Advances in Artificial Intelligence and Machine Learning; Research 6 (1) 4947-4958 Received 24-07-2025; Accepted 23-01-2026; Published online 30-01-2026

Automatic Malaria Parasite Detection Using Machine Learning and Image Classification

Ibukun O. Eweoya eweoyai@babcock.edu.ng
Babcock University - Department of Software Engineering
Ilisan Remo - Nigeria

Abayomi Agbeyangi aagbeyangi@wsu.ac.za
Walter Sisulu University, South Africa - Department of Business and Applications Development
Buffalo City Campus, East London - South Africa

Jose Lukose jlukose@wsu.ac.za
Walter Sisulu University, South Africa - Department of Business and Applications Development
Buffalo City Campus, East London - South Africa

Corresponding Author: Ibukun Eweoya

Abstract

Microscopic examination of stained blood smears remains a reference method for malaria diagnosis, but is labour-intensive and operator-dependent. We evaluated image-based machine-learning models for parasite detection in single-cell images and report a deployment-oriented configuration. We used the NIH “Cell Images for Malaria” set (27,558 labelled cells; 13,780 parasitised; 13,778 uninfected). We compared a lightweight CNN against SVM and Naïve Bayes baselines. Variants with HOG features and dimensionality reduction (PCA/LDA) were explored for analysis only. The best CNN achieved 95.0% accuracy, 94.7% precision, 95.3% recall, and 95.0% F1 on the held-out test set, with a ROC-AUC of 0.91. We showed that a compact CNN trained on publicly available cell images yields high diagnostic performance and can be embedded in a simple desktop GUI for point-of-care assistance.

Keywords: Malaria, Malaria detection, Plasmodium falciparum, Convolutional neural networks, Image classification.

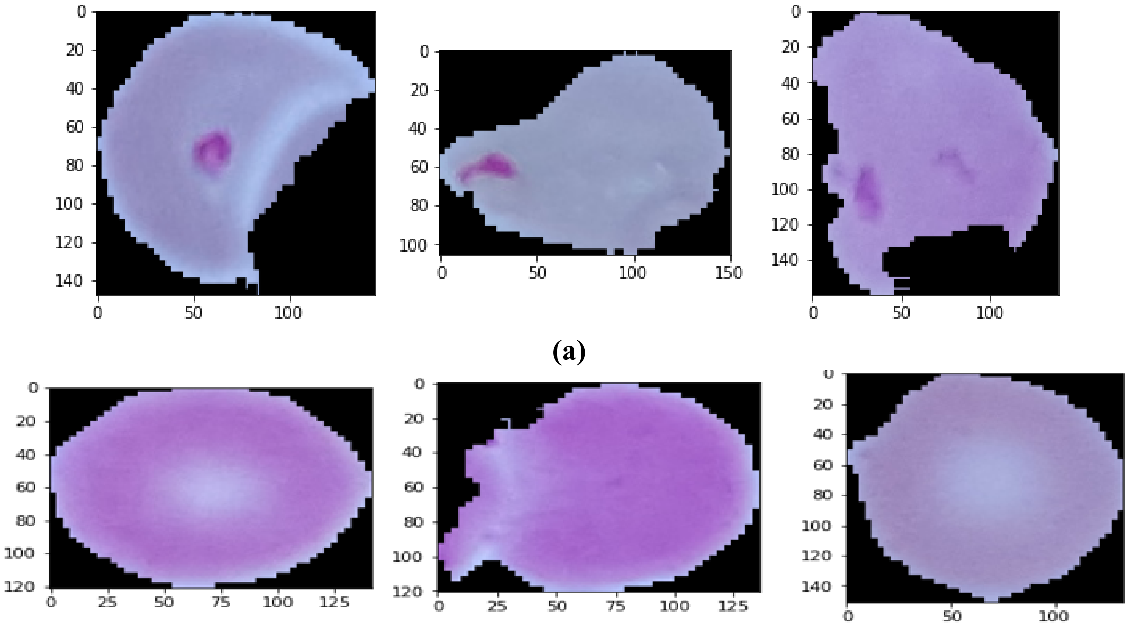
1. INTRODUCTION

Malaria remains a major cause of febrile illness and preventable death in sub-Saharan Africa [1–3]. Case management depends on prompt, accurate diagnosis, yet routine microscopy is vulnerable to workload, slide quality and operator variability. In low-resource settings, these constraints translate directly into missed infections and treatment delays.

Malaria remains a major cause of febrile illness and preventable death in sub-Saharan Africa [4]. Case management depends on prompt, accurate diagnosis, yet routine microscopy is vulnerable to

Citation: Ibukun O. Eweoya, et al. Automatic Malaria Parasite Detection Using Machine Learning and Image Classification. **Advances in Artificial Intelligence and Machine Learning.** 2026;6(1):274. https://dx.doi.org/10.54364/AAIML.2026.61274

workload, slide quality and operator variability. In low-resource settings, these constraints translate directly into missed infections and treatment delays. According to the WHO, transmission of malaria parasites persists across many regions, with the most burden noted in sub-Saharan Africa. In 2023, there were hundreds of millions of cases globally and an estimated ~600,000 deaths. Children are seen as having the majority of malaria-related mortality, making them an easy prey [5, 6]. In countries such as Nigeria, delayed or inaccurate diagnoses remain a leading cause of preventable deaths from malaria [7]. Microscopic examination of stained blood smears is still considered labour-intensive, time-consuming, and prone to human error, even though it is still the gold standard. FIGURE 1, shows examples of infected and uninfected red blood cell images, showing examples of visual features that underpin automated parasite detection.



(a)

(b)

Figure 1: Sample microscopic blood smear images. (a) Images of malaria-infected red blood cells showing visible parasites. (b) Images of uninfected red blood cells with no visible parasitic presence.

These challenges show that there is a need for more accurate, scalable, and accessible diagnostic tools that is suitable for low-resource environments. Traditional methods, which are often seen as effective under controlled conditions, frequently fail to meet these requirements at scale [8, 9]. This study addresses that gap by evaluating a machine learning-based model that can help automate the detection of malaria parasites in blood smear images. The approach handles large or complex datasets and takes care of common issues, such as class imbalance and outliers, that many conventional techniques struggle to address. Recent advances in computer vision and deep learning have shown strong potential in medical image analysis and disease prediction [10, 11]. Leveraging these advances, we aim to develop a reliable image-based detection system that supports healthcare workers and strengthens malaria diagnosis in low-resource settings.

The main research question guiding this study is: "How accurately can a machine learning-based model detect malaria parasites from microscopic blood smear images compared with traditional diagnostic methods?". Since the eradication of malaria parasite disease is mostly influenced by environmental and biological factors, modern technology tools can provide more effective ways for its control and management. Automated diagnostic systems, in particular, offer a promising pathway to reducing preventable malaria deaths by improving both the speed and accuracy of diagnosis.

2. RELATED WORKS

Literature has noted automation of malaria detection with digital image analysis and machine learning in recent years [9–20]. These efforts basically target shortages of trained microscopists, observer fatigue, and the throughput limits of manual microscopy. Methods span classical image processing to modern deep learning, with an emphasis on reliable, scalable parasite detection.

Traditional pipelines most of the time begin with image enhancement, conversion to greyscale, histogram equalisation, and noise suppression, followed by segmentation to isolate red blood cells and flag parasitised cells. Thresholding, region growing, and watershed algorithms are common choices. Oliveira et al. [11] proposed a colour-driven approach using HSV components that adapts well in noisy or imperfect slides. Manning et al. [16] combined an improved watershed method with a feed-forward neural network in a graphical interface and reported 92% accuracy on a dataset exceeding 27,000 images. These studies underscore the foundational role of preprocessing and segmentation in dependable computer-assisted diagnosis.

Deep learning has reduced reliance on old and manual features by learning discriminative morphology directly from pixels. Maturana et al. [10] developed a CNN tailored for mobile deployment, improving portability and access. Shambhu et al. [19] reported 96.02% accuracy for CNN-based classification of blood-smear images. Object-detection architectures have also been explored by Shewajo and Fante [18], and achieved a 95.3% recall using YOLOv4-tiny on thick smears. More recently, Chibuta et al. [20] and Bae et al. [21] demonstrated smartphone-compatible systems with 97–99% detection accuracy in field-like conditions, highlighting the feasibility of low-cost, near–real-time tools in resource-limited settings.

Clinical evaluations of machine-learning-enhanced microscopy have shown encouraging results. Yoon et al. [12] reported perfect sensitivity and specificity using fluorescence microscopy together with automated image analysis. Das et al. [13] evaluated EasyScan GO across 11 sites, yielding 91% sensitivity and 76% specificity. Holmström et al. [17] found that portable fluorescence microscopes combined with deep learning correlated strongly with manual assessments ($r > 0.99$) and performed comparably to PCR. Despite these gains, reliance on specialised optics or illumination can hinder adoption where such hardware is scarce.

Public datasets have been central to benchmarking. The NIH Malaria Dataset (~27,000 labelled single-cell images) has been widely used by Dwivedi et al. [22], Behera et al. [23], Asif et al. [24], and Boit and Patil [25], with reports of greater than 95% accuracy alongside strong sensitivity and specificity. These benchmarks facilitate a related comparison but also expose recurring challenges, class imbalance, stain variability, and dataset bias, that may limit generalising a routine clinical slide.

Overall, existing research demonstrates that automated systems can reach high accuracy, sensitivity and specificity. However, many models are validated under controlled conditions or depend on non-standard hardware and substantial computing resources, which constrain their scalability. Building on these backgrounds, this study presents a resource-efficient, image-based detection model that balances diagnostic performance with practical deployment needs, showcasing workable and high-performing prototypes, as well as a clinically valuable resource for low-resource settings.

3. METHODS

A structured machine learning approach is used in this study to create a model for finding malaria parasites in images of red blood cells. The approach has four main steps: gathering data, analysing and preprocessing it, training the model, and testing it.

3.1 Data Collection

This study worked with a publicly available dataset hosted on Kaggle, originally obtained from the NIH[1] repository [26]; consisting of 27,558 images of red blood cells, split almost evenly between parasitised (13,780) and uninfected (13,778) samples. Each image is a coloured blood smear showing whether malaria parasites are present. This dataset has been widely employed in prior research, making it a dependable benchmark for malaria detection studies that rely on image classification.

3.2 Data Preprocessing, Analysis and Visualisation

Preprocessing was a crucial stage in the model training process. The image pixel values were normalised to a 0–1 range using min–max scaling, which helps with faster convergence during training. The images were resized uniformly to 64 by 64 pixels in order to maintain consistency across the dataset. For classification purposes, labels were encoded numerically, with parasitised cells assigned the value one (1) and uninfected cells assigned zero (0). An exploratory analysis was conducted to understand the dataset's structure and distribution. This included checking for class balance and verifying the integrity of the images. Visual inspections helped identify any anomalies or patterns. An 80:20 ratio dataset was used to train the model. Stratified sampling ensured that each subset included a proportional amount of each class.

3.3 Model Training

The malaria detection model was trained using a CNN architecture (see FIGURE 2). Since CNNs can automatically learn and extract spatial characteristics straight from visual data, they are especially well-suited for image classification. The architecture designed for this study comprised multiple convolutional layers activated with ReLU functions, interspersed with pooling layers to

---

[1] https://ceb.nlm.nih.gov/repositories/malaria-datasets/

progressively reduce dimensionality. These layers were followed by fully connected dense layers responsible for the final classification.

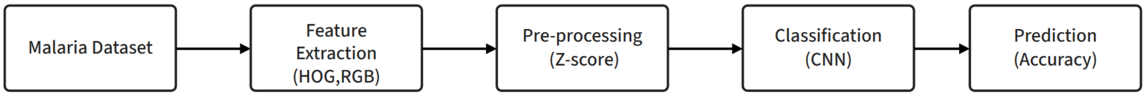
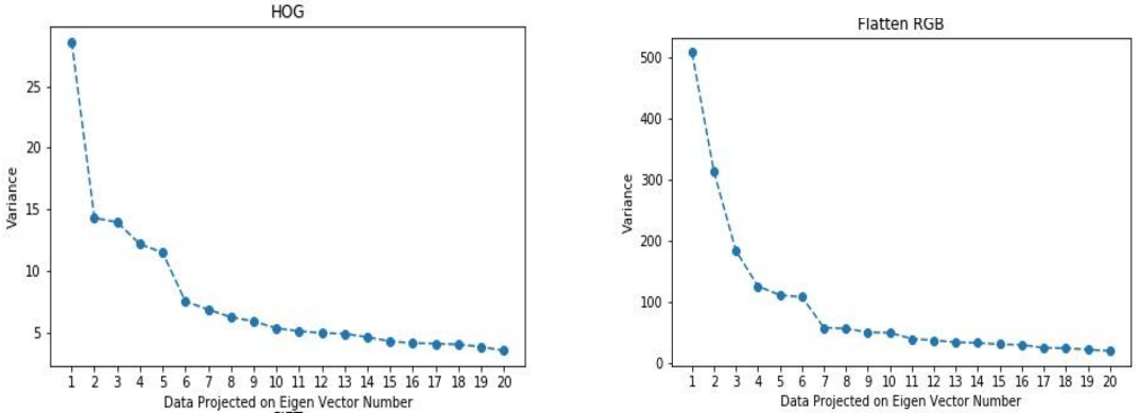


Figure 2: Model training pipeline

To mitigate overfitting, dropout regularisation was applied at selected layers. The model was compiled using the Adam optimiser, chosen for its efficiency in handling gradient descent. Model performance was primarily monitored using accuracy. Training was carried out over several epochs, with early stopping employed to prevent overtraining once validation performance ceased to improve.

FIGURE 3(a) and 3(b), present the variance of Z-score normalised data for HOG features and RGB features, respectively. These plots illustrate the distribution and consistency of the feature values, highlighting how the different feature extraction techniques contribute to the robustness of model training.



(a) (b)

Figure 3: (a) Variance of Z-score normalised data for Histogram of Oriented Gradients (HOG) features. (b) Variance of Z-score normalised data for Red-Green-Blue (RGB) colour channels

The model was cross-validated. Furthermore, for performance and generalisation enhancement, random cropping data augmentation was employed during the training; therefore, variations are better handled, and the model's robustness is improved through the data augmentation.

By combining HOG and RGB variant plots in the CNN, the hybrid model achieves higher accuracy and enhanced generalisation. It is crucial in medical image analysis as it captures geometric patterns and more complex patterns, enhancing semantic features through CNN automatic learning. HOG is a more interpretable feature descriptor, making the model more explainable.

3.4 Model Evaluation

The dataset for testing was used for model evaluation after training for performance based on — accuracy, precision, recall, and the F1-score—to provide a complete sense of its predictive strength. To better understand how the system classified samples, a confusion matrix was generated, showing the distribution of correct and incorrect predictions across infected and uninfected cases. Also, plotted are the ROC curves and the AUC, which offer a measure of how reliably the model distinguishes between the two classes. Looking at these results together allowed for a balanced assessment, drawing attention not only to overall accuracy but also to the important balance between sensitivity and specificity, both of which are central in medical diagnosis.

The summary of the critical details of VGG-19 architecture as employed in the model training are: Total layers (19), being 16 conv and 3 dense; total image size is 224 x 224 x 3 of RGB; filter size of 3 x 3; stride of 1. And the pooling entails MaxPooling 2 x 2, stride 2, the dense layers were of 512 units, 128 units sigmoid output, the dropout rates were based on 0.5 and 0.3. Also, it was based onAdam optimiser, learning at the rate of 0.001, with a batch size of 32, and the stopping criteria was an early stopping based on validation loss.

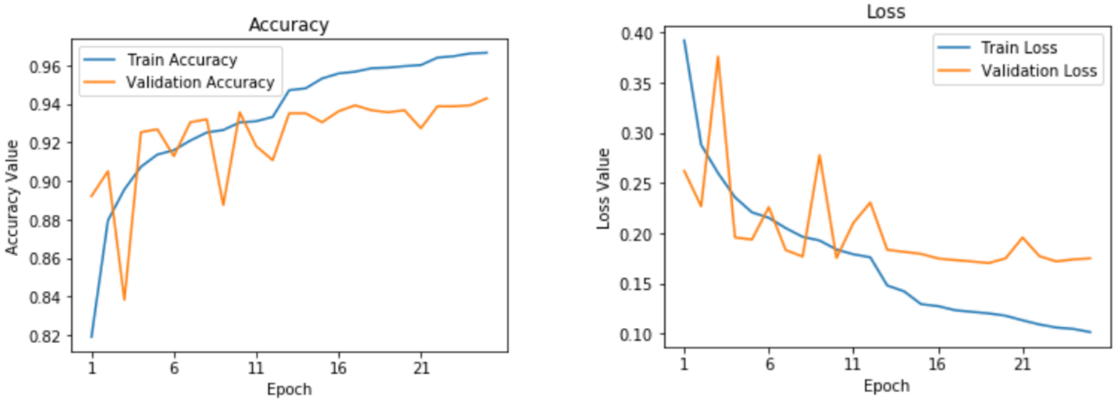
4. RESULTS AND DISCUSSION

Stained blood smear images were used to train the CNN model, which was built on the VGG19 architecture, to identify malaria parasites. With an equal number of parasitised and uninfected samples, the collection included 27,558 images. The model's strength and ability to generalise improved through preprocessing techniques like image scaling, RGB and HOG colour-space transformations, and Z-score normalisation. The Adam optimiser was used for training, and the loss function for binary cross-entropy. With an overall accuracy of 95%, the model showed a strong ability for prediction. The validation performance was almost the same as the training performance, and the accuracy and loss curves (FIGURE 4), confirmed steady convergence. This indicates that the model did not overfit and that it generalised well to new data. The plots display how performance improved over epochs, with loss consistently decreasing, showing effective learning.

To evaluate classification performance in more detail, we created ROC curves (FIGURE 5), for both classes using PCA and Linear Discriminant Analysis (LDA), which are two dimensionality reduction techniques. The ROC plots visualise the trade-off for the true positive rate and the false positive rate at varying thresholds. Larger Area Under the Curve (AUC) values indicate a stronger ability to discriminate. These results show the model's effectiveness in reliably distinguishing between parasitised and uninfected cells.

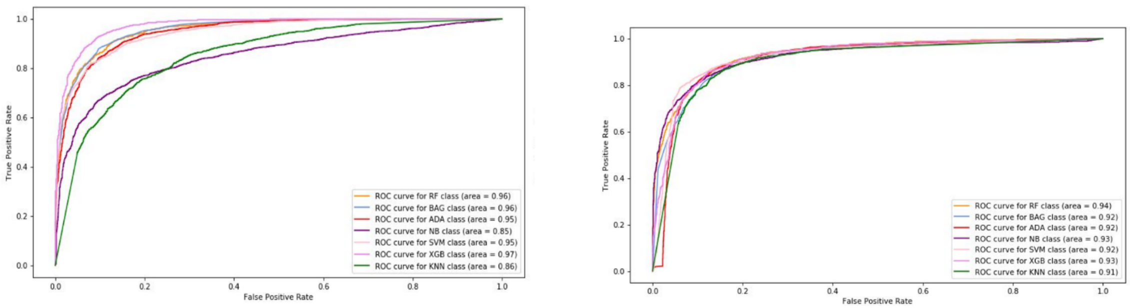
The confusion matrix (FIGURE 6), displayed high true positive and true negative counts. This provided further evidence of the model's effectiveness. It indicates that the system can accurately classify most infected and uninfected samples. Reducing incorrect diagnoses is essential for clinical applications.

Beyond testing, the model was implemented as a user-friendly Tkinter GUI-based system to show an example of practical usability. The application permits users to upload single-cell images and it returns immediate diagnostic feedback. As shown in FIGURE 7(a), the interface provides an



(a) (b)

Figure 4: (a) The accuracy plot of the CNN model. (b) The training and validation loss of the CNN model



(a) (b)

Figure 5: (a) Receiver Operating Characteristic (ROC) curves for classifiers using Principal Component Analysis (PCA) reduced features. (b) ROC curves for classifiers using Linear Discriminant Analysis (LDA) reduced features.

image-upload control; FIGURE 7(b), shows the classification output, clearly indicating *infected* or *uninfected*. This integration links research with real-world use by delivering a simple, reliable support tool for facilities with limited resources.

Table 1: Performance Comparison of the Classification Algorithms

Algorithm	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**
Naive Bayes	88.2	86.9	87.5	87.2
Support Vector Machine (SVM)	91.3	90.8	90.5	90.6
Convolutional Neural Network (CNN)	**95.0**	**94.7**	**95.3**	**95.0**

Performance on a widely used, real-world dataset shows that the practical value of AI-assisted diagnostics is limited in settings with few trained microscopists or limited infrastructure. While

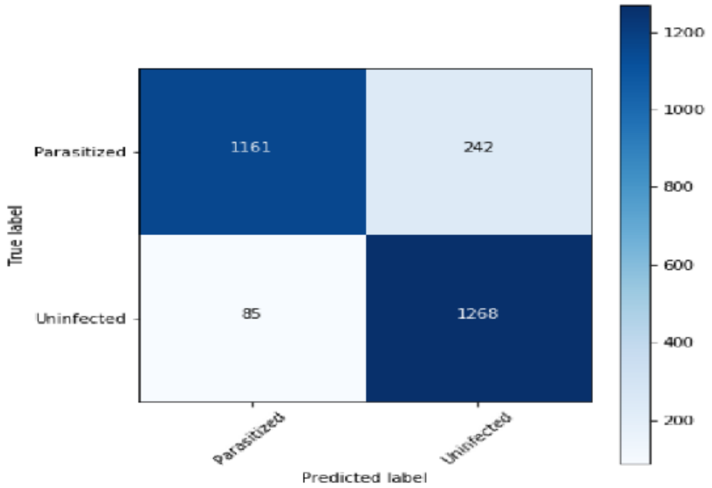
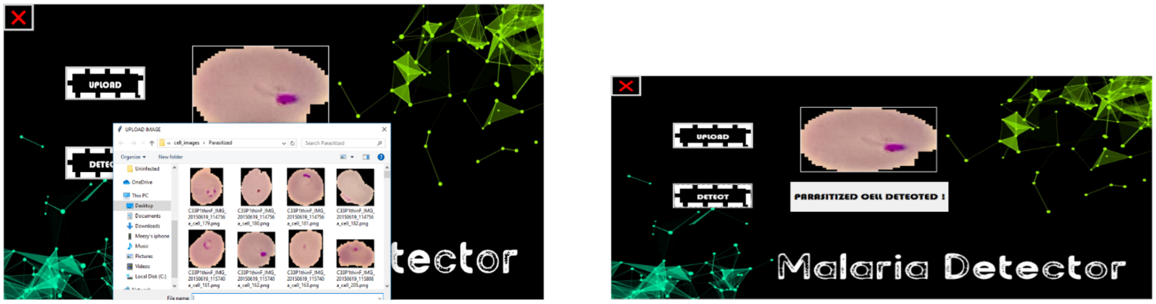


Figure 6: The Model Confusion Matrix



(a) (b)

Figure 7: (a) *Cell Image Upload View* – Interface for selecting and uploading malaria cell images for analysis. (b) *Detect View* – Output display showing the classification result (infected or uninfected) after processing the uploaded image through the trained CNN model.

light microscopy and rapid diagnostic tests (RDTs) remain the mainstays of field diagnosis, their sensitivity and specificity can be below expectations, particularly for non-falciparum and mixed-species infections that are frequently missed by microscopy and RDTs but detected by more sensitive molecular assays such as PCR [13, 15]. In this context, image-based CNNs offer a good complement to routine workflows, potentially improving accuracy and reducing turnaround time. The graphical interface lowers the barrier to use: uploading a cell image and receiving a result in near real time is feasible for health workers without specialist training.

Prior studies report encouraging performance from deep learning and image-processing pipelines. Shambhu et al. [19] achieved 96.02% accuracy with a customised CNN. Yoon et al. [12] demonstrated 100% sensitivity and specificity using fluorescence microscopy with automated analysis. Kundu and Anguraj [15] reported 90.33% accuracy using Local Derivative Radial Patterns and a Random Forest optimised by Particle Swarm Optimisation. However, many such systems depend on complex models or some specialised, costly approach, limiting their use for low-resource deployment.

As noted previously, traditional methods, such as microscopy and Rapid Diagnostic Tests (RDTs), are still widely used in most cases. The reason is that they are affordable and suitable for use in the field. However, these methods often struggle to detect non-falciparum species and mixed infections reliably. They are also likely to encounter issues such as human error and inconsistent slide quality [13, 15, 27]. These challenges highlight the need for practical, scalable, and reliable diagnostic systems.

In contrast, the present study employs a lightweight CNN with straightforward preprocessing and examines the role of dimensionality reduction (PCA and LDA) in enhancing analysis and interpretability. The approach achieves very competitive accuracy and a high ROC-AUC, comparable to the results reported by Oliveira et al. [11] and Holmström et al. [17], while being tuned for modest hardware, including laptops and embedded devices. Unlike object-detection models that prioritise real-time localisation (e.g., YOLOv4-tiny [18], Faster R-CNN [28]), our focus is on binary classification (infected vs. uninfected) with fewer parameters. While several works tend to use the NIH malaria dataset [22–25], our approach provides an evaluation of feature-space reduction and its impact on classification, highlighting the relationship between dimensionality and interpretability. This emphasis on efficiency and clarity enhances accessibility, practicality, and deployment potential in low-resource environments, addressing gaps noted in earlier studies.

To position the work with the state-of-the-art architecture, specifically the lightweight CNN architectures TABLE 2, offers a comparison

Table 2: Model Evaluation Comparison with Lightweight CNN Architectures

Model	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1- Score (%)**
VGG-19 CNN	95.0	94.7	95.3	95.0
MobileNetV2	96.3	96.1	96.4	96.2
EfficientNetB0	97.5	97.3	97.6	97.4

The improvement in performance is primarily due to the architectural efficiency and optimised feature extraction mechanisms employed by the newer models.

4.1 Limitations

Despite the promising results in this study, there are still areas for future work. These include improving the model's generalisability across different image qualities, extending support to multi-species malaria detection, and deploying lightweight versions of the model on mobile devices for offline use. Additionally, further clinical validation is necessary to establish the model's efficacy in real-world healthcare workflows. Furthermore, the model was tested on one dataset (NIH malaria dataset), though it is a standard one for the purpose; more datasets are better for enhanced comparison. Also, there are ethical concerns in clinical Artificial Intelligence adoption.

5. CONCLUSION

This study developed and tested a machine learning model for detecting malaria parasites in stained blood smear images. The system has an accuracy of 95%, which points to its promise as a dependable diagnostic tool. Beyond accuracy, its easy-to-use and modest computational status suggest that it could be integrated into everyday diagnostic practice, to shorten turnaround times, lower costs, and support clinical decision-making. It should be noted that the approach is designed with low-resource environments in mind, where skilled microscopists and well-equipped laboratories are often unavailable. Providing health workers with such an accessible tool has the potential to make a meaningful contribution to malaria control and treatment.

Looking ahead, several areas of improvement remain. The model will need to be tested on images with different staining qualities and under varied field conditions to assess its robustness. Extending its capacity to recognise multiple Plasmodium species and different stages of infection would further increase its clinical value. We are also looking at designing a lightweight version of the system for mobile or portable use, which could help rural and hard-to-reach communities use it easily. Finally, we also noted a need for clinical validation and regular updates with new data to ensure the tool remains accurate and reliable over time.

6. ACKNOWLEDGMENT

The authors thank the University for supporting this research and Omiete Brown-West for the implementation.

References

[1] López-Otín C, Kroemer G. Hallmarks of Health. Cell. 2021;184:33-63.

[2] Somoskovi A, Ahmedov S, Salfinger M. It Is Health That Is Real Wealth & Not Pieces of Gold & Silver. Indian J Med Res. 2013;137:437-439.

[3] Yaqoubi A, Sabouhi F, Bozorgi-Amiri A, Amalnick MS. A Light Robust Optimization Model for a Hierarchical Healthcare Network. Kybernetes. 2024;53:990-1014.

[4] Oladipo HJ, Tajudeen YA, Oladunjoye IO, Yusuff SI, Yusuf RO, et al. Increasing Challenges of Malaria Control in Sub-Saharan Africa: Priorities for Public Health Research and Policymakers. Ann Med Surg. 2022;81:104366.

[5] https://www.who.int/publications/i/item/97892401044.

[6] Venkatesan P. The 2023 Who World Malaria Report. Lancet Microbe. 2024;5:e214.

[7] Okereke PU, Ayodeji BM, Bolarinwa RT, Ayodeji OD, Popoola IO, et al. Assessing the Influence of Rapid Diagnostic Test (Rdt) Accuracy on Malaria Misdiagnosis and Antimalarial Resistance in Nigeria. Ann Med Surg (Lond). 2025;87:658-662.

[8] Anjorin ET, Olulaja ON, Osoba ME, Oyadiran OT, Ogunsanya AO, et al. Overtreatment of Malaria in the Nigerian Healthcare Setting: Prescription Practice, Rationale, and Consequences. Pan Afr Med J. 2023;45:111.

[9] Abdul-Rahman T, Ajetunmobi OA, Bamigbade GB, Ayesiga I, Shah MH, et al. Improving Diagnostics and Surveillance of Malaria Among Displaced People in Africa. Int J Equity Health. 2025;24.

[10] Maturana CR, De Oliveira AD, Nadal S, Bilalli B, Serrat FZ, et al. Advances and Challenges in Automated Malaria Diagnosis Using Digital Microscopy Imaging With Artificial Intelligence Tools: A Review. Front Microbiol. 2022;13:1006659.

[11] de Souza Oliveira A, Costa MG, Barbosa MD, Costa Filho CF. A New Approach for Malaria Diagnosis in Thick Blood Smear Images. Biomed Signal Process Control. 2022;78:103931.

[12] Yoon J, Jang WS, Nam J, Mihn DC, Lim CS. An Automated Microscopic Malaria Parasite Detection System Using Digital Image Analysis. Diagnostics. 2021;11:527.

[13] Das D, Vongpromek R, Assawariyathipat T, Srinamon K, Kennon K, et al. Field Evaluation of the Diagnostic Performance of Easyscan Go: A Digital Malaria Microscopy Device Based on Machine-Learning. Malar J. 2022;21:122.

[14] Maturana CR, de Oliveira AD, Nadal S, Serrat FZ, Sulleiro E, et al. Imaging: A Novel Automated System for Malaria Diagnosis by Using Artificial Intelligence Tools and a Universal Low-Cost Robotized Microscope. Front Microbiol. 2023;14:1240936.

[15] Kundu TK, Anguraj DK. Optimal Machine Learning Based Automated Malaria Parasite Detection and Classification Model Using Blood Smear Images. Traitement Signal. 2023;40:91-99.

[16] Manning K, Zhai X, Yu W. Image Analysis and Machine Learning-Based Malaria Assessment System. Digit Commun Netw. 2022;8:132-142.

[17] Holmström O, Stenman S, Suutala A, Moilanen H, Kücükel H, et al. A Novel Deep Learning-Based Point-Of-Care Diagnostic Method for Detecting Plasmodium Falciparum With Fluorescence Digital Microscopy. In: Advances in Medical Imaging, Detection, and Diagnosis. Jenny Stanford Publishing. 2023:979-1000.

[18] Shewajo FA, Fante KA. Tile-Based Microscopic Image Processing for Malaria Screening Using a Deep Learning Approach. BMC Med Imaging. 2023;23:39.

[19] Shambhu S, Koundal D, Das P. Deep Learning-Based Computer Assisted Detection Techniques for Malaria Parasite Using Blood Smear Images. Int J Adv Technol eng explor. 2023;10:990.

[20] Chibuta S, Acar AC. Real-Time Malaria Parasite Screening in Thick Blood Smears for Low-Resource Setting. J Digit Imaging. 2020;33:763-775.

[21] Bae CY, Shin YM, Kim M, Song Y, Lee HJ, et al. Embedded-Deep-Learning-Based Sample-To-Answer Device for On-Site Malaria Diagnosis. Front Bioeng Biotechnol. 2024;12:1392269.

[22] Dwivedi S, Pandey K, Shubham K, Pandey OJ, Tripathi AM, et al. MalaNet: A Small World Inspired Neural Network for Automated Malaria Diagnosis.IEEE Trans Artif Intell. 2025;6:2486-2497.

[23] Behera N, Das S, Singh AP, Swain AK, Rout M. CNN-Based Medical Image Classification Models for the Identification of Pneumonia and Malaria. In: International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC). IEEE. 2024:1-6.

[24] Asif S, Khan SU, Zheng X, Zhao M. MozzieNet: A Deep Learning Approach to Efficiently Detect Malaria Parasites in Blood Smear Images. Int J Imaging Syst Technol. 2024;34:e22953.

[25] Boit S, Patil R. An Efficient Deep Learning Approach for Malaria Parasite Detection in Microscopic Images. Diagnostics. 2024;14:2738.

[26] https://www.kaggle.com/iarunava/cell-images-for-detecting-malaria

[27] Vatandoost H, Hanafi-Bojd AA, Nikpoor F, Raeisi A, Abai MR, et al. Situation of Insecticide Resistance in Malaria Vectors in the World Health Organization of Eastern Mediterranean Region. Toxicol Res . 2022;11:1-21.

[28] Nakasi R, Mwebaze E, Zawedde A, Tusubira J, Akera B, et al. A New Approach for Microscopic Diagnosis of Malaria Parasites in Thick Blood Smears Using Pre-Trained Deep Learning Models. SN Appl Sci. 2020;2:1255.