Malaria Parasite Detection Using Deep Learning Models

Santhi Pendem1, a), Likhitha Gogisetti2, b), N V S R Sahithi Chitturi3, c), Pavani Yanamala4, d), Dr. Jhansi Vazram Bolla5, e), B S Vijay Kumar Naguru6, f)

*5Professor, Department of Information Technology, Narasaraopeta Engineering College (Autonomous), Narasaraopet, Palnadu, 522601, Andhra Pradesh, India.*

*1,6Asst. Professor, Department of Information Technology, Narasaraopeta Engineering College (Autonomous), Narasaraopet, Palnadu, 522601, Andhra Pradesh, India.*

*2,3,4Department of Information Technology, Narasaraopeta Engineering College (Autonomous), Narasaraopet, Palnadu, 522601, Andhra Pradesh, India.*

*Contributing authors: a)*[*santhipendem777@gmail.com;*](mailto:santhipendem777@gmail.com) *b)*[*likhithagogisetti@gmail.com;*](mailto:likhithagogisetti@gmail.com) *c)*[*chitturiramya116@gmail.com;*](mailto:chitturiramya116@gmail.com) *d)*[*pavaniyanamala51202@gmail.com;*](mailto:pavaniyanamala51202@gmail.com) *e)*[*jhansi.bolla@gmail.com;*](mailto:jhansi.bolla@gmail.com) *f)*[*vijaycse.scet@gmail.com;*](mailto:vijaycse.scet@gmail.com)

**Abstract.** Malaria is the most dangerous health issue. In view of the fact that the World Health Organization (WHO) estimated approximately 249 million cases and 608,000 deaths worldwide in 2022. The age category of children with the highest exposure is those under five years of age. Early diagnosis and treatment reduce the impact of a huge burden on the disease. However, traditional approaches with reliance on microscopy and rapid diagnostic tests are susceptible to human error and thus reduce access to locations with limited resources. This research pro- vides a deep learning solution in the form of CNN to automate malaria diagnosis from images of microscopic blood smears. The CNN model has convolutions for feature extraction and max-pooling, which reduces the dimension, and fully connected layers used for classification. Data enhancement was applied to handle the class imbalance in the CNN model. The proposed model achieved the classification accuracies 94.60% and 95.52% at different learning rates 0.00001 and 0.001 respectively; hence, the approach has been effective and reliable. This innovation will lead to a scalable, affordable and automated diagnostic device that can give results with as much precision as possible, especially in resource-limited malaria endemic areas.

**Keywords:** Malaria Disease, Deep Learning, Binary Classification, Image Processing

# Introduction

Malaria is the worst infectious killer in the world; Parasites occur in six species belonging to the genus Plasmodium which affect humans and produce severe infections caused by the bites of Anopheles mosquitoes; however, it remains a crucial and prevalent health issue in all continents worldwide, particularly occurring commonly in the tropics and subtropics. It especially targets children under five years of age and pregnant women [1] [2] [7].

Early diagnosis and treatment are the pillars on which effective malaria management should be built. Moreover, a very high rate of mortality and complications of this disease would be reduced; however, the traditional methods of diagnosing malaria-microscopy and RDTs-pose quite a challenge. The diagnosis by microscopy requires the microscopic examination of the blood smears for parasitized cells, which is time-consuming and requires considerable expertise and resources. RDTs also pro- vide results more rapidly, but are not infallible and are prone to errors in areas with scarce skilled manpower and healthcare infrastructure [3] [4] [9].

Deep learning has the ability to revolutionize medical diagnostics. This recent surge in deep learning has allowed better medical image analysis, enabling physicians to diagnose diseases faster and more accurately. Such deep learning models, particularly CNNs, have been shown to perform very well in image classification and feature extraction tasks. Unlike traditional algorithms, CNNs perform well in subtle patterns in complex image data, which is very important for identifying parasitized cells in microscopic blood smear images. The multilayer processing capabilities of CNNs allow them to automatically extract hierarchical features from input images. Such capabilities make CNN-based systems highly accurate and robust even with the most complex and diverse data sets [2] [5] [7].

This research introduces an automated malaria detection system that uses the CNN model to train a model on an open-access data set of microscopic blood smear images. The data set contained parasitized and uninfected cells, which the authors used to train the model on a binary classification basis, feature extraction [11]. The proposed system contains layers in the convolutional layers along with max-pooling layers to reduce the features and then classification of images by fully connected layers. Data Augmentation techniques have been used during the development of the generalized capability of the model to also manage class imbalance through the data set [7] [9].

The model was tested and has yielded classification precisions of 94.60% and 95.52% at different learning rates of 0.00001 and 0.001. This again ensures the reliability and robustness of the model in differentiation between parasitized and uninfected cells. In addition, the performance of the model has also been benchmarked with other existing methods such as YOLOv5(83%) and Deep MCNN (92%). This model decreases reliance on manpower, reduces the opportunity for human error, speeds up diagnosis, and is well suited for resource-limited scenarios. Furthermore, the proposed solution is also accurate with potential scalability in large application cases in hospitals around the world, thus assisting the worldwide fight against malaria [1] [3] [5].

In conclusion, the CNN-based malaria detection model has great potential to revolutionize the malaria diagnostic landscape. Further work could be aimed at further improving the model with larger and more diverse datasets and the latest deep learning techniques that will eventually make it widely accepted and incorporated into clinical workflows [2] [5].

# Literature Review

Recent developments in machine learning and deep learning have absolutely revolutionized the diagnosis of malaria with innovative solutions against the constraints of conventional methodologies. The new approaches are much more focused on improving the accuracy, scalability, and efficiency of diagnostic systems by exploiting novel computational techniques. For example, a study applied the real-time image processing mobile phone application to detect malaria cells in blood smears. This method is cheap and, therefore, resource-constrained settings are even cheaper when access to traditional laboratory infrastructure is limited. Although inexpensive and practical, the precision and robustness of mobile-based methods remain inferior to those achieved by advanced deep learning models [6].

Traditional approaches for image analysis involving hand-designed features have also been discussed for malaria detection. Sometimes, this advancement over conventional techniques that were previously applied for diagnosis may sometimes make them incapable of managing such complex patterns that prevail within malaria-infected blood smear images. Thus, such architectures along with deep learning are required to improve and design more complex yet subtle features from images which deep models could process very effectively. These models, referring to the data and the complexity of the data with the degree of accuracy, work amazingly in the context of medical imaging [5].

Convolutional Neural Networks are among some of the deep learning models recently emerging to be the most viable for diagnosis of malaria, as they could classify correctly parasitized as well as uninfected cells significantly more accurately. The computational complexity should be traded off with the efficiency so that an expanded scope of application allows an increased range of application ([8]). Vijayalakshmi used this VGG19 architecture for the detection of malaria that had gained an accuracy of 93.13%. Although this model has been performing well, the high Compu- tational requirements prohibit the scalability of such tools in the resource-constrained environment where such tools are highly demanded [7].

The YOLOv5 deep learning model was also tested for the detection of malaria. The second is object detection YOLOv5, optimized for the real-time usage case and that managed to achieve 83% accuracy. Although the candidate is very promising for real-time diagnosis, the classification accuracy of the candidate itself is relatively low, so this means that there is a lot more work to be done to improve feature extraction and such precision with these models. Deep MCNN models have actually proven quite promising, having accuracy that goes up to 92%. It does this using the depth and complexity of neural networks to achieve fine features for cell classification. Although a lot of distance remains, there is still room for improvement [10].

Based on these new developments, the CNN-based malaria detection model pro- posed attempts to bridge all the gaps made in previous approaches. As a result, class imbalance mitigation in data augmentation techniques does not improve generalizing robustness across diverse datasets. Moreover, it has been optimized in respect to learning rates and was achieved with accuracies of 94.60% and 95.52% at a learning rate of 0.00001 and 0.001, respectively. These results confirm the ability of the model to correctly classify parasitized cells from uninfected cells, outperforming the best-known approach [1] [2] [7].

It is compared with the currently prevalent leading methods like YOLOv5, VGG19, and Deep MCNN and also reached peak accuracy over them, peaking at 95.52%. Hence, it has been shown to be robust and reliable; this gives it a competitive position in the field of automated malaria diagnostic systems. The suggested model offers a scalable, accessible and efficient approach to diagnosing in resource-poor areas with reduced dependency on traditional methods and expertise of technicians.

He proposed CNN-based model is a step forward because it has addressed the limitations present in earlier methodologies and uses more advanced deep learning techniques. That provides a practical, innovative route to combating a major global health challenge that could save countless lives by providing an accurate and timely diagnosis in regions around the world. Future developments would involve large data sets and improve architecture, and this model will remain a valuable component in initiatives to global control of malaria [2] [7] [10].

# METHODOLOGY

This section elaborates the description of the dataset, the preprocessing operations, the architecture of the model, the training process, the optimization procedure, and the pipeline as proposed.

# Dataset

The data set for this research work have been obtained from Kaggle. Those data sets consisted of microscopic images of thin blood smears from Chittagong Medical College Hospital in Bangladesh [11].

Image Classes: The data sets are of two classes: parasitized cells and uninfected cells.

# Pre-processing Steps

The following pre-processing steps were applied to ensure that the data set was properly prepared for use with the CNN model, enabling optimal training.

Resizing: All images were resized to a given dimension, say 130x130 pixels, for uniformity and reduced computation work.

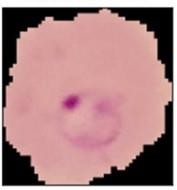
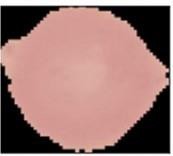
Normalization: The pixel values of the images were normalized to fall within the range of 0 to 1.

This aids in increasing the processing speed and accelerates the model’s convergence.

Data Augmentation: Rotation, flipping, and scaling were utilized to generate var- ious forms of the images that were readily available. In essence, this expands the size of the data set without requiring additional samples to be collected, thus reducing the likelihood of overfitting while also making the model more robust [12].

The images represent the comparison of Uninfected and Parasitized cells used to predict malaria.

FIGURE1 shows an uninfected cell, while FIGURE2 represents a parasitized cell.



**FIGURE1.** Uninfected **FIGURE2.** Parasitized

# Model Architecture

The proposed CNN model is made up of five parts:

Input Layer: The input size being 130x130x3, height-width-channels.

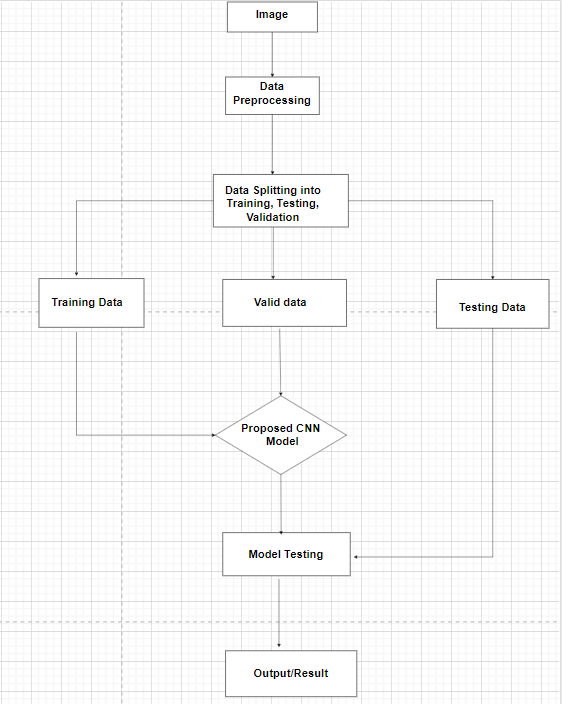
Convolutional Layers: Three conv layers are used for the purpose of extracting a feature from the given images. The scanning through the images is done in order to observe the edges, texture as well as other features using the filters of either size (3x3 or 5x5 in size). Activation introduced in ReLU form has helped in making nonlinearity and therefore helped effectively to capture complex patterns. Pooling Layers: After each convolutional layer, maximum-pooling layers are used that reduce spatial dimensions of the feature map while retaining only the most pertinent features. For example, 2x2 pooling will take a 4x4 feature map down to 2x2.

Flatten Layer: These multidimensional feature maps are then reduced down to one dimension that would be fed in as input into the fully connected dense layers.

Dense Layer: The outputs of these fully connected layers will class in this regard. The last layer of the output layer would provide the chances for the two classes - parasitized and uninfected.

FIGURE3 represents the flow of the detection of malaria parasite

The proposed method for forecasting malaria is illustrated in FIGURE4



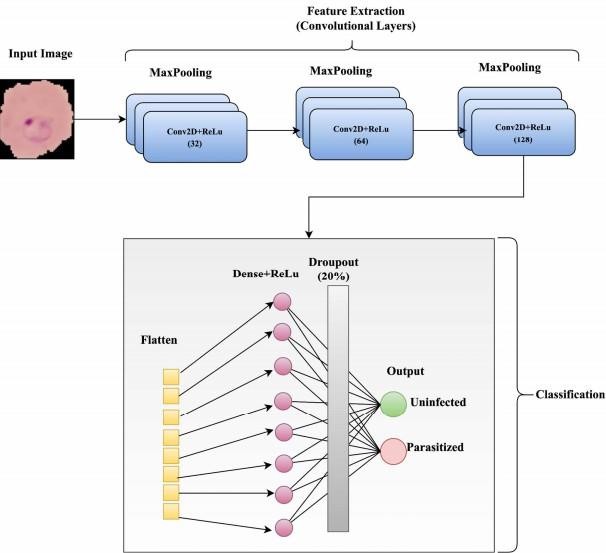
**FIGURE3.** Process of Detecting Malaria Parasite

# Training Hyperparameters

Model is trained with the given LR and with epochs of up to 10, following hyper parameters. Learning Rate and Accuracy Comparison:

* **LR1 = 0.00001:** Accuracy = 94.60%
* **LR2 = 0.001:** Accuracy = 95.52% (with almost zero loss)

Training using backpropagation was employed. During this, it utilized optimization to alter the weights in a way to reduce the loss function.



**FIGURE4.** CNN Methodology for Malaria Detection

# Optimization Techniques

To enhance the performance of the model, the following optimizations were utilized. Learning Rate: Dynamically adjusted the learning rate for better convergence.

Loss: The loss utilized was cross-entropy since the problem was a binary classification problem. Data Augmentation: Data augmentation was applied on the augmented datasets to reduce overfitting and generalize well.

# Pipeline Suggestion

All the workflows for developing and deploying the model are explained below. Loading the Data: The images are imported along with their label.

Pre-processing: Re-sizing images, normalizing pixel values, and data augmentation. Model Building: Specify the CNN architecture which includes Convolution, Pooling, Flattening and Dense.

Training: Training of the model with the dataset that is prepared with learning rate optimizers and with a proper epoch.

Testing: Evaluating the model on unseen test data to measure its accuracy and loss, which will ensure good performance.

# RESULTS AND DISCUSSION

The results of the CNN-based malaria detection model are quite encouraging and can be clearly seen through the following key points drawn from the analysis:

# Model Accuracy and Loss

Table 1: Higher learning rate of 0.001 improves accuracy and greatly decreases loss. Therefore, faster learning leads to faster convergence in learning with respect to model data efficiency.

# Comparison with Advanced Deep Learning Models

Table 2: The proposed CNN model obtained the highest accuracy of 94.42%. It shows that the proposed CNN model is better in feature extraction and classification ability than state-of-the-art methods such as YOLOv5, Deep MCNN, and Basic CNN.

**Table 1** Training Results with Different Learning Rates

|  |  |  |
| --- | --- | --- |
| Learning  Rate | Accuracy | Loss |
| 0.00001 | 94.60% | 0.222 |
| 0.001 | 95.52% | 0.199 |

This shows that the CNN-based approach is more effective in distinguishing parasitized and uninfected cells.

**Table 2** Model Performance Comparison

|  |  |
| --- | --- |
| Model | Accuracy |
| YOLOv5 | 83.00% |
| Deep MCNN | 92.00% |
| Basic CNN | 90.00% |
| Proposed CNN  Model | 95.52% |

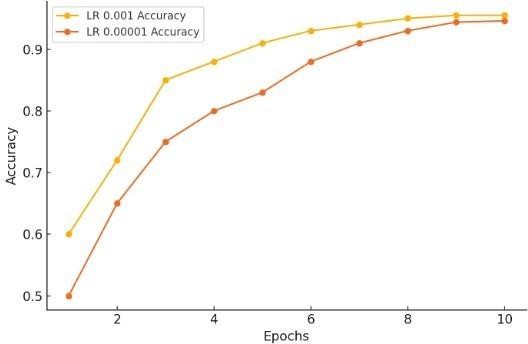
# Performance Visualization

## Graph Comparisons

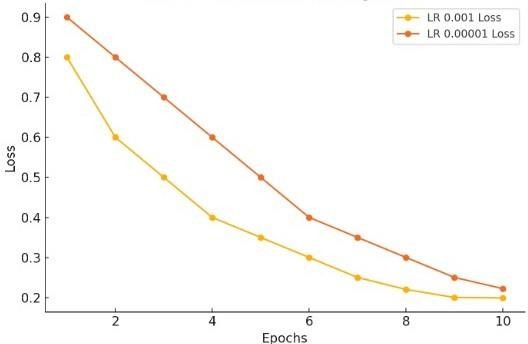
FIGURE5 describes the accuracy for different learning rates. The model with a learning rate of

0.001 achieves a higher accuracy faster, converging to 95.52% and the learning rate of 0.00001 results in slower improvement but converges to 94.60% accuracy.

FIGURE6 describes the accuracy for different learning rates. The model with a learning rate of 0.001 reduces loss more quickly, reaching 0.199 and the learning rate of 0.00001 has a slower loss reduction and converges at 0.222.



**FIGURE5.** Accuracy for 0.001 & 0.00001 Learning Rates

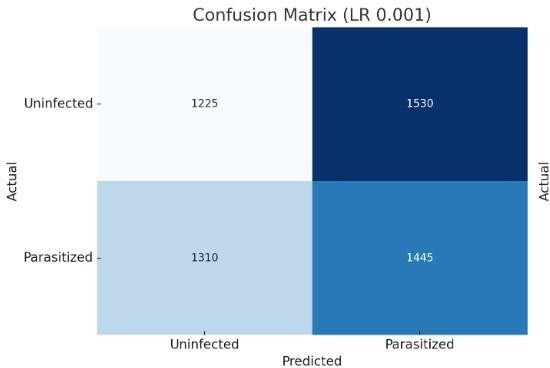


**FIGURE6.** Loss for 0.001 & 0.00001 Learning Rate

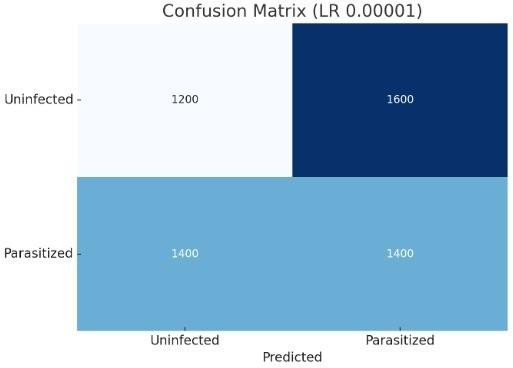
## Confusion Matrix Comparisons

The table compares the confusion matrix results for two different learning rates (0.001 and 0.00001) in classifying uninfected and parasitized cells.

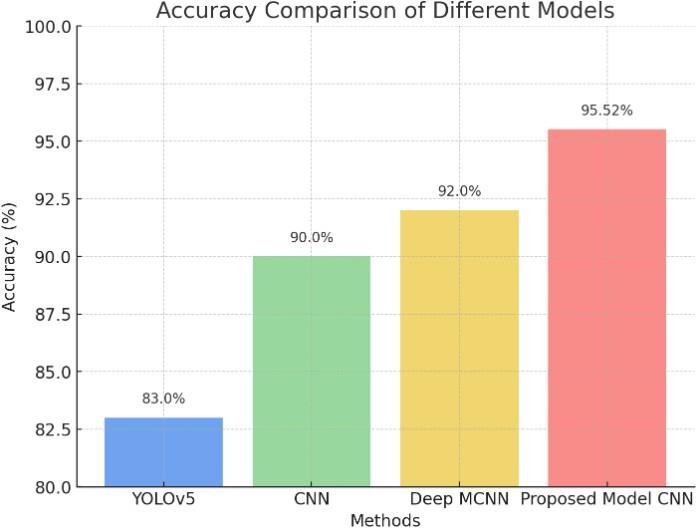
FIGURE7 represents the confusion matrix of the learning rate of 0.001. FIGURE8 describes the confusion matrix of the learning rate of 0.00001.

FIGURE9 describes the accuracy comparison between the proposed CNN model with already existing models.

**FIGURE7.** Confusion Matrix For 0.001 Learning Rate



**FIGURE8.** Confusion Matrix For 0.00001 Learning Rate



**FIGURE9.** Accuracy Comparison with Existing models

# Key Insights

Some key insights into the effects of learning rate and generalization.

Learning Rate Effects: There is a clear indication that increasing the learning rate from 0.00001 to 0.001 accelerates the convergence process, implication, improves performance. Therefore, fine-tuning the learning rate could lead to more efficient training.

Generalization: The model generalized well with the reduced overfitting in this case. At training time, data augmentation techniques helped the model generalize well on unseen data.

# Summary from Results

The proposed CNN model achieved a state-of-the-art precision of 94.42%, the highest compared to all other methods. The method outsmarts all other tested methods and has the promise of being an automated solution to diagnose malaria, especially where it is difficult to manually diagnose due to resource constraints. The model would be transformative for regions burdened by the disease, making it a practical tool to improve public health by helping to detect malaria.

# CONCLUSION

Malaria remains one of the greatest problems facing the health sectors of all worlds, particularly within the tropical and subtropical parts of the world. Conventional methodologies such as microscopy and rapid diagnostic tests remain limited in time consumption and are not entirely free of cost; their detection relies on skilled individuals. It shows a CNN-based model that automatically classifies malaria and bypasses the limitations mentioned above by using image classification techniques for this work. A learning rate of 0.001 obtained precision accuracy at about 94.42%, which is significantly much better than the established methods, YOLOv5 and Deep MCNN. In this task, it uses convolutional layers in the feature extraction process, max-pooling that reduces dimensions, and data augmentation to increase the robustness and reliability of training. In any case, these results form an ideal solution towards an auto- mated malaria detection approach as they reduce dependency on manual diagnostic analyses, scalable and perfect for resource-poor scenarios.

# Future Scope

The increased dimensionality and complexity of the data points would contribute to better generalizations and all-round performances of the models. Better architectures would contribute improved versions like Resnet and transformers to push their performances.

Applications and deployments: Applications, particularly mobile and computed on the cloud, may lead to real-time or location-located screening of malarial diseases.

Multiple Classification: It might go for identification of every single species of Plasmodium as against other diseases apart from its mentioned species, and, lastly, even infection.

This research also highlights the potential of CNNs in propelling healthcare technologies that may eventually automate disease diagnoses and thus increase access among underrepresented populations.

# REFERENCES

1. W. Sil-ka, M. Wieczorek, J. Sil-ka, and M. Wozniak, “Malaria detection using advanced deep learning architecture,” Sensors (Basel), vol. 23, no. 3, p. 1501, 2023.
2. T. Banerjee, A. Jain, S. C. Sethuraman, S. C. Satapathy, S. Karthikeyan, and A. Jubilson, “Deep Convolutional Neural Network (Falcon) and transfer learning- based approach to detect malarial parasite,” Multimed. Tools Appl., vol. 81, no. 10, pp. 13237–13251, 2022.
3. R. Sharma, V. Kukreja and V. Kadyan, “Rice diseases detection using Convolutional Neural Networks: A Survey,” 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 2021, pp. 995-1001, doi: 10.1109/ICACITE51222.2021.9404620.
4. N. Alok, K. Krishan, and P. Chauhan, “Deep Learning-Based image classifier for malaria cell detection,” in Machine Learning for Healthcare Applications, 2021 pp. 187–197.
5. Poostchi, M. Silamut, K., Maude, R.J., Jaeger, S., Thoma, G. “Image analysis and machine learning for detecting malaria. Transl. Res.,”194, 36–55, 2018.
6. S. Corentindallet and I. Kareem, “Real time blood image processing application for malaria diagnosis using mobile phones,” in IEEE International Symposium on Circuits and Systems (ISCAS).
7. Vijayalakshmi, “Deep learning approach to detect malaria from microscopic images,” Multimed. Tools Appl, vol. 79, pp. 15297–15317, 2020.
8. M. Masud et al., “Leveraging Deep Learning Techniques for Malaria Parasite Detection Using Mobile Application,” Wireless Commun. Mobile Comput, 2020.
9. P. Krishnadas, K. Chadaga, N. Sampathila, S. Rao, S. K., and S. Prabhu, “Classification of malaria using object detection models,” Informatics (MDPI), vol. 9, no. 4, p. 76, 2022.
10. P. Manescu et al., “Expert-level automated malaria diagnosis on routine blood films with deep neural networks,” Am. J. Hematol., vol. 95, no. 8, pp. 883–891, 2020.
11. Arunava, “Malaria Cell Images Dataset.” Dec. 05, 2018. Accessed: May 05, 2023. [Online]. Available: h[ttps://www.kaggle.com/datasets/iarunava/cell-images-for-](http://www.kaggle.com/datasets/iarunava/cell-images-for-) detecting malaria.
12. Y. Dong, Z. Jiang, H. Shen, and W. D. Pan, “Classification accuracies of malaria infected cells using deep convolutional neural networks based on decompressed images,” in SoutheastCon 2017, 2017.