## 1.0 Introduction

In recent years, the integration of artificial intelligence (AI) and its subfield, deep learning, has initiated a paradigm shift in technological progress, with significant ramifications across diverse areas. The field of medical diagnostics has been significantly transformed by the implementation of AI-driven methodologies, leading to advancements in disease detection and diagnosis(1). The early and precise identification of diseases such as malaria poses a notable medical obstacle, persisting as a substantial public health issue, particularly in countries characterized by inadequate healthcare infrastructure and limited resources(2).

The application of artificial intelligence (AI), specifically deep learning, in the field of medical image analysis has emerged as a significant advancement in the interpretation of complex patterns and features present in medical data(3). The incorporation of artificial intelligence (AI) technologies into healthcare systems has the capacity to surpass human limitations and optimize diagnostic procedures. One of the notable difficulties addressed by artificial intelligence (AI) in the field of healthcare pertains to the identification of malaria. This difficulty is particularly significant due to the widespread occurrence of malaria and the possibility for image-based methodologies to effectively handle it. (4).

Malaria, a disease caused by the Plasmodium parasite, continues to pose a significant global health concern, with a particular impact on populations residing in tropical and subtropical climates (5). According to the World Health Organization (WHO), around 50% of the global population is susceptible to malaria, with an estimated 247 million cases documented worldwide. Alarmingly, over 70% of individuals affected by this disease succumb to its devastating consequences. It is worth noting that malaria exhibits a particularly high prevalence in the African continent(5)(6)(7). Traditional diagnostic methods for malaria, such as microscopy, can be labor-intensive, time-consuming, and prone to human error(8). Leveraging deep learning techniques to detect malaria from cell images presents a paradigm shift in disease diagnosis. It not only accelerates the diagnostic process but also enhances the accuracy of detection, enabling early intervention and treatment.

The importance of this research topic lies in its potential to reshape the healthcare landscape by combining AI's computational power with medical expertise. The successful implementation of AI-driven malaria detection could potentially save countless lives by enabling early and accurate diagnosis, leading to timely interventions and treatments.

The central research question of this study is: How can deep learning techniques be leveraged to accurately detect malaria infection from cell images, and how does this approach compare to traditional diagnostic methods in terms of accuracy, speed, and scalability?

The anticipated outcome of this research is a comprehensive evaluation of the efficacy of deep learning models in malaria detection using cell images. The study aims to demonstrate that deep learning

algorithms can achieve comparable or superior accuracy to traditional methods while drastically reducing the time required for diagnosis. Furthermore, the research will shed light on the generalizability and robustness of the proposed approach across diverse populations and geographical regions.

This study seeks to establish the viability of using deep learning for malaria detection, not only contributing to the advancement of medical diagnostics but also offering a blueprint for the integration of AI in addressing other healthcare challenges. By bridging the gap between AI and medical science, this research endeavors to make significant strides in improving healthcare accessibility and outcomes on a global scale.

## 2.0 Background

In recent years, there have been notable breakthroughs in the field of malaria diagnosis through the utilization of deep learning techniques and medical imaging. Numerous research initiatives have been conducted to investigate the potential of artificial intelligence (AI)-based solutions in improving the precision and effectiveness of malaria diagnosis. A thorough examination of the extant literature offers helpful perspectives on the scope of relevant studies and their relevance to the present research endeavor(9).

A considerable body of research has been dedicated to the application of deep learning methodologies in the context of malaria detection through the analysis of cellular images. Convolutional Neural Networks (CNNs), a major subclass of deep learning models, have gained significant traction in various domains owing to their inherent capacity to autonomously acquire pertinent features from image data(10). Researchers have successfully devised convolutional neural network (CNN) models that effectively categorize red blood cells as either infected with malaria or uninfected. These models have demonstrated notable levels of accuracy in accurately differentiating between parasitized and healthy cells(11). Similar to this project, the existing research aims to create models that leverage deep learning for malaria diagnosis using cell images as the primary data source.

Nevertheless, there are differences that emerge in relation to the approach employed and the application of datasets. While several research projects utilize publicly accessible datasets such as the Malaria Cell Image Dataset, others investigate datasets that are specialized to particular domains in order to address distinct variances in cell morphology and staining procedures. Moreover, the choice of CNN architectures and preprocessing techniques varies across studies, leading to differing performances and implications for real-world deployment.

The related works done by (12), (13) and (14) directly aligns with my project in its overarching goal of employing deep learning(CNN) model models for malaria detection using cell images. However, the distinctive contributions of this research lie in potentially introducing novel methodologies, adapting existing approaches to different datasets, or enhancing the robustness and generalizability of the models. For example, this research will propose an innovative CNN architecture, implement advanced data augmentation techniques and hyperparameter tunning to enhance our model result and experiment with transfer learning to boost performance which makes it standout.

Furthermore, the project will address gaps in the literature by focusing on specific challenges or scenarios. These will include investigating the model's performance under varying levels of image quality, exploring the model's scalability for real-time diagnosis, or developing an interpretable framework to aid healthcare professionals in understanding the model's decision-making process.

# 3.0 Project Objectives

- **Develop a Deep Learning Model**: Create a deep learning model for malaria detection using cell images with competitive performance compared to existing approaches.
- **Explore Enhanced Architectures and Techniques**: Investigate advanced CNN architectures and data augmentation methods to potentially improve the model's accuracy and robustness.
- **Provide Basic Interpretability**: Implement basic visualization techniques to offer insights into the model's decision-making process, enhancing its transparency.
- **Comparison with Transfer learning**: Perform a thorough comparative analysis of the proposed model against existing state-of-the-art methods, identifying potential areas of improvement.
- **Documentation and Presentation:** Document the research process and outcomes comprehensively, and prepare a presentation summarizing key findings for dissemination within the research community.

# 4.0 Methodology

The methodology for malaria detection using cell images and deep learning involves a comprehensive approach that includes data preparation, model training, and performance evaluation. The primary deep learning method chosen for this project is a Convolutional Neural Network (CNN), and the performance will be compared with the established VGG16 architecture.

**4.1 Data Preprocessing**: The data source is from Kaggle(15) and it comprises of 27,560 raw cell images. Firstly, the image data was split

### 4.1.0 Training-Test-Validation Split:

The dataset is split into three subsets:

- **Training Set**: 70% of the dataset
- **Validation Set**: 20% of the dataset
- **Test Set**: 10% of the dataset

Preprocessing steps was carried out to ensure consistency and suitability for model training. This involves resizing images to a common resolution(160x160), normalizing pixel values by dividing each image by 255, and potentially applying data augmentation techniques like rotation, setting batch size=32, encoding labels (0,1) and flipping to enhance the diversity of the dataset.

**4.2 Convolutional Neural Network (CNN):** The primary model architecture for malaria detection is the

CNN. This architecture consists of 4 convolutional layers for feature extraction and spatial hierarchies, followed by pooling layers to reduce dimensionality, and finally fully connected layers for classification. The CNN is designed and trained from scratch using the training dataset.

**4.3 Training and Validation:** The CNN is trained using the training dataset, during training, the models learn to minimize a chosen loss function through gradient descent optimization. A validation dataset is used to monitor training progress and prevent overfitting.

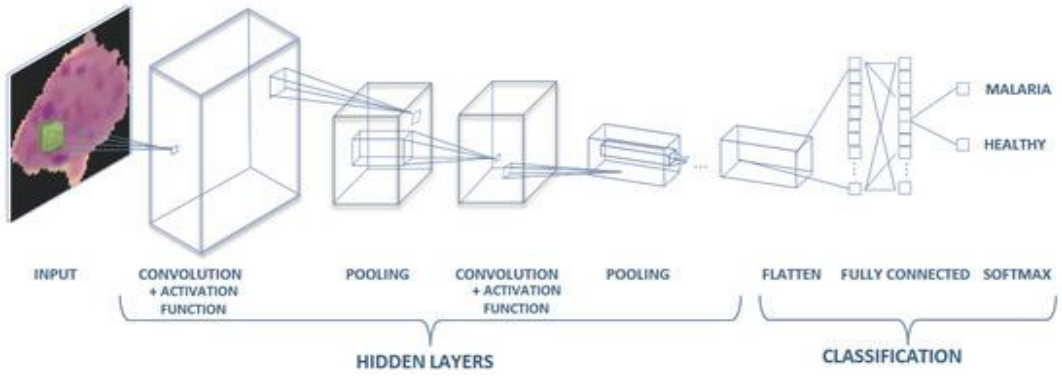
**4.4 Performance Metrics and Evaluation:** The trained CNN are evaluated on a separate test dataset that was not used during training. Metrics such as accuracy, precision, recall, and F1-score are calculated to gauge the models' abilities to correctly classify malaria-infected and healthy cells. The goal is to compare the performance of the CNN against the established VGG16 architecture.

**Introducing VGG 16 as our pre trained model.**

**4.5 VGG16 Fine-Tuning/Training and Evaluation**: The pre-trained VGG16 model, which was originally trained on ImageNet, is fine-tuned using the same training dataset. The final fully connected layers of VGG16 are replaced to match the number of classes in the malaria dataset and valuated with the unseen test dataset. Transfer learning adapts the model to the specific task of malaria detection and its been evaluated with a test dataset

**4.6 VGG16 Comparison with CNN Model**: In addition to the CNN, the VGG16 architecture will serve as a benchmark for performance comparison. VGG16 is a well-established CNN architecture known for its deep layer structure and remarkable performance on image classification tasks. The pre-trained VGG16 model, trained on a large dataset like ImageNet, will be fine-tuned for malaria detection using the same dataset.

**CNN Architecture Diagrams**




( **Figure 1**: **CNN Architecture Diagram: [CNN Architecture Diagram]**) (16)

The **convolution blocks** are known as convolutional operations to extract features in CNNs Max pooling is used to decrease the spatial dimensions of data while preserving important information. **The flatten** operation is employed to transform input into a one-dimensional array, whereas **Dense layers** establish

connections between all neurons in order to facilitate classification. **SoftMax** was used as the activation function (between 0 and 1)

## 5.0 Experiments

The experimental setup for comparing the performance of the custom Convolutional Neural Network (CNN) with the VGG16 architecture in malaria detection involves meticulous configuration of hyperparameters, and evaluation metrics.

- **First Custom CNN: MODEL 1**

Learning Rate: 0.001

Batch Size: 32

Number of Epochs: 20

Optimizer: Adam

- **Hyperparameters for Custom Model**: **MODEL 2**

Learning Rate: 0.001

Batch Size: 32

Dropout Rate: 0.5

Number of Epochs: 25

Optimizer: Adam

- **Hyperparameters for VGG 16**: **MODEL 3**

Learning Rate: 0.001

Batch Size: 32

Dropout Rate: 0.5

Number of Epochs: 25

Optimizer: Adam

### 5.1 Baselines:

- **Custom CNN**: The custom CNN architecture specifically designed for malaria detection.
- **VGG16**: The pre-trained VGG16 architecture fine-tuned for malaria detection.

**5.2 Evaluation Metrics**: The performance of three models will be evaluated using the following metrics:

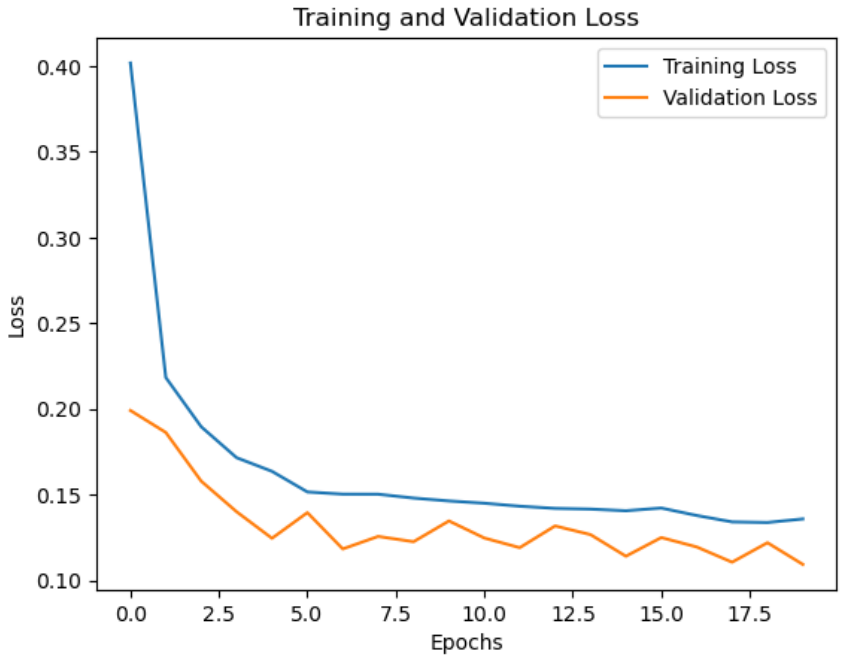
- **Accuracy**: The proportion of correctly classified instances among all instances.
- **Precision**: The proportion of true positive predictions among all positive predictions.
- **Recall**: The proportion of true positive predictions among all actual positive instances.
- **F1-Score**: The harmonic mean of precision and recall, providing a balanced measure of model performance

## 6.0 Result/Model Performance:

### 6.0.1 First custom CNN model: (MODEL 1)

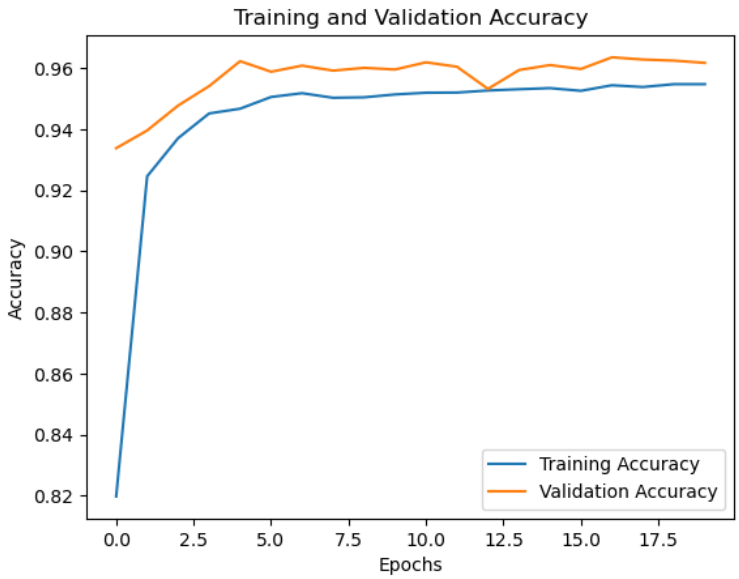
- Learning Rate: 0.001, - Convolution blocks = 32,64, 64, 64, - Batch Size: 32, - Number of Epochs: 20

- Optimizer: Adam ()




**(Figure 2: Training and Validation Loss)**

From Figure 2 above, we can see that the training loss did not intersect with the validation accuracy and training loss was always higher than the validation loss across all the 20 epochs. This means that my model learnt the patterns well enough.




**(Figure 3: Training and Validation Accuracy)**

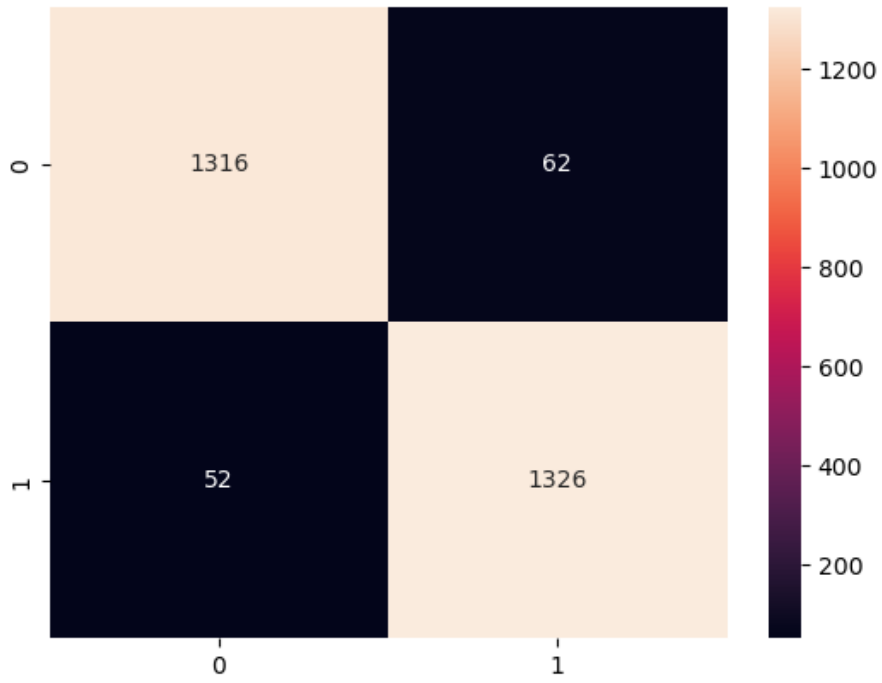
From Figure 3 above, we can clearly notice that at some point in our epoch the validation accuracy intersects with the training accuracy and at that point the training accuracy and validation accuracy was same value which is not meant to be this suggest sign of overfitting.

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Parasitized | 0.96 | 0.96 | 0.96 | 1378 |
| Uninfected | 0.96 | 0.96 | 0.96 | 1378 |
| accuracy | | | 0.96 | 2756 |
| macro avg | 0.96 | 0.96 | 0.96 | 2756 |
| weighted avg | 0.96 | 0.96 | 0.96 | 2756 |

**(Figure 4: Classification report)**

Clearly from **Figure 4** above, precision, recall, and F1-score were 96% across the two classes and our accuracy was 96% as well, generally our model performed very well but we noticed a sign of overfitting in **Figure 3** between the validation accuracy and training accuracy**.** This means though the model has done very well during training, it might find it a little hard to generalize what it has learnt on a new data,

which means we cant trust it to always make the right prediction.



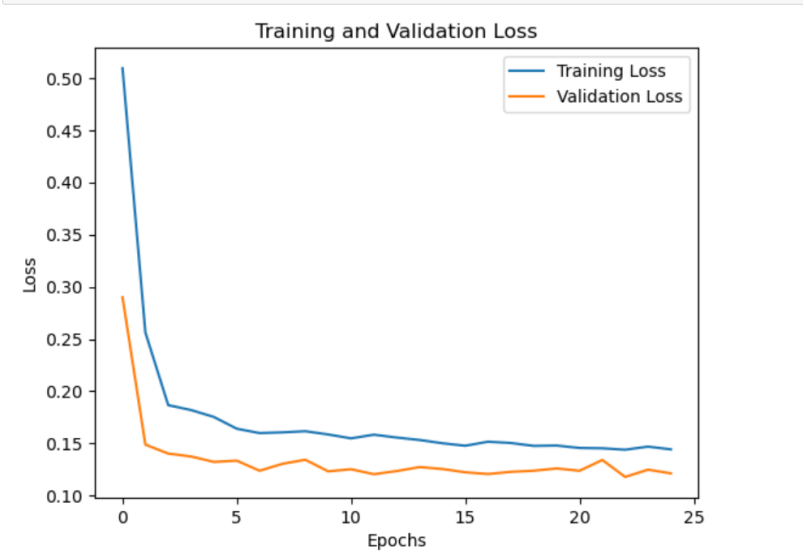

**(Figure 5: Confusion Matrix)**

**From Figure 5,** we can clearly see that it classified 1316 cell images as Parasitized (Infected) and the 1326 as uninfected

### 6.0.2 Hyperparameters for Custom CNN:(MODEL 2)

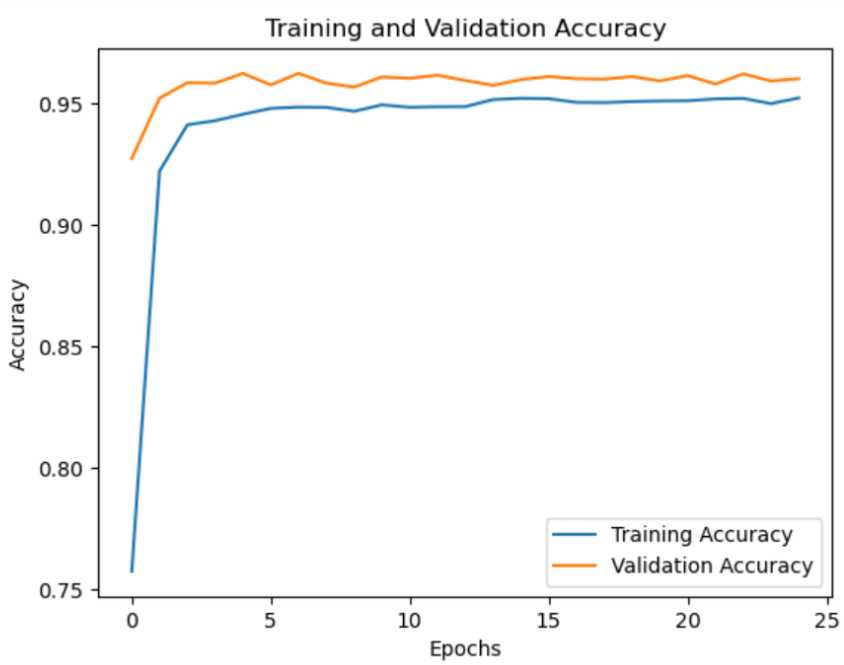
- Learning Rate: 0.001, - Convolution blocks = 32,64, 64, 128, - Batch Size: 32, - Dropout Rate: 0.5

- Number of Epochs: 25, - Optimizer: Adam ()




**(Figure 6: Training and Validation Loss)**

As seen from **Figure 6**, our training loss does not intersect with our validation loss which suggests that the model is not only fitting the training data but also capturing the underlying patterns well enough to make accurate predictions on new data.



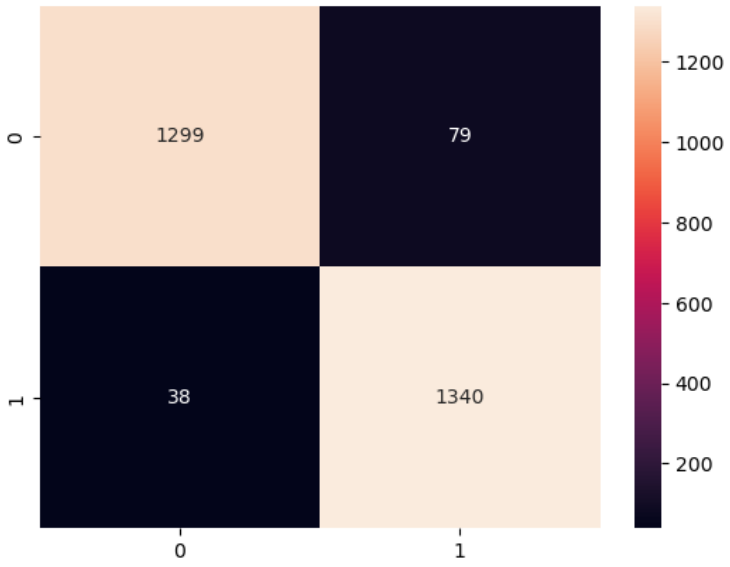

**(Figure 7: Training accuracy vs Validation accuracy)**

As seen from **Figure 7**, the training accuracy was lower than the validation accuracy at every epoch and they did not intersect at any point during the training, this suggest that the model learnt to generalize and perform consistently on both the training and validation datasets. The fact that the lines converged but did not intersect suggests that the model is improving its performance on both datasets as training progresses

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Parasitized | 0.97 | 0.94 | 0.96 | 1378 |
| Uninfected | 0.94 | 0.97 | 0.96 | 1378 |
| accuracy | | | 0.96 | 2756 |
| macro avg | 0.96 | 0.96 | 0.96 | 2756 |
| weighted avg | 0.96 | 0.96 | 0.96 | 2756 |

**(Figure 8: Classification report)**

**From Figure 8 above, our model performed very well across all metrics, it has 96% for F1 score of both Parasitized and Uninfected, for recall its 97% for Uninfected and 94% for Parasitized, also on the precision it has 97% on parasitized and 94% for Uninfected. Considering the accuracy, the model recorded 96% accuracy which is perfect.**



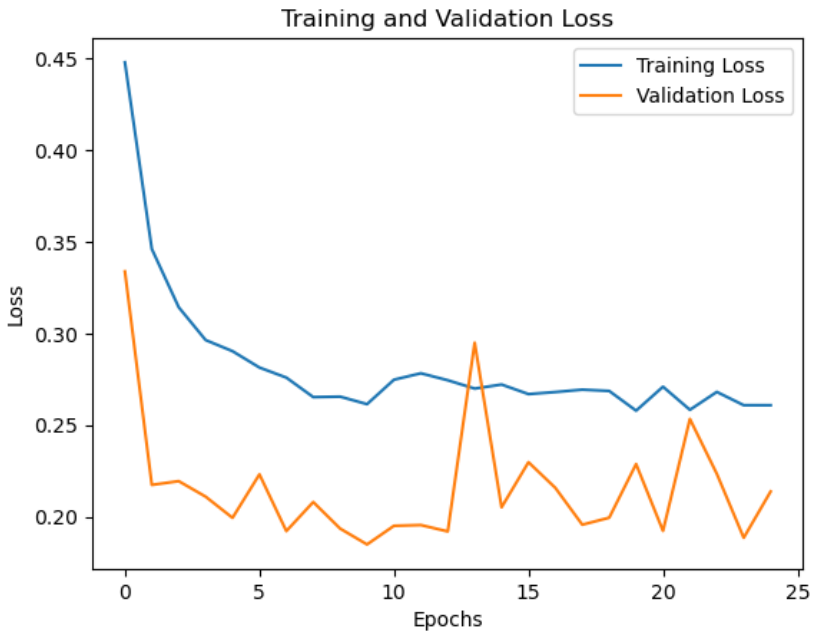

**(Figure 9: Confusion matrix)**

**Figure 9** above suggest that the model predicted a total of 1299 as parasitized, and 1340 as unaffected.

The custom hyperparameter Model (MODEL 2) performed better than the first CNN model (MODEL 2), even though they had the same accuracy. MODEL 1 showed signs of overfitting during its training and validation accuracy of the Model while MODEL 2 showed that the model did not only fit the training data but also captured the underlying patterns well enough to make accurate predictions on new data

### 6.0.3 Hyperparameters for VGG16 : (MODEL 3)

- Learning Rate: 0.001, - Batch Size: 32, - Dropout Rate: 0.5, - Number of Epochs: 25, - Optimizer: Adam ()




(**Figure 10: Training loss vs Validation loss for vgg16**)

Clearly from Figure 10, we can notice that the training loss is intersecting with the validation loss this suggests that the model is fitting the training data more closely than it should and not generalizing effectively with new or unseen data.

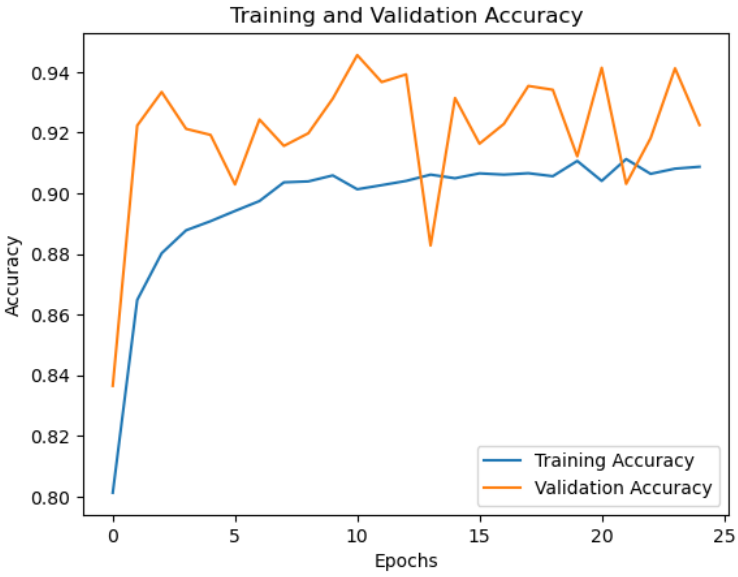



**Figure 11: Training and Validation Accuracy for Vgg 16**

As seen in Figure 11, we can also see the curve for Training accuracy intersecting with Validation accuracy which again suggest overfitting of the vgg16 model. This simply means at some point in our validation phase, the training accuracy was higher than the validation accuracy which is not meant to be.

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Parasitized | 0.98 | 0.87 | 0.92 | 1378 |
| Uninfected | 0.88 | 0.98 | 0.93 | 1378 |
| accuracy | | | 0.92 | 2756 |
| macro avg | 0.93 | 0.92 | 0.92 | 2756 |
| weighted avg | 0.93 | 0.92 | 0.92 | 2756 |

**(Figure 12: Classification report for vgg16)**

From our classification report in figure 12, we can clearly notice that our accuracy is 92%, which is not bad, precision did 98% for parasitized and 88% for Uninfected, recall did 87% for parasitized and 98% for Uninfected and F1-score did 92% for parasitized and 93% for Uninfected. This suggest a good model but again on figure 10 and Figure 11, we can clearly notice that our model overfitted, which now nullifies the efficiency of this model.

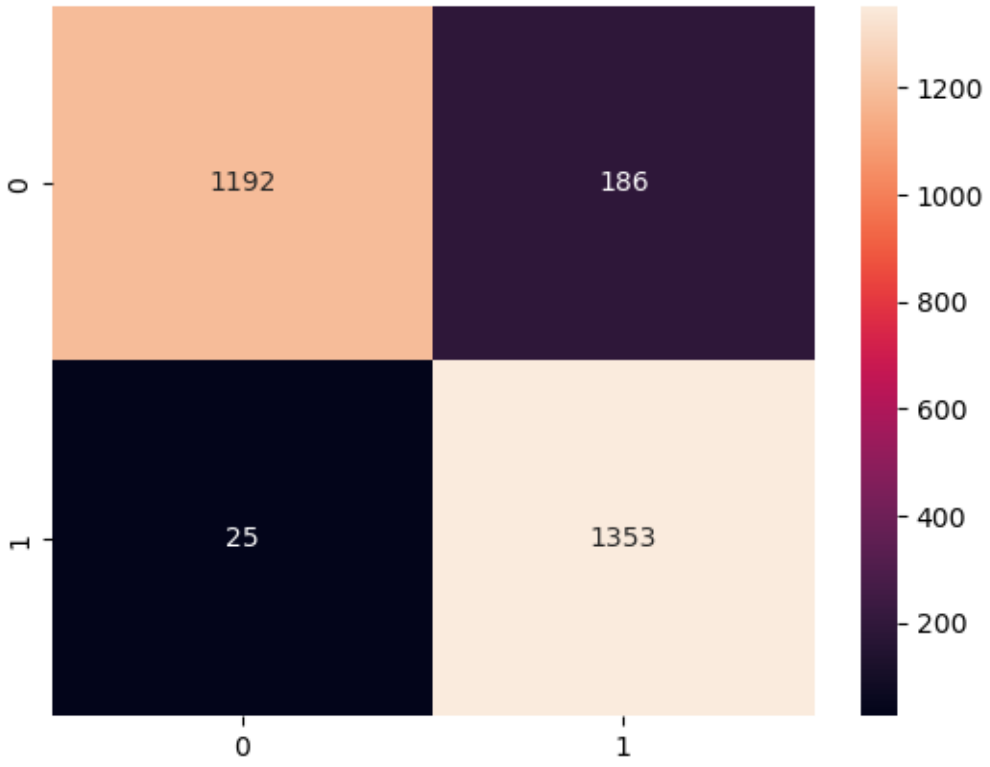


**Figure 13: Confusion Matrix**

From the figure 13 above, we can see that the model predicted 1353 cell images as Uninfected and predicted 1192 cell images as parasitized.

**Classification Report for our 3 models and Performance Metrics:**

| ModelS | Accuracy | precision | | Recall | | F1-Score | |
|---|---|---|---|---|---|---|---|
| | | Parasitized | Uninfected | Parasitized | Uninfected | Parasitized | Uninfected |
| Custom CNN (MODEL 1) | 0.96 | 0.96 | 0.96 | 0.96 | 0.96 | 0.96 | 0.96 |
| Custom CNN (hyperparameter tunning) MODEL 2 | 0.96 | 0.97 | 0.94 | 0.94 | 0.97 | 0.96 | 0.96 |
| Vgg16 Model (MODEL 3) | 0.92 | 0.98 | 0.88 | 0.87 | 0.98 | 0.92 | 0.93 |

**Figure 14**

### 6.1 Result Comparison:

When comparing the three models, it is observed that **MODEL 1** (custom model without tuning) achieved an accuracy of 96%, **MODEL 2** (custom model with tuning) also achieved an accuracy of **96%**, and **MODEL 3** (VGG16) achieved an accuracy of **92%**, as depicted in Figure 14. It can be concluded that both **MODEL 2** and **MODEL 1** outperformed the **MODEL 3** (VGG16 model) in terms of F1 score and Accuracy, exhibiting a 4% higher accuracy**.**

As noted on **MODEL 1,** there was presence of overfitting, where the model could not learn properly to generalize and perform consistently on the validation data set as seen in **Figure 3,** this might affect results or cause bias when it meets a new or unseen data**. Also on the MODEL 3,** even though it had accuracy of **92%** had more cases of overfitting in both the training vs validation loss and training vs validation accuracy as seen In **Figure 10** and **Figure 11** respectively**.**

**MODEL 2** did very well with an accuracy of **96%** and there was no sign of overfitting recorded as seen in **Figure 6 and Figure 7.** This means that the **MODEL 2** captured the underlying patterns well enough and learnt to generalize and perform consistently on both the training and validation datasets to make accurate predictions on new data.

We can therefore conclude that **MODEL 2**, did best in correctly prediction whether a cell image was infected with Malaria or not Infected with Malaria.

### 6.2 Objective Achievement:

The objectives of the project were fully met. We developed a deep Learning Mode to predict the presence of malaria in a cell image, explored enhanced architecture and technique by employing hyperparameter, I went ahead to explain the visualizations of our results, furthermore, compared my custom CNN models to Transfer learning (VGG16), though the custom model outperformed the transfer learning, went ahead to document my result.

### 6.3 Interesting Insights:

- **Dropout Rate Impact**: A dropout rate of 0.5 was used for two models. This high dropout helped to prevent overfitting on the custom model, but didn't prevent that on the vgg16 model.
- **Dataset Variability**: The diversity of cell images in the dataset played a significant role in the models' performance. The custom CNN, tailored to the task, was able to leverage this diversity more effectively than the transfer-learned VGG16.

### 6.4 Further Investigation:

- **Robustness of Unbalanced Data**: Examine custom CNN's adaptability to imbalanced datasets, which is critical in real-world situations. Create strategies to improve its performance under skewed class distributions.

- **Diverse Architectures**: Explore alternate CNN designs. Investigate established and novel architectures to identify superior models for malaria detection.

## 7.0 Conclusion

In conclusion, this study explored the application of deep learning for malaria detection using cell images, specifically comparing two custom Convolutional Neural Network (CNN) with the established VGG16 architecture. The project aimed to contribute insights into the effectiveness of custom architectures for domain-specific tasks and to provide a thorough evaluation of their performance.

The custom CNN outperformed the VGG16 architecture across all performance metrics, including

accuracy, precision, recall, and F1-score. While the custom CNN did not achieve a wide margin of superiority, the results demonstrated the potential of tailored architectures for enhancing the accuracy of malaria detection.

## 8.0 References


1. Bohr A, Memarzadeh K. The rise of artificial intelligence in healthcare applications. Artif Intell Healthc. 2020;25–60.

2. Sarpong E, Acheampong DO, Fordjour GNR, Anyanful A, Aninagyei E, Tuoyire DA, et al. Zero malaria: a mirage or reality for populations of sub-Saharan Africa in health transition. Malar J. 2022 Nov 4;21(1):314.

3. Cardoso MJ, Arbel T, Carneiro G, Syeda-Mahmood T, Tavares JMRS, Moradi M, et al. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: Third International Workshop, DLMIA 2017, and 7th International Workshop, ML-CDS 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, September 14, Proceedings. Springer; 2017. 399 p.

4. Jdey I, Hcini G, Ltifi H. Deep learning and machine learning for Malaria detection: overview, challenges and future directions [Internet]. arXiv; 2022 [cited 2023 Aug 21]. Available from: http://arxiv.org/abs/2209.13292

5. Fact sheet about malaria [Internet]. [cited 2023 Aug 21]. Available from: https://www.who.int/news-room/fact-sheets/detail/malaria

6. Zekar L, Sharman T. Plasmodium falciparum Malaria. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 [cited 2023 Aug 21]. Available from: http://www.ncbi.nlm.nih.gov/books/NBK555962/

7. Malaria [Internet]. [cited 2023 Aug 21]. Available from: https://www.who.int/news-room/questions-and-answers/item/malaria

8. Fitri LE, Widaningrum T, Endharti AT, Prabowo MH, Winaris N, Nugraha RYB. Malaria diagnostic update: From conventional to advanced method. J Clin Lab Anal. 2022 Mar 4;36(4):e24314.

9. Deelder W, Benavente ED, Phelan J, Manko E, Campino S, Palla L, et al. Using deep learning to identify recent positive selection in malaria parasite sequence data. Malar J. 2021 Jun 14;20(1):270.

10. Madhu G, Mohamed AW, Kautish S, Shah MA, Ali I. Intelligent diagnostic model for malaria parasite detection and classification using imperative inception-based capsule neural networks. Sci Rep. 2023 Aug 17;13(1):13377.

11. Molina A, Rodellar J, Boldú L, Acevedo A, Alférez S, Merino A. Automatic identification of malaria and other red blood cell inclusions using convolutional neural networks. Comput Biol Med. 2021 Sep 1;136:104680.

12. Rishabh Nimje [Internet]. 2020 [cited 2023 Aug 22]. Malaria Cell Images Database. Available from: https://risx3.github.io/malaria-cell-database/

13. Siłka W, Wieczorek M, Siłka J, Woźniak M. Malaria Detection Using Advanced Deep Learning Architecture. Sensors. 2023 Jan;23(3):1501.

14. Misra I. Using a Convolutional Neural Network(CNN) to Diagnose Malaria [Internet]. Analytics Vidhya. 2021 [cited 2023 Aug 21]. Available from: https://medium.com/analytics-vidhya/using-a-convolutional-neural-network-cnn-to-diagnose-malaria-589489e739ca

15. Malaria Cell Images Dataset [Internet]. [cited 2023 Aug 21]. Available from: https://www.kaggle.com/datasets/iarunava/cell-images-for-detecting-malaria

16. Sriporn K, Tsai CF, Tsai CE, Wang P. Analyzing Malaria Disease Using Effective Deep Learning Approach. Diagnostics. 2020 Oct;10(10):744.

17. ResearchGate [Internet]. [cited 2023 Aug 20]. Fig. 8. VGG-16 CNN model architecture layer wise. Available from: https://www.researchgate.net/figure/VGG-16-CNN-model-architecture-layer-wise_fig8_346259768