reader manually annotated the slides.

For training and testing, we use a 70–30 split. 70% of the total dataset is used for training purposes, and the rest 30% of the data is used for validation and testing purposes.

3.2 | Pre-processing and data augmentation

Various pre-processing methods were tried out to decide the optimal methods which improve the performance of Mosquito-Net. The data augmentations used were Random Rotation, Random Horizontal Flip, Random vertical flip, and colour jitter of 0.05% was used on the training images.
It helps us increase the dataset’s size and introduce variability in the dataset without collecting new data. Furthermore, the input images were normalized with mean of $[0.5, 0.5, 0.5]$ and standard deviation of $[0.5, 0.5, 0.5]$.

Figure 3(a) Shows a raw blood-smear image. Figure 3(b–e) shows the results of various augmentation techniques used on the dataset for the pre-processing procedure.

### 3.3 | Mosquito-Net architecture

The primary building block of our computer-aided diagnostic system is Mosquito-Net. Most of the previous methods use transfer learning techniques, where models previously trained on ImageNet are instantiated and fine-tuned according to malaria diagnosis requirements. Rajaraman et al. (2018) used transfer learning techniques on five pre-trained CNNs. They achieved impressive results, but the main issue with their method is it is highly inefficient, and it requires high computation power. Suppose we need to run malaria tests in a remote location where there are not enough medical facilities. Ideally, the device would require a lightweight model that consumes less power and gives desirable accuracy in predicting malaria cases. It can be efficiently run on devices incorporating the IoT and edge devices.

The Mosquito-Net proposed by us depicted in Figure 4 uses about 10× less time inference time on CPU than state of the art (SOTA) ResNet50 besides being about 5× times lighter. We propose a sequential convnet (CNN) for the task of malaria detection. The proposed Mosquito-Net shown in Figure 4 consists of three convolutional layers and three fully connected layers. The model is fed with an RGB image of $(120 \times 120 \times 3)$ dimensions. The convolutional layers use a kernel of $5 \times 5$ pixels and a stride of 1 pixel. Besides, padding of $2 \times 2$ pixels is also
used. The convolution block's output is batch normalized and passed through the ReLU activation function (Shang, 2016) before being max pooled by a kernel size of $2 \times 2$ pixels and a stride of 2. The output of the third conv block is fed to the F.C. layers having 512 and 128 neurons. Dropout is used with a probability of 0.2 to avoid overfitting (LeCun et al., 1989).

### 3.4 Training details

The proposed Mosquito-Net is trained on the malaria dataset (Rajaraman et al., 2018), and a scheduling L.R. policy is used, where the learning rate is decreased by a small amount when training metrics remain the same for an extended period. ADAM (Kingma & Ba, 2019) optimizer is used for optimizing the Mosquito-Net. The Mosquito-Net was built and evaluated using the PyTorch, a deep learning library.

The hyperparameters used for training were:

1. LR = $1e^{-3}$
2. Epochs = 20
3. Batch size = 32

We use Binary Cross-Entropy (1) as the loss function for training and converging our deep learning model.

$$\mathcal{L}_{\text{binary cross-entropy}}(x_i) = - (y_i \log(f(x_i)) + (1 - y_i) \log(1 - f(x_i)))$$

where $x_i$ and $y_i$ are the input features and labels and $f(x_i)$ is the Mosquito-Net’s output, the probability of whether it is infected or uninfected.

The Mosquito-Net converges, as can be inferred from the loss curve in Figure 5. The plot of training loss decreases to the point of stability along with the plot of validation loss. Also, the validation loss has a small gap with the training loss, indicating no overfitting.

Convolution operation can be summarized as a series of convolutional products using various kernels of filters, which is applied on an input image, and the output is passed through an activation function. In the end, we have fully connected layers for the classification of the convolution outputs.

### 3.5 Web-based deployment

We also introduce a web-enabled platform that can be accessed at [https://zyik.ml/malaria/templates/index.html](https://zyik.ml/malaria/templates/index.html). It can be deployed on IoT devices like Raspberry Pi and Nvidia Jetson-Nano. The platform receives the input image where it is converted to a byte stream and tested with a pixel similarity function to verify whether the image uploaded has an appropriate blood smear or not (The similarity function compares the input
image with sample images from the dataset). If verification fails, the user/client is redirected to the home page with an appropriate message. The input image is initially passed through pre-processing steps that include reshaping and converting to a tensor before carrying out the analysis through the deep learning model. We have tested the web server on an Amazon EC2 instance (t2.micro instance) and on very constrained resource instances (t3a.nano and t2.nano instances) to evaluate its efficiency. The web server uses Flask and PyTorch frameworks (Flask and Caffe2 framework version also present) (Figure 6).
4 RESULTS

The malaria dataset from Rajaraman et al. (2018) is used to train and test the proposed Mosquito-Net. The dataset contains a total of 27,558 cell images with equal instances of parasitized and uninfected cells. Figure 7 shows the bar chart of the average inference time of all the CNN architectures. The hyperparameter optimization was done by manual grid search. We have evaluated the proposed method in terms of accuracy, AUC, F1 score, sensitivity, specificity, and MCC. Equations for these measures are given in Equations (1)–(6). All criteria report more than a 90% success rate. Sensitivity and specificity for the results are around 96%–98%.

1. Accuracy

\[
\text{Accuracy} = \frac{\text{Number of Correct predictions}}{\text{Total no of predictions made}} \tag{2}
\]

2. F1-Score

\[
F1 = 2 \times \frac{1}{\frac{1}{\text{precision}} + \frac{1}{\text{recall}}} \tag{3}
\]

3. Sensitivity

\[
\text{Sensitivity} = \frac{\text{True Positive}}{\text{False Negative + True Positive}} \tag{4}
\]

4. Specificity

\[
\text{Specificity} = \frac{\text{False Positive}}{\text{False Positive + True Positive}} \tag{5}
\]

5. MCC

\[
\text{MCC} = \frac{\text{TP} \times \text{TN} - \text{FP} \times \text{FN}}{\sqrt{(\text{TP} + \text{FP})(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TN} + \text{FN})}} \tag{6}
\]

**FIGURE 7** Depiction of average inference times of various CNN models
4.1 Quantitative analysis

Table 1 presents a performance comparison to other studies using the same malaria dataset from (Rajaraman et al. (2018)). The comparison criteria are accuracy, AUC, F1 score, sensitivity, specificity, and MCC. Our preliminary results outperform all of the other approaches. Mosquito-Net’s results against other transfer learning methods were presented in Table 1, which indicates that Mosquito-Net obtained the best accuracy, AUC, sensitivity, specificity, F1, and MCC. The definition of metrics can be found in Equations (1)–(6). Hence, Mosquito-Net gives the best performance compared to other transfer learning methods.

Our preliminary results outperform all of the other approaches. Our Mosquito-Net was compared with five state-of-the-art approaches: DenseNet-121 (Huang et al., 2018), Res-Nets (He et al., 2016), Alex-Net (Krizhevsky, 2014), and VGGs (Simonyan, 2015). All performances were compared on the test set and presented in Table 2 in terms of the number of parameters and average inference times of CPU and GPU. We can observe that carrying out inferences with Mosquito-Net could significantly provide better performance than other transfer learning methods in terms of all three indicators. Comparison performance is presented in Table 2.

A reduce learning rate on the plateau with respect to the validation loss approach was used (Chollet, 2016) for model training. Mosquito-Net outperformed all other models mentioned in the Table 1 across all metrics. As can be inferred from Table 2, the Mosquito-Net is the smallest in terms of parameters and took ~20× less time for inference on CPU and ~27× less inference time on GPU. GPUs boost inference times significantly, but more focus is pointed to CPU inference times due to the latter’s higher costs. The same training conditions were maintained and checked before computing inference times for all the models.

4.2 Qualitative analysis

Deep learning can provide rapid preliminary results and improve access to quality malaria diagnoses in the absence of specialist pathologists. To get a better insight into how Deep Learning models make their decision, we take a look at the output of GRADCAM, which highlights the input image areas in the form of heat maps that impact the findings of the Mosquito-Net.

To ensure the Mosquito-Net model was learning pertinent features, we generated class activation mappings (CAMs) to highlight the image areas most indicative of the disease. It is generated by a method known as GradCam. In Figure 8, the first column displays a blood-smeread input image fed into the Mosquito-Net. The corresponding right columns consist of a heatmap pointing out the image areas that indicate the presence of malaria infection. Thus, the brightest regions of the CAMs are the regions that most influence the model’s prediction.

TABLE 1 Quantitative analysis in terms of accuracy, AUC, F1-score, sensitivity, specificity

| Model                 | Accuracy  | AUC      | Sensitivity | Specificity | F1-score | MCC     |
|-----------------------|-----------|----------|-------------|-------------|----------|---------|
| Mosquito-Net          | 0.966 ± 0.010 | 0.990 ± 0.009 | 0.976 ± 0.004 | 0.958 ± 0.002 | 0.967 ± 0.001 | 0.933 ± 0.002 |
| Alex-Net              | 0.927 ± 0.15  | 0.978 ± 0.09  | 0.939 ± 0.18  | 0.931 ± 0.44  | 0.939 ± 0.11  | 0.869 ± 0.24  |
| VGG-16                | 0.951 ± 0.12  | 0.979 ± 0.08  | 0.941 ± 0.19  | 0.949 ± 0.02  | 0.939 ± 0.16  | 0.891 ± 0.30  |
| ResNet-50             | 0.961 ± 0.07  | 0.989 ± 0.03  | 0.942 ± 0.23  | 0.968 ± 0.08  | 0.961 ± 0.08  | 0.909 ± 0.14  |
| Xception-Net         | 0.889 ± 0.10  | 0.956 ± 0.05  | 0.941 ± 0.40  | 0.841 ± 0.22  | 0.889 ± 0.10  | 0.781 ± 0.33  |
| DenseNet-121         | 0.929 ± 0.18  | 0.967 ± 0.04  | 0.939 ± 0.24  | 0.919 ± 0.29  | 0.929 ± 0.17  | 0.889 ± 0.36  |

TABLE 2 Quantitative analysis in terms of no. of parameters and inference time (in ms) on CPU and GPU

| Model name           | No. of parameters | Avg. inference time on CPU | Avg. inference time on GPU |
|----------------------|-------------------|---------------------------|---------------------------|
| Mosquito-Net         | 7,472,002         | 0.016                     | 0.002                     |
| DenseNet121          | 7,978,856         | 0.327                     | 0.054                     |
| ResNet-18            | 11,689,512        | 0.136                     | 0.008                     |
| ResNet-34            | 21,797,672        | 0.226                     | 0.015                     |
| ResNet-50            | 25,557,032        | 0.268                     | 0.02                      |
| Alex-Net             | 61,100,840        | 0.05                      | 0.005                     |
| VGG-16               | 138,357,544       | 0.748                     | 0.02                      |
| VGG-19               | 143,667,240       | 0.862                     | 0.026                     |
| VGG-19 BN            | 143,678,248       | 0.913                     | 0.027                     |
Mosquito-Net’s fully trained convolutional network is fed with an image to generate the CAMs. Then the feature maps are extracted from the output of the last convolutional layer.

We are currently carrying out a detailed analysis of deploying the customized Mosquito-Net model in smartphones and IoT devices and analyzing the performances. Now, various Android and iOS ML libraries with Firebase technology offer flexibility for dynamic memory allocation of CPU and GPU based on the user’s requirement. Our pilot studies observed that the proposed model occupied only 46 M.B. and took little time to make predictions on the test data. The deployed model could serve as a triage tool and minimize delays in disease-endemic/resource-constrained settings.

5 | CONCLUSION

Mosquito-Net, a Deep CNN-based malaria disease classification model from microscopic blood-smeared images, is introduced in this paper. Medical image processing has been gaining much attention recently due to the emergence of deeper and high-accuracy classification networks that can compete against humans and speed-up biomedical research to a great extent. Deep learning models do not require manual pre-processing or feature engineering, nor do they require any manual parameter tweaking. Our proposed model, Mosquito-Net, incorporates the latest deep learning advancements for accurate malaria classification from blood-smeared images in this paper. Our mosquito-net achieves an accuracy of 96.6% on the validation set. Tackling the problem of automated malaria classification detection can improve the recovery rate and enable faster cures, thus improving people’s overall health and quality of life.

Providing pathologists with predictions from a deep learning model could improve the quality and consistency of malaria diagnosis. By leveraging automation by experts, computer-aided diagnostics systems can help healthcare services and reduce the radiologists’ and pathologists’ burden in developing countries where healthcare services are limited.

CONFLICT OF INTEREST
There are no conflicts of interest among the authors.

DATA AVAILABILITY STATEMENT
We do not have any data to share.

ORCID
Suresh Chandra Satapathy https://orcid.org/0000-0001-8236-4104
AUTHOR BIOGRAPHIES

Aayush Kumar is doing his undergraduate study in Computer Science Engineering from School of Computer Engg, KIIT Deemed to be University, Bhubaneswar, Odisha, India. He has research interest in Machine Learning for Healthcare Applications.

Sanat B. Singh is doing his undergraduate study in Computer Science Engineering from School of Computer Engg, KIIT Deemed to be University, Bhubaneswar, Odisha, India. He has research interest in machine learning and AI for Healthcare.

Suresh Chandra Satapathy is a professor in School of Computer Engg, KIIT Deemed to be University, Bhubaneswar, Odisha, India. His research interests are on machine Learning, soft computing, swarm intelligence etc. He has authored more than 150 research articles in various journals and conferences.

Minakhi Rout is a faculty in School of Computer Engg. KIIT Deemed to be University, Bhubaneswar, Odisha, India. Her research works are in the domain of Machine Learning and Soft Computing. She has guided many undergraduate and graduate projects and also published papers in her domain of expertise.

How to cite this article: Kumar A, Singh SB, Satapathy SC, Rout M. MOSQUITO-NET: A deep learning based CADx system for malaria diagnosis along with model interpretation using GradCam and class activation maps. Expert Systems. 2022;39:e12695. https://doi.org/10.1111/exsy.12695