Journal of Innovative Image Processing (ISSN: 2582-4252)
www.irojournals.com/iroiip/

# XMal-CNN: An Explainable Deep Neural Model for Automated Malaria Detection from Blood Smear Images

**RiddhiKumari Patel[1], Safvan Vahora[2]**

[1]Computer Engineering, [2]Information Technology Engineering, Government Engineering College Modasa, Gujarat, India.

**E-mail:** [1]priddhi002@gmail.com, [2]safvan465@gmail.com

## Abstract

Malaria is a severe and critical health issue widespread throughout the globe. Malaria must be diagnosed correctly and efficiently in its initial stage in order to treat and cure it before it becomes a terminal illness. The current paper describes XMal-CNN, a novel deep learning approach to be utilized in automated malaria diagnosis from microscopic blood smear images. The proposed structure utilizes a depth-wise Convolutional Neural Network (CNN) with a Squeeze and Excitation (SE) block to increase feature representation and perform classification of images. The suggested approach model performs in such a way that it surpasses baseline CNNs and currently existing state-of-the-art approaches, achieving 95.26% accuracy, 93.97% precision, 96.73% recall, and 95.33% F1-score. To improve model interpretability and explainability, Explainable AI (XAI) techniques such as LIME and Grad-CAM++ are used, providing useful insights and understanding of the decision making process of the model. Systematic and extensive evaluations on benchmark blood smear image datasets are conducted to validate the performance and explainability of the proposed model. Due to its superior diagnostic precision and interpretability, XMal-CNN becomes a trustworthy and important AI-assisted tool, aiding healthcare experts in making informed and data-driven decisions to diagnose and treat malaria.

**Keywords:** Blood Smear Images, Deep Learning, Depthwise Convolutional Neural Network, Explainable AI (XAI), Medical Image Classification, Malaria Detection.

## 1. Introduction

Malaria is a potentially fatal disease caused by Plasmodium parasites, which are transmitted to humans through the bites of infected female Anopheles mosquitoes. In countries such as India, the combination of suitable climate, environmental conditions and socio-economic challenges allows the disease to exist as a major public health problem [1].

Effective treatment of malaria depends heavily on timely and accurate diagnosis. This is because the different Plasmodium species most notably P. falciparum and P. vivax respond to different anti-malarial drugs. Administering the wrong treatment can lead to complications and contribute to the growing problem of drug resistance [2].

Although malaria can be cured and preventive medications are available, an effective vaccine has yet to be developed. Symptoms of fever, chills, headache, and vomiting usually


*DOI: https://doi.org/10.36548/jiip.2025.3.010*
*Received: 15.07.2025, received in revised form: 16.08.2025, accepted: 01.09.2025, published: 10.09.2025*

develop one to two weeks after infection. Since these symptoms are very similar to those seen in other diseases, accurate diagnostic testing is very important. Microscopic inspection of stained blood smears is still the gold standard, but it requires qualified laboratory staff able to accurately recognize the parasite species of malaria and their life cycle stages [2].

Global trends in malaria incidence and mortality based on recent data indicates that between 2020 and 2023, cases of malaria around the world increased to 263 million from 247 million. During the same period, deaths related to malaria fell from 622,000 to 597,000[1]as presented in figure 1. This trend is indicative of the success in prevention and treatment that results in saving lives, while transmission is still rising. This improvement in trends of malaria is due to improved and increased reach of diagnostics, better treatments as compared to the past and wider distribution of malaria preventive measures such as insecticide treated bed nets and modern anti-malarial drugs and medications.

Several different approaches are used to detect malaria and its stages. Out of these different methods, Microscopic Blood Smear Examination method, is the most widely used method in medical settings throughout the world. In this method, thick smears of a patient's blood are used to identify the presence of malaria parasites, whereas thin smears are used to identify the species of Plasmodium present in the blood [3]. The trained lab-technician or microbiologist then looks at the stained slide under microscope and counts infected red blood cells to measure malaria parasite load. This method is highly cost effective and provides detailed information about malaria species and parasite stages. The drawback of this method is that the accuracy of diagnostics depends heavily on the examiner's skill and expertise.

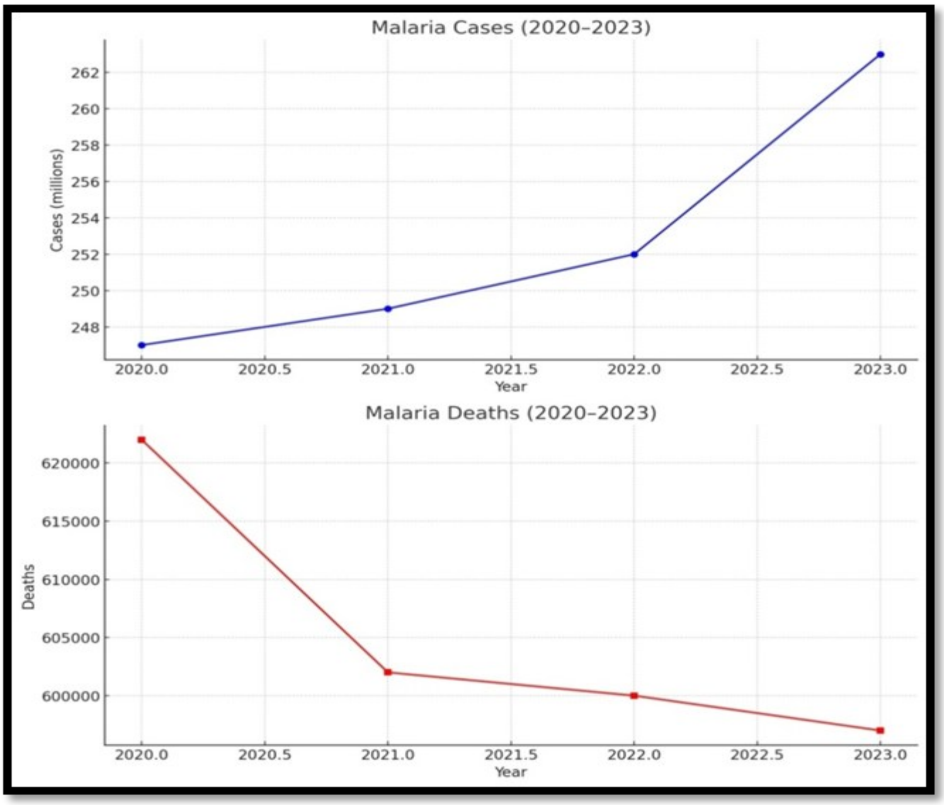


**Figure 1.** Yearly Trends of Malaria Cases and Deaths in Millions (2020–2023)

## A. Microscopic Blood Smear Examination

This is a traditional and conventional method of malaria detection that involves making thick and thin smears of a patient's blood on a slide manually using laboratory chemicals. The thick smear identifies the presence of malaria parasites in the blood sample and the thin smear helps in identifying the specific species of Plasmodium in the blood sample. A skilled and expert lab-technician or microbiologist observes the blood smear slides under a microscope and counts infected RBCs, as pictorially represented in Figure 2[3].

Microscopic diagnosis of malaria is not costly and also enables accurate detection of malaria causing Plasmodium species and life stages of the parasite. As a result, this method is more accurate for the quantification of infected RBCs than RDTs. Its performance, however, is very much dependent on the skill and expertise of the person handling and analyzing the sample of the patient’s blood.

Although microscopy offers the most precise identification, it is also labour intensive and time consuming. Inexperienced examiners may produce less reliable results, and processing large numbers of samples can strain resources in high burden areas [5].

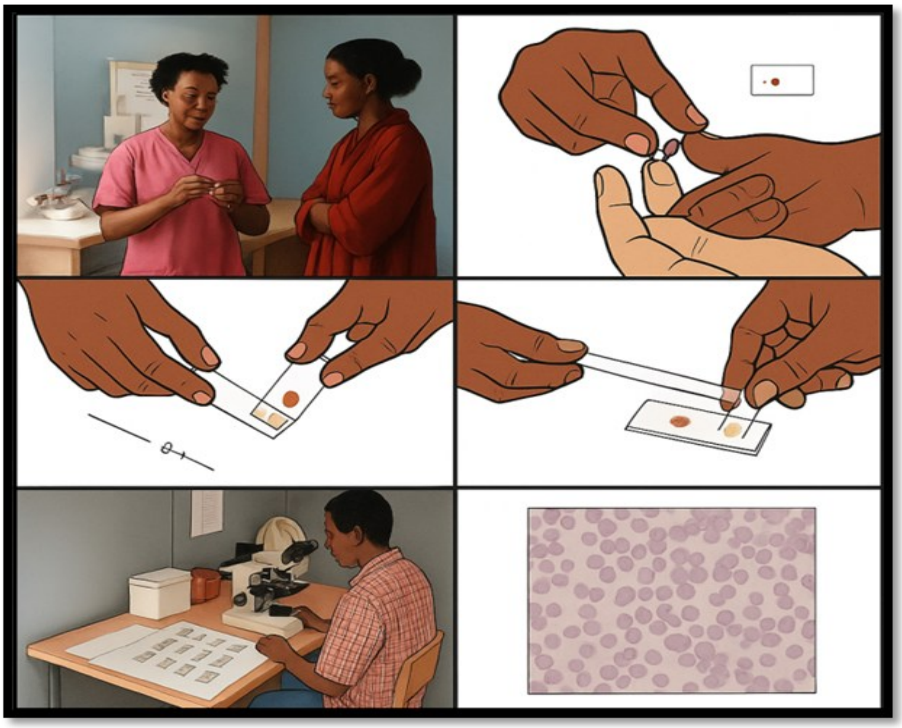

**Figure 2.** Blood Smear Microscopy Process for Malaria Diagnosis [3]

## B. Rapid Diagnosis Test (RDT)

A Rapid Diagnostic Test (RDT) kit is a compact device that detects antigens from malaria parasites present in blood. A blood sample is inserted into the kit, which processes it internally to provide quick results. The RDT kit for malaria detection is shown in Figure 3 [3].

RDT kits offer key advantages like delivering rapid results, requiring no specialized expertise, and being especially useful in malaria-endemic regions. However, they have limitations. Studies show that RDTs can lack accuracy, potentially leading to incorrect treatment decisions [5]. A major drawback is their inability to identify specific Plasmodium species. As a result, researchers recommend developing computer-assisted malaria detection systems to improve accuracy, speed, and reduce reliance on expert medical personnel.

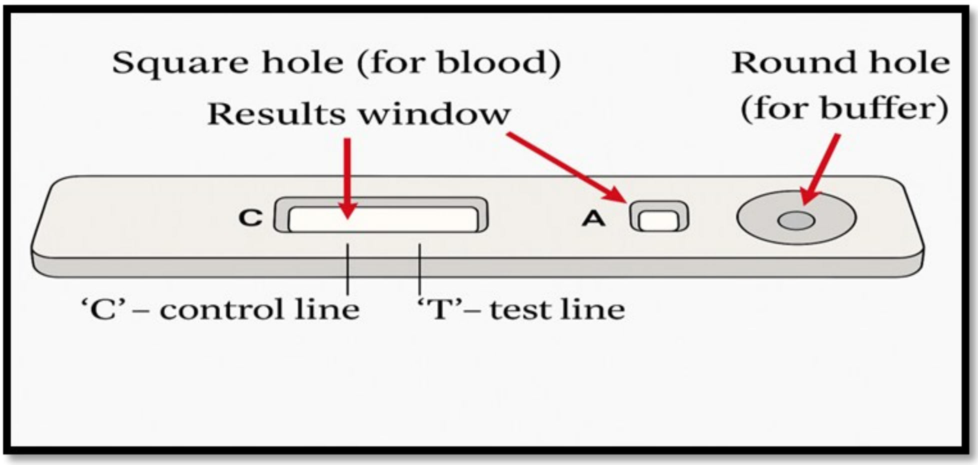


**Figure 3.** Rapid Diagnostic Testing Kit (RDT) for Malaria Detection [3]

After evaluating the strengths and limitations of existing malaria diagnostic methods, researchers suggest that computer-assisted detection is necessary. Such systems can overcome issues related to accuracy, speed, and dependence on skilled microbiologists, thereby enhancing the overall performance of malaria diagnosis.

### C. Polymerase Chain Reaction (PCR)

The PCR (Polymerase Reaction) method of diagnosis is the most sensitive malaria test, detecting as low as 1 parasite/µL compared to 50 parasites/µL by microscopy. However, it is also the most labour-intensive, expensive, and time-consuming method. While commonly used in developed countries, its availability is limited in developing regions.

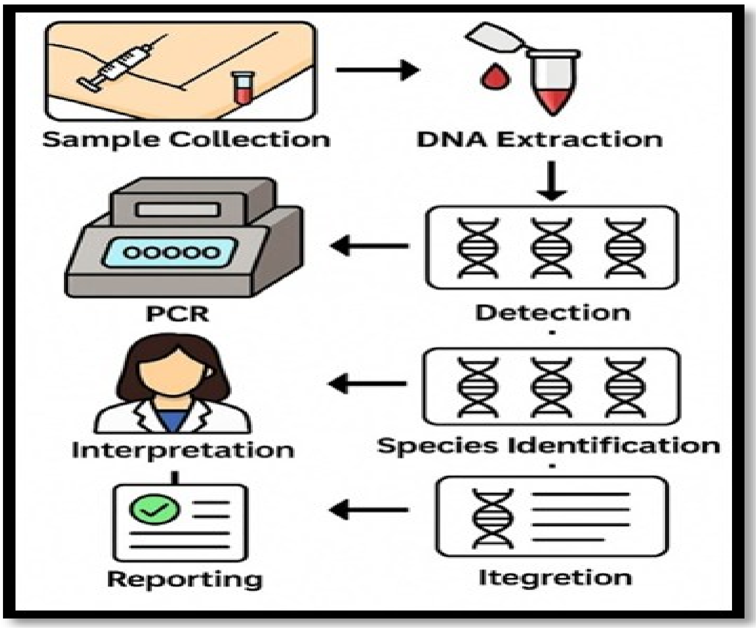


**Figure 4.** PCR Technique for Detecting Malaria Parasites

PCR detects malaria by amplifying the parasite's DNA using a polymerase enzyme, allowing detection from very small or dried blood samples as presented in figure 4. While highly sensitive, it is costly, time-consuming, and less accessible in low-resource settings. Microscopic examination of Giemsa-stained blood smears remains the traditional benchmark standard to date, due to its high sensitivity and specificity, but it is labour-intensive and prone to human error, especially where skilled personnel are scarce [4] [28]. RDTs offer a quicker alternative but are limited by species and parasite density sensitivity [3].

To overcome these limitations, automated diagnostic platforms using Artificial Intelligence, Machine Learning, and Deep Learning have emerged as promising solutions. These technologies enable fast, accurate analysis of blood smear images and medical data, reducing human error and supporting diagnosis in underserved areas.

AI-powered systems are increasingly used in healthcare to diagnose diseases like malaria, tuberculosis, and COVID-19, as well as neurological diseases such as Alzheimer's disease, Parkinson's disease, epilepsy, and stroke by analyzing various medical images and signals [5].

AI technologies are transforming healthcare by improving diagnosis, risk assessment, and treatment planning across various medical domains. In cardiovascular care, AI-based electrocardiogram (ECG) analysis aids in detecting arrhythmias and coronary artery disease, enhancing patient management. Deep learning models have significantly improved diagnostic accuracy in tumor and cancer detection—such as brain tumors, breast cancer, lung cancer, cancers—through the analysis of pathological and radiological images [26]. AI plays a significant role in the diagnosis and management of many health conditions through advanced image and data analysis. In the treatment and care of Diabetes, AI based technologies help in predicting and managing the disease by analyzing medical histories and retinal images of patients, enabling to create patient specific tailored treatment strategies. In AI based Diabetes care and treatment systems, Convolutional Neural Networks (CNN) are commonly used, for the detection of diabetic retinopathy by recognizing indicators like microaneurysms and hemorrhages in the retina through image analysis [6].

Ophthalmic conditions such as diabetic retinopathy, glaucoma, and macular degeneration can be effectively diagnosed using deep learning models trained on Fundus images and Optical Coherence Tomography (OCT) scans obtained from medical devices. [5]. The use of transfer learning, especially using pre-trained analysis models like ResNet and VGG, improves the accuracy and efficiency of disease detection. Convolutional Neural Networks (CNNs) are of great importance and value in this domain, as these models automatically extract and learns intricate features from raw medical images obtained from medical devices. Automation removes the need for manual feature selection in images and improves the overall diagnostic performance of detection systems.

For respiratory diseases such as pneumonia, asthma and COPD; deep learning models analyze lung images and sensor data from medical imaging machines to support the early and accurate diagnosis of disease [7]. In dermatology, AI systems classify skin conditions, including melanoma and psoriasis, by processing and analyzing high resolution skin images. Remarkable precision is obtained. It can be appropriately deduced that Explainable AI (XAI) is being increasingly adopted in medical diagnostics across the globe to build trust and transparency in AI systems. In malaria detection, AI models trained on blood smear images can identify infected cells with high accuracy. However, understanding the model's decisions is critical. Techniques like SHAP (SHapley Additive exPlanations) assign importance to image regions that influence predictions, providing visual explanations of processing, such as heat maps [8]. This improves trust and helps medical professionals align AI outputs with clinical judgments. For example, a study shows that a lightweight CNN model using SHAP outperformed existing models in malaria diagnosis and proved effective for rapid and interpretable diagnostics [7].

In this study, we present a customized approach of using depth-wise CNN integrated with a Squeeze-and-Excitation (SE) block, achieving superior performance over traditionally used CNNs, SE-CNNs and depth-wise CNNs without SE blocks. To enhance the transparency of the model, we have incorporated XAI techniques such as LIME and Grad-CAM++. These techniques offer detailed insights into the model's decision-making process. By combining high classification accuracy with explainability, our proposed approach strengthens existing malaria detection, offering both reliability and interpretability for improved diagnostic outcomes in situations with limited resources.

## 2. Related Work

Recent advances in data driven technologies such as machine learning and artificial intelligence have assisted in the enhancement of malaria diagnosis through the automation of blood smear image analysis. Deep learning techniques, particularly Convolutional Neural Networks (CNNs), transfer learning, and ensemble models have accomplished high accuracy in the detection of infected red blood cells. Interpretability tools such as LIME and Grad-CAM++ enhance clinical trust by visualizing model decisions. These systems are currently being mapped onto mobile and cloud platforms, making them more accessible in low-resource environments [9].

Traditional diagnostic methods like microscopy, RDTs-(Rapid Diagnostic Tests), and PCR each have significant drawbacks. Microscopy is precise but requires skilled personnel and is time-consuming. RDTs provide faster results but tend to be less reliable based on their sensitivity to parasitaemia levels and variability in the quality of tests. PCR is highly precise and able to detect low-level infections as well as mixed infections but is too costly and technically demanding for use in areas with limited resources on a large scale. Previous methods of computation employing machine learning sought to be more efficient by hand extracting features such as texture, colour, and shape to determine infected cells. Even though they were more scalable than traditional manual microscopy, these methods were not as robust, flexible, or precise as possible now with the aid of current deep learning methods [9].

Today's deep learning models present a more potent solution for detecting malaria. One research study conducted by Sidharthan [9] presented an AI-driven diagnostic system utilizing the EfficientNet-B2 architecture, which was 93.57% accurate in detecting malaria from images of blood smears. This model performed better compared to other previous pre-trained networks based on precision, F1-score and AUC as the metrics of evaluation. It was also found in the study that employing an 80:20 divisions between training and testing data produced better model performance compared to a 90:10 divisions, underlining the significance of having a well-balanced dataset in the case of model training.

A number of exclusive innovative research studies have been performed by researchers globally to further enhance malaria diagnosis using artificial intelligence based on a wide range of various machine learning (ML) and deep learning (DL) methods and models. An important innovation in this respect is by Liu et al., who presented AIDMAN, an exclusive mobile device compatible platform based on YOLOv5 and Transformer-based models for thin blood smear image analysis. This system reported remarkable accuracies of 93.62% [10] and 98.62% in independent trials. Despite the potential for field diagnostics, its actual time application and deployment are hindered by the need for high end GPUs and considerably high computational capabilities, limiting its availability and utilization in low resource environments.

Conventional machine learning techniques have also been investigated by researchers employing image pre-processing. Kunwar et al. [11] proposed a diagnostic model paradigm by integrating image processing techniques like edge detection and morphological filtering with machine learning classifiers, obtaining a total accuracy of approximately 90%. Similar work has been done by Motwani et al. [12], using colour, texture and shape based feature extraction and subsequent application of classifiers such as ANN and k-NN, obtaining similar performance compared to earlier work. In spite of their simplicity, these models struggle with the ability to cope with varied clinical image datasets due to variations in image quality and poor generalization capability.

Some researchers have used a different study approach by including clinical and epidemiological information rather than or along with image information. Srivastava and Saurabh [13] used techniques such as SVM, decision trees and random forest and attaining an accuracy of 98.46% with the random forest-based model. Lee et al. [14] constructed predictive models with patient records such as fever history, anemia and travel patterns and attained an accuracy of 96.2% with logistic regression and gradient boosting techniques. These methods emphasize the point, how structured data can be used to augment and enhance image based diagnostics, although their efficacy relies significantly on the availability and caliber of clinical patient records.

Deep learning, especially CNN and transfer learning, has proven to be extremely effective in the detection of malaria. Shekar et al. [15] employed transfer learning with CNNs and reported 94% accuracy, whereas Abubakar et al. [16] compared classic k-NN with CNNs, with CNNs emphatically outperforming at 96%. Nayak et al. [17] achieved further improvement in CNNs by combining ResNet architectures, ensemble methods, and data augmentation, reporting an accuracy of 94.5%. These deep learning networks provide better diagnostic accuracy, but their computational requirements typically preclude real-time usage, particularly in rural areas.

Even more advanced architectures such as Capsule Networks (Caps Net) have been investigated. Madhu et al. [18] proposed a CNN + Caps Net model with 96.1% accuracy, leveraging the capacity of Caps Net to retain spatial relationships in images more effectively. Hoyos et al. [19], on the other hand, enhanced CNN generalization using data augmentation (rotation, flipping, contrast variation) to achieve 94.57% accuracy. Both research studies emphasize the relevance of architectural advancements and image diversity in constructing solid diagnostic systems.

Comparative studies of deep learning models have shown additional findings. Hemachandran et al. [20] established that CNN performed better than Mobile Net and ResNet50, with 94.8% accuracy. Alnussairi and Abdulaziz obtained 93% using a CNN model but expressed issues with the explainability and generalization of the model [21]. Abubakar et al. [16] proposed Deep FMD, a CNN-based system that achieved superior performance on benchmark datasets (92.5% accuracy), but had lower accuracy when applied to images from varying staining methods. These results highlight the requirement for interpretable, flexible models trained across various datasets.

Several research groups have focused on developing CNN architectures that are optimized for real-time inference and scalable deployment across various cloud environments. Kumar and Kumar [22] implemented a CNN model trained specifically for thin blood smear images with 95.2% accuracy but at a high computational cost. Taye et al. [23] integrated Xception and InceptionV3 architectures within an ensemble model also with 95.2% accuracy but again with high memory and training expenses. Sukumarran et al. [24] used an optimized YOLOv4 model for malaria parasite object detection with 92% accuracy. These studies provide performance and speed at the cost of scalability and low-latency deployment in the real world.

These researches combined represent the great leaps AI has made towards malaria diagnosis, from simple machine learning techniques to complex deep learning and object detection architectures. Though most systems have demonstrated excellent diagnostic accuracy, in many cases well over 90%, their practical deployment in real-world settings continues to be hampered by challenges such as computational costs, interpretability of the models, and lack of robustness under varied clinical environments. The need for expensive hardware, inconsistency of imaging conditions, and necessity of high-quality annotated data further complicates field deployment. Nonetheless, the major challenges determined from the analyzed literature are dataset imbalance, inter-class similarity, and intra-class variance, which have drastic effects on malaria classification performance. To close this gap, future work should endeavor to develop models that are lightweight, interpretable, and generic, that retains high performance while being practicable for applicability in low-resource settings. Such advancements will be crucial to facilitating scalable, artificial intelligence-based diagnostic technologies that can efficiently aid worldwide efforts in malaria elimination and control.

## 3. Proposed Methodology

This paper introduces a novel deep learning-based system for automated malaria diagnosis using microscopic blood smear images, which has the potential to overcome time and skill-related limitations of routine microscopic diagnosis.

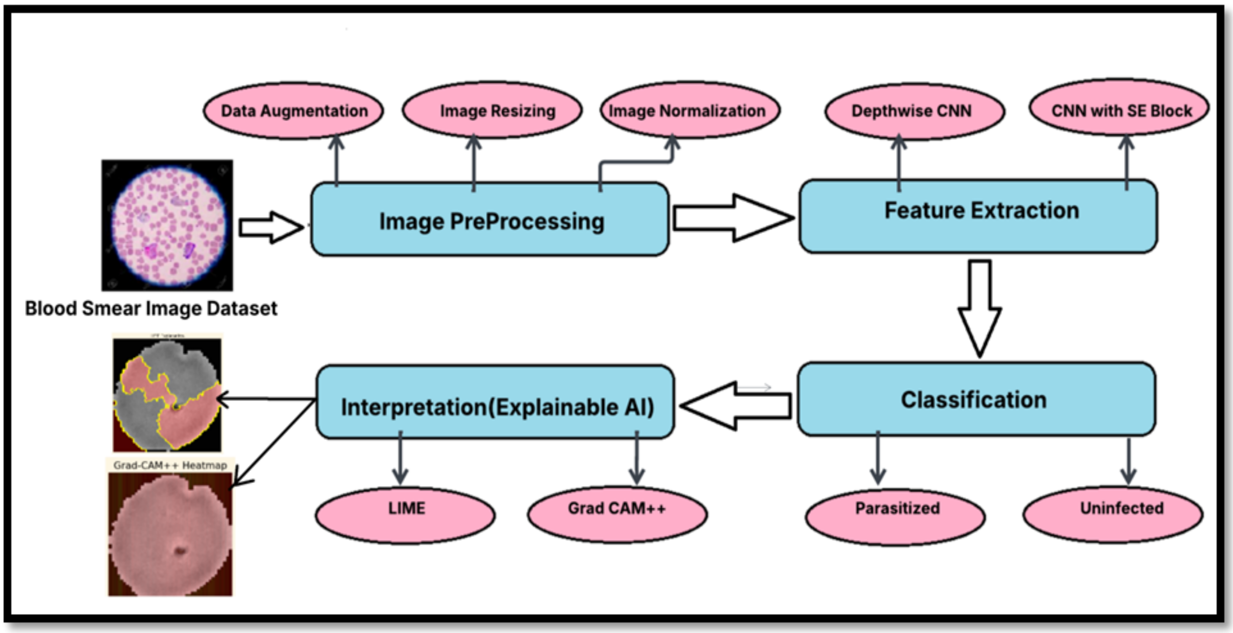


**Figure 5.** Proposed XMal-CNN Model Architecture for Malaria Detection

The proposed process includes dataset collection, image preprocessing (augmentation, normalization, resizing), and feature extraction leveraging CNNs augmented with Depthwise Convolutions and SE blocks. An optimized image classifier distinguishes infected and uninfected cells, suppressing false positive outcomes. Explainable AI methods like Grad-CAM++ and LIME are used to make the decision-making process visibly more accurate and enable trustworthiness. The architecture of the proposed model for malaria detection, henceforth named XMal-CNN, is tested on benchmark datasets. The model architecture, as depicted in Figure 5, shows outstanding performance in terms of diagnosis and is very well suited for implementation in low-resource medical setups and situations.

### A. Dataset Description

The data are very important and essential in the design of an AI-based malaria diagnosis system that provides automated and precise identification of malaria parasites, especially in resource-constrained health facilities.

In the present study, a publicly accessible dataset from the United States National Library of Medicine [27] has been used. This Malaria Cell Images Dataset [2] is the most widely used benchmark for the study of automatic malaria diagnosis systems. Figures 6 and 7 depict sample representative microscopic images of parasitized malarial cells and non-infected red blood cells from the dataset. The dataset consists of 27,558 high-quality annotated images of microscopic blood smears. The dataset is evenly distributed into two classes: (1) 13,779 images of infected red blood cells (with Plasmodium) and (2) 13,779 images of uninfected red blood cells. The class balance and high quality of the labels of the images make this dataset perfect for the training and testing of deep learning models.

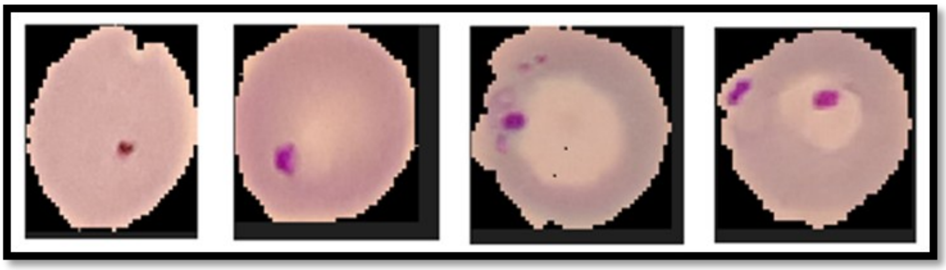

**Figure 6.** Microscopic Images of Parasitized Cells in Malaria

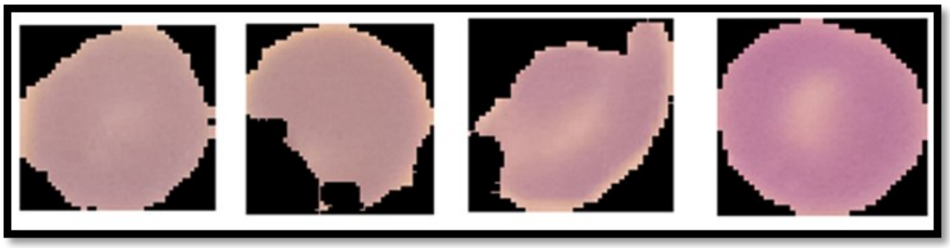

**Figure7.** Microscopic Images of Uninfected Red Blood Cells

### B. Data Pre-processing

The blood smear image, which contains microscopic red blood cells (RBCs) that can be infected with the Plasmodium parasites, is analyzed. Various pre-processing processes and techniques are used on images to improve their quality and boost the performance of the model. The following steps are used for data pre-processing:

**Data Augmentation:** Discrete geometric translations of dataset images are performed in data augmentation. E.g. flipping, rotation, and contrast changes are applied to enhance the model's ability to generalize images from sample to sample in each case. This practice avoids overfitting.

**Image Resizing:** All the images in the dataset are standardized to a specific resolution of 150×150 pixels to ensure consistency and compatibility with deep learning architecture.

**Image Normalization:** The pixel intensity levels of an image are normalized to a fixed interval for variance elimination. Variance in images results from light exposure and variation in staining while making blood smear slides.

## C. Feature Extraction

Image feature extraction is done using Deep Learning based Convolutional Neural Networks (CNNs). A suggested Depthwise CNN with Squeeze-and-Excitation Blocks improves computational efficiency as well as the high accuracy of malaria classification.

**Depthwise Separable Convolution:** Traditional Convolution layers contain a large number of parameters and therefore, they contain high computational complexity. Depthwise separable convolutions divide traditional convolutions into Depthwise Convolution and Pointwise Convolution. Depthwise Convolution, applies a single filter to all input channels, whereas Pointwise convolution applies a 1×1 convolution to collect outputs from Depthwise convolution. Mathematically, if the input features map has dimensions of $H \times W \times C_{in}$. This notation represents the shape of an input image or feature map in Deep Learning, where H denotes height, W denotes width, and $C_{in}$ denotes the number of input channels in the above notation (e.g., 3 for an RGB image) [10]. This format defines how data flows through convolutional layers, with operations being applied across both spatial dimensions and channels. In standard convolution, if the kernel size is K×K, the total number of computations required is $H \times W \times C_{in} \times K^2 \times C_{out}$, where $C_{out}$ denotes number of the output channels (filters). This shows the compuational cost of applying multiple filters (each spanning all input channels) across the entire input. In contrast, Depthwise separable Convolution reduces this compuational cost by dividing the operation into two parts: (1) depthwise convolution, which applies a single K×K filter per input channel, and (2) pointwise convolution, which uses 1×1 convolutions to combine outputs across channels. The total computation is thus reduced to $H \times W \times C_{in} \times K^2 + H \times W \times C_{in} \times C_{out}$, significantly improving the efficiency of the model, while maintaining the ability to extract spatial features from dataset images.

**Squeeze-and-Excitation (SE) Blocks:** SE blocks recalibrate the feature maps of images dynamically by giving weightage based on the importance of each channel. This is achieved through three steps:

**Squeeze operation:** It uses a global average pooling (GAP) to reduce spatial dimensions. The squeeze function is given by:

$$S_c = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} x_{i,j,c} \tag{1}$$

The global average pooling operation is applied to a convolutional feature map in a deep learning model. Specifically, it calculates the average value of all the elements within the spatial dimensions (height H and width W) of a particular channel c. Here, $x_{i,j,c}$ represents the pixel value at a position (i,j) in channel c and the double summation iterates over all spatial locations of that channel. By dividing the total sum by the product H×W, the formula gives a single scalar value $S_c$, which effectively summarizes the entire channel. This operation is commonly used in Squeeze-and-Excitation (SE) blocks, where it serves as the "squeeze" step, reducing each 2D feature map into a single representative value per channel. This allows the model to learn and apply attention to different channels by later "exciting" or reweighting them based on their importance, enhancing the network's ability to focus on more relevant features.

**Excitation operation:** A Fully Connected (FC) layer with two activation functions (ReLU and Sigmoid) is used to scale the extracted features:

$$z = \sigma(W_2 \times \delta(W_1 \times s)) \quad (2)$$

Where: W1 and W2 are weight matrice, δ denotes the ReLU activation function, σ denotes the Sigmoid activation function.

**Scaling operation:** Original feature maps of dataset images are multiplied by the recalibrated values to emphasize highly informative regions or features in the image.

## D. Classification

The extracted features are passed to a classifier that determines whether the cell is Infected (parasitized) or Uninfected (healthy). This classification helps in automated malaria diagnosis, reducing the need for manual expert analysis.

## E. Interpretation (Explainable AI - XAI)

To improve model explainability and gain the trust among healthcare professionals, the interpretability of malaria classification in proposed model was introduced using various advanced Explainable AI models such as LIME and Grad-CAM++.

LIME: LIME is an approach that breaks down individual prediction of any Machine Learning model locally by approximating it with an interpretable model. It proceeds by perturbing the input data and noticing the changes in the predictions and thus determining the most influential pieces of the input in the model's decision. This method is handy in the case of complex, black box models, as it gains insights into what they do without having access to their internal operation [8].

Grad-CAM++ (Gradient-weighted Class Activation Mapping): Grad-CAM++is a refined and advanced version of the Grad-CAM algorithm, designed to provide visual explanations for Convolutional Neural Network (CNN) decision-making [29]. Grad-CAM++ generates heat maps that show the regions of an image which is most relevant to the model's prediction. By computing the gradients of the output with respect to feature maps in the final convolutional layer, Grad-CAM++ identifies the image regions with the highest influence on the classification outcome or decision making. It increases the interpretability of CNNs, so their decisions simpler to understand and believable.

By using these representations provided by XAI methods, medical professionals can understand the decision making of AI-based diagnosis models with better clarity, improving trust and results in more accurate use of AI-based diagnostics.

## 4. Results and Discussion

Execution of the suggested model was done by utilizing the Kaggle platform. An NVIDIA A800 GPU with 16 GB RAM was used to facilitate smooth training and testing of the suggested method.

**Performance Metrics:** Proper and suitable performance metrics are needed to precisely measure the reliability and applicability of the suggested malaria detection model in real-world situations, such as diagnosis scenarios. The suggested model for malaria detection is assessed in terms of different significant performance metrics [10]

**Accuracy (Overall Classification Performance):** Accuracy, in plain terms, estimates the ratio of instances or outcomes, both malaria-infected and not infected, that are appropriately classified out of the total cases. Accuracy provides an overall impression of how correct the model's results are in classification.

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \quad (3)$$

Here, TP stands for True Positives (cases of malaria that were correctly predicted), TN for True Negatives (cases of non-malaria that were correctly predicted), FP for False Positives (cases of malaria that were incorrectly predicted), and FN for False Negatives (cases of malaria that were missed).

**Positive Predictive Value, or precision**: Precision measures the proportion of correctly predicted malaria cases to all positive cases. Precision demonstrates the suggested model's capacity to reduce erroneous predictions or errors, preventing a healthy patient from receiving an incorrect diagnosis. The precision formula in mathematics is:

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \quad (4)$$

where FP stands for False Positives and TP for True Positives. The number of predicted positive cases that are actually cases of malaria is known as precision. To put it simply, high precision means that the model makes fewer false positives, or that fewer healthy patients are incorrectly identified as infected.

**True Positive Rate/Sensitivity Recall:** The model's accuracy in classifying actual cases of malaria infection is gauged by recall. It shows how well the model detects all true positive infections and the fewest false negative cases overall. Recall can be expressed mathematically as:

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \quad (5)$$

where FN stands for False Negatives and TP for True Positives. The recall of the model shows how well it can identify actual cases of malaria worldwide. A high and robust catch rate of infected cases and a low percentage of false negatives or omitted infections can be inferred from high recall.

**F1 Score:** When class distributions are skewed or when false positives and false negatives are both critically important, the F1 Score (Harmonic Mean of Precision and Recall) is a very useful evaluation metric that best balances precision and recall. The F1 score, which represents the trade-off between precision and recall, is mathematically calculated as the harmonic mean of the two.

$$\text{F1 Score} = 2 * \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (6)$$

Recall and precision are balanced in the F1 score. It ensures that neither false positives nor false negatives dominate the evaluation, which is especially helpful when the dataset is unbalanced.

**ROC-AUC (Receiver Operating Characteristic-Area Under Curve):** The ROC-AUC metric measures the diagnostic ability of the model by plotting the trade-off between sensitivity and the false positive rate across different classification thresholds. The ROC (Receiver Operating Characteristic) curve depicts the relationship between the True Positive Rate (TPR) and the False Positive Rate (FPR).

True Positive Rate (TPR) / Recall:

$$\text{TPR} = \frac{\text{TP}}{\text{TP+FN}} \quad (7)$$

False Positive Rate (FPR):

$$\text{FPR} = \frac{\text{FP}}{\text{FP+TN}} \quad (8)$$

The ROC Curve indicates the True Positive Rate (Recall) against the False Positive Rate at different classification thresholds. The AUC (Area Under the Curve) value describes the overall model performance. The higher value of the AUC closer to 1 implies a high discrimination to differentiate between malaria infected and non-infected cases.

For improved model performance, the Adam optimizer with adaptive learning rate (as 0.001 initially) is utilized, supporting faster convergence by using momentum and gradient updates. An 80:20 train-validation split with stratified sampling ensures balanced class distribution and precise generalization of evaluation.

Experimental assessment involves comparing various CNN models based on two Explanation methods: LIME and Grad-CAM++. LIME indicates the importance of local features by input perturbation, whereas Grad-CAM++ shows visual information about the crucial image areas driving the predictions. Model performance is assessed in terms of metrics such as accuracy, precision, recall, and F1-score in identifying parasitized and uninfected cells, and comparisons are made between both explanation approaches.

**Table 1.** Performance Evaluation of Different CNN Variants

| **Models** | **Accuracy (%)** | **Precision** | **Recall** | **F1-Score** |
|---|---|---|---|---|
| Custom CNN | 93.03 | 91.73 | 94.41 | 94.10 |

| | | | | |
|---|---|---|---|---|
| SE Block CNN | 94.69 | 92.81 | 94.93 | 94.84 |
| Depthwise CNN | 50.02 | 50.00 | 01.00 | 00.67 |
| Depthwise CNN + SE Block(Xmal- CNN) | 95.26 | 93.97 | 96.73 | 95.33 |

The performance comparison analysis of different CNN variants is presented in Table 1. The proposed XMal-CNN (Depthwise CNN with SE Block) delivers the best results, achieving 95.26% accuracy, 93.97% precision, 96.73% recall, and a 95.33% F1-score. The SE Block with CNN also performs well with 94.69% accuracy, highlighting the effectiveness of attention mechanisms. In contrast, the standalone Depthwise CNN performs poorly, with only 50.02% accuracy and an F1-score of 0.67, suggesting underfitting and underperformance. The Custom CNN shows moderate performance (93.03% accuracy) but is outperformed by both SE Block CNN and XMal-CNN.

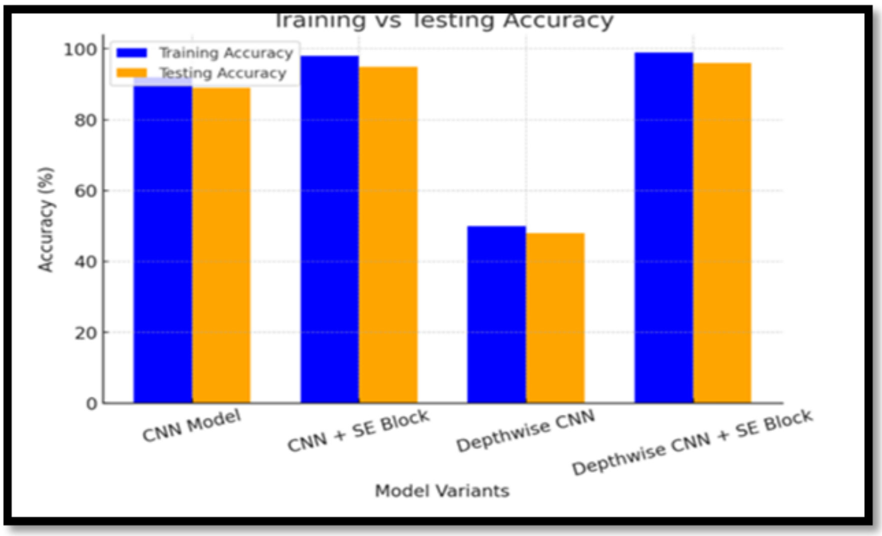


**Figure 8.** Accuracy Comparison of Different CNN Model Variants

Figure 8 illustrates the training and testing accuracy of different CNN-based model variants named standard CNN, CNN with SE Block, Depthwise CNN, and Depthwise CNN with SE Block. All models demonstrate strong performance, with the CNN and CNN + SE Block models achieving good accuracy levels, indicating effective feature learning and classification capabilities. The Depthwise CNN + SE Block model significantly outperforms the baseline Depthwise CNN, achieving 95.26% testing accuracy. This improvement demonstrates the SE Block's effectiveness in enhancing feature selection by adaptively recalibrating channel-wise responses. In contrast, the basic Depthwise CNN shows poor performance (~50% accuracy), likely due to limited parameter complexity. Adding the SE Block addresses this limitation by emphasizing important features, leading to a substantial boost in performance.

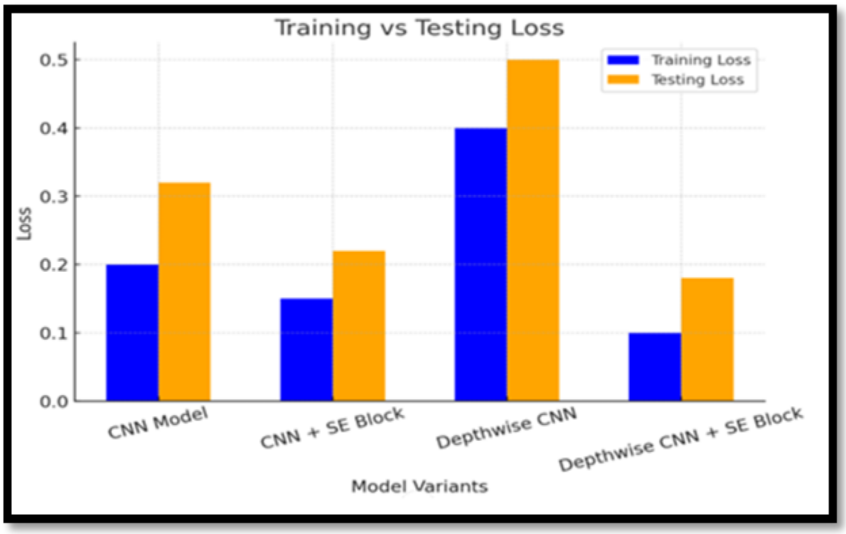


**Figure 9.** Loss Comparison of Different CNN Model Variants

Figure 9 presents the training and testing losses of different CNN model variants. The Depthwise CNN + SE Block model achieves the lowest training and testing losses, confirming its effectiveness in minimizing errors and enhancing generalization. This highlights that integrating an SE Block with Depthwise CNN significantly enhances the model's performance, striking a balance between computational efficiency and accuracy.

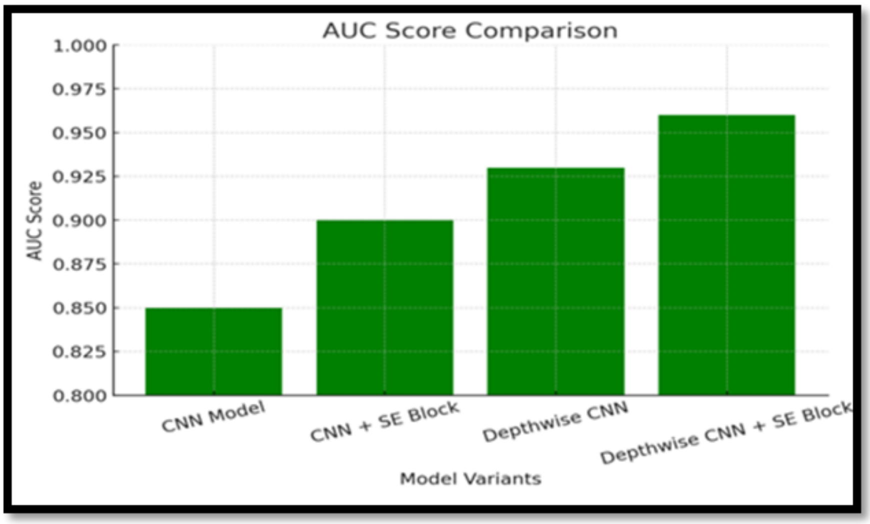


**Figure 10.** Comparison of AUC Scores Across CNN Model Variants

Moreover, Figure 10 presents a comparison of AUC scores of different CNN model variants, highlighting that Depthwise CNN + SE Block achieves the highest AUC(~0.96), indicating superior classification performance.

The confusion matrix shown in Figure 11 reflects the model's effectiveness in distinguishing between parasitized (infected) and uninfected cells. Despite a few false positives, the model demonstrates strong and reliable classification performance. The model correctly identifies 2,665 parasitized cases as true positives (TP) and 2,584 uninfected cases as true negatives (TN).

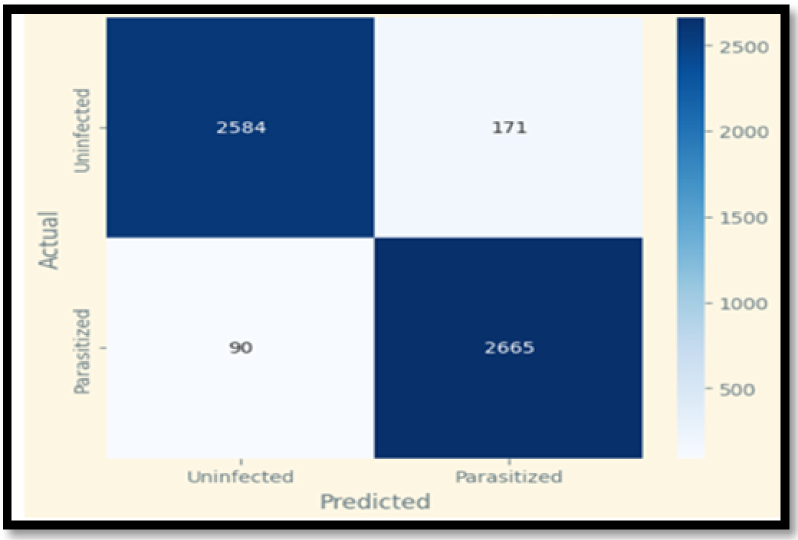


**Figure 11.** Confusion Matrix of XMal-CNN for Malaria Detection

It records 171 false positives (FP), where uninfected cases are misclassified as parasitized, and 90 false negatives (FN), where parasitized cases are incorrectly labeled as uninfected. Although a slightly elevated FP rate suggests occasional misclassification of healthy samples, the model maintains excellent reliability, interpretability, and diagnostic effectiveness.

Additionally, explainable AI models such as LIME and Grad-CAM++ are used to provide visual interpretability of the proposed model. Figure 12 illustrates the model's interpretability using LIME and Grad-CAM++. Figure 12(a) shows the original grayscale scan, with a dark central region likely indicating a lesion. Figure 12(b) presents visual interpretation using the LIME model, highlighting key areas influencing the model's decision, with red indicating high importance and grey indicating less importance, outlined by yellow contours. Figure 12(c) shows the visual interpretation of the proposed model using the Grad-CAM++ heatmap, where bright red and yellow mark regions of high model attention. Together, LIME and Grad-CAM++ enhance explainability by confirming that the model's focus aligns with medically relevant features, which is crucial for building trust and supporting clinical decisions.

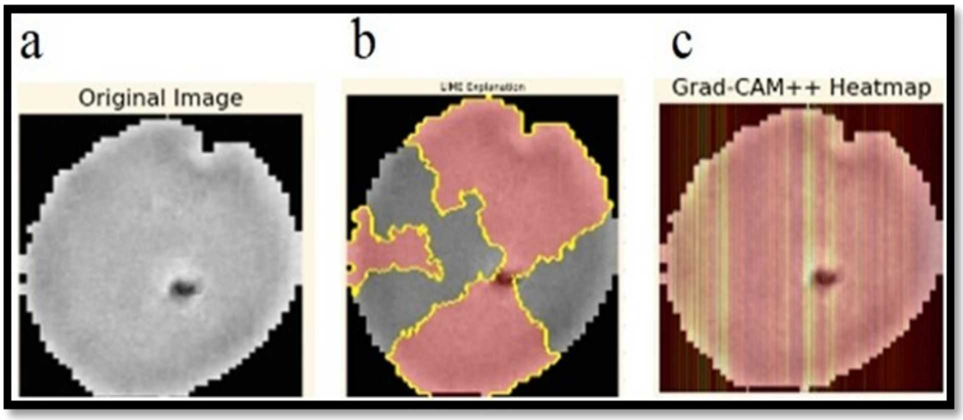


**Figure 12.** Visual Interpretability Comparison: (a)Original Image, (b)LIME Explanation, and (c) Grad-CAM++ Heatmap

**Table 2.** Comparative Analysis of the Proposed Model with State-of-the-Art Models

| Methods | Accuracy |
|---|---|
| Malaria Detection custom CNN model [12] | 90% |
| Transfer Learning with Pre-trained CNN Models [24] | 91.5% |
| YOLOv5 and Attention Models [10] | 92.00% |
| ML - Based Malaria Prediction [14] | 93.03% |
| DL Features + ML Classifiers[17] | 92.5% |
| Proposed XMal-CNN | 95.26% |

Despite obtaining state-of-the-art performance on the benchmark dataset, the proposed XMal- The CNN model's generalizability to other real-world malaria classification datasets has to be confirmed. Table 2 highlights a comparison of the proposed XMal-CNN model with the most sophisticated malaria detection models. The comparison indicates that the proposed XMal-CNN model is more efficient in detecting malaria compared to existing techniques. Accuracy varies from 90% to 91.5% for traditional CNN models and transfer learning methods, whereas 92% to 93% for more complex models like YOLO-based and ML hybrid models. With the highest accuracy rate of 95.26%, XMal-CNN proves itself to be an outstanding feature extraction and classification network. Squeeze-and-Excitation blocks and Depthwise Convolutions, responsible for enhancing feature representation as well as model accuracy, are the main cause behind the improvement. Also, the more complicated task of multiclass classification during analysis, such as distinguishing between specific Plasmodium species, has not yet been addressed by the existing model, which is designed for binary classification (infected and uninfected classification). To improve the clinical applicability of the model, future research can investigate and expand this architecture to use more diverse datasets and increase the number of diagnostic categories. Although XMal-CNN demonstrates good diagnostic performance, its inference time and memory requirements on low-power devices have not yet been addressed, as the study focuses on accuracy and interpretability. The use of depth-wise convolutions with SE suggests potential for efficient deployment, and future work will involve assessing scalability on hardware-constrained systems.

## 5. Conclusion

This work proposes XMal-CNN, an interpretable deep learning architecture that uses depth-wise convolutional neural networks (CNNs) and squeeze-and-excitation (SE) blocks to perform automatic malaria diagnosis from microscopic blood smear images. It can be concluded that, with 95.26% accuracy, 93.97% precision, 96.73% recall, and an F1-score of 95.33%, the proposed architecture outperforms current state-of-the-art methods and conventional CNN models in some evaluation metrics. One of the key contributions of the current work is the employment of explainable AI (XAI) techniques like LIME and Grad-CAM++ that visually analyze and identify the parts of the input images that influence the predictions. These techniques make the existing models more transparent and reliable. In clinical practice, healthcare applications of AI-based systems need end-result interpretability to ensure that decisions made by AI do not vary from clinical decision-making and knowledge-based reasoning. Additional future research will explore multiclass image classification to separate different Plasmodium species, improve the model to support edge and mobile deployment, and assess its performance based on real clinical datasets. International initiatives to eradicate malaria can be significantly supported by XMal-CNN, a new AI-aided diagnostic system that facilitates quality and affordable healthcare provision.

**References**

[1] World Health Organization. (2023). World Malaria Day 2023: Background. https://www.who.int/campaigns/world-malaria-day/2023/background

[2] Hawkes, Michael, and Kevin C. Kain. "Advances in malaria diagnosis." Expert review of anti-infective therapy 5, no. 3 (2007): 485-495.

[3] Shambhu, Shankar, Deepika Koundal, Prasenjit Das, Vinh Truong Hoang, Kiet Tran-Trung, and Hamza Turabieh. "Computational methods for automated analysis of malaria parasite using blood smear images: recent advances." Computational intelligence and neuroscience 2022, no. 1 (2022): 3626726.

[4] Fisher, J. F. (2022). Testing for malaria: Rapid diagnostic tests and PCR. Medmastery. Retrieved August 25, 2025, from https://www.medmastery.com/guides/malaria-clinical-guide/testing-malaria-rapid-diagnostic-tests-rdts-and-polymerase-chain

[5] Thakur, Gopal Kumar, Abhishek Thakur, Shridhar Kulkarni, Naseebia Khan, and Shahnawaz Khan. "Deep learning approaches for medical image analysis and diagnosis." Cureus 16, no. 5 (2024).

[6] Kong, Mingui, and Su Jeong Song. "Artificial Intelligence Applications in Diabetic Retinopathy: What We Have Now and What to Expect in the Future." Endocrinology and Metabolism 39, no. 3 (2024): 416-424.

[7] Goni, Md Omaer Faruq, and Md Nazrul Islam Mondal. "Explainable AI based malaria detection using lightweight CNN." In 2023 International Conference on Next-Generation Computing, IoT and Machine Learning (NCIM), IEEE, (2023): 1-5.

[8] Attai, Kingsley, Moses Ekpenyong, Constance Amannah, Daniel Asuquo, Peterben Ajuga, Okure Obot, Ekemini Johnson et al. "Enhancing the interpretability of malaria and typhoid diagnosis with explainable AI and large language models." Tropical Medicine and Infectious Disease 9, no. 9 (2024): 216.

[9] Sidharthan, C. (2024, June 12). AI revolutionizes malaria diagnosis with 97.57% accuracy using EfficientNet. News-Medical. https://www.news-medical.net/news/20240612/AI-revolutionizes-malaria-diagnosis-with-975725-accuracy-using-EfficientNet.aspx

[10] Liu, R., Liu, T., Dan, T., Yang, S., Li, Y., Luo, B., Zhuang, Y., Fan, X., Zhang, X., Cai, H., & Teng, Y. (2023). AIDMAN: An AI-based object detection system for malaria diagnosis from smartphone thin-blood-smear images. Patterns, 4(9), 100806. https://doi.org/10.1016/j.patter.2023.100806

[11] Kunwar, S. (2018). Malaria detection using image processing and machine learning. arXiv (Cornell University). https://doi.org/10.48550/arxiv.1801.10031

[12] Motwani, K., Kanojiya, A., Gomes, C., & Yadav, A. (2021). Malaria detection using image processing and machine learning. International Journal of Engineering Research & Technology (IJERT), 9(3). https://www.ijert.org/malaria-detection-using-image-processing-and-machine-learning

[13] Srivastava, S., & Debnath, D. (2020). Detection of malaria using machine learning. International Journal for Research in Applied Science & Engineering Technology, 8(7), 295–303. https://doi.org/10.22214/ijraset.2020.730310

[14] Lee, Y. W., Choi, J. W., & Shin, E. (2020). Machine learning model for predicting malaria using clinical information. Computers in Biology and Medicine, 129, 104151. https://doi.org/10.1016/j.compbiomed.2020.104151.

[15] Shekar, G., Revathy, S., & Goud, E. K. (2020). Malaria Detection using Deep Learning. 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI), 746–750. https://doi.org/10.1109/icoei48184.2020.9143023

[16] Abubakar, A., Ajuji, M., & Yahya, I. U. (2021). DeepFMD: Computational analysis for malaria detection in Blood-Smear images using Deep-Learning features. Applied System Innovation, 4(4), 82. https://doi.org/10.3390/asi4040082

[17] Nayak, S., Kumar, S., & Jangid, M. (2019). Malaria Detection Using Multiple Deep Learning Approaches. 2019 2nd International Conference on Intelligent Communication and Computational Techniques (ICCT), 292–297. https://doi.org/10.1109/icct46177.2019.8969046

[18] Madhu, G., Mohamed, A. W., Kautish, S., Shah, M. A., & Ali, I. (2023). Intelligent diagnostic model for malaria parasite detection and classification using imperative inception-based capsule neural networks. Scientific Reports, 13(1). https://doi.org/10.1038/s41598-023-40317-z

[19] Hoyos, K., & Hoyos, W. (2024). Supporting malaria diagnosis using deep learning and data augmentation. Diagnostics, 14(7), 690. https://doi.org/10.3390/diagnostics14070690

[20] Hemachandran, K., Alasiry, A., Marzougui, M., Ganie, S. M., Pise, A. A., Alouane, M. T., & Chola, C. (2023). Performance analysis of deep learning algorithms in diagnosis of malaria Disease. Diagnostics, 13(3), 534. https://doi.org/10.3390/diagnostics13030534

[21] Alnussairi, M. H. D., & İbrahim, A. A. (2022). Malaria parasite detection using deep learning algorithms based on (CNNs) technique. Computers & Electrical Engineering, 103, 108316. https://doi.org/10.1016/j.compeleceng.2022.108316

[22] Kumar, S., Priya, S., & Kumar, A. (2023). Malaria detection using Deep Convolution Neural Network. arXiv (Cornell University). https://doi.org/10.48550/arxiv.2303.03397

[23] Taye, A. G., Yemane, S., Negash, E., Minwuyelet, Y., Abebe, M., &Asmare, M. H. (2024, June 27). Automated Web-Based Malaria Detection System with Machine Learning and Deep Learning Techniques. arXiv.org. https://arxiv.org/abs/2407.00120

[24] Sukumarran, D., Hasikin, K., Khairuddin, A. S. M., Ngui, R., Sulaiman, W. Y. W., Vythilingam, I., & Divis, P. C. S. (2024). An optimised YOLOv4 deep learning model for efficient malarial cell detection in thin blood smear images. Parasites & Vectors, 17(1). https://doi.org/10.1186/s13071-024-06215-7.

[25] Mujahid, M., Rustam, F., Shafique, R., Montero, E. C., Alvarado, E. S., De La Torre Diez, I., & Ashraf, I. (2024). Efficient deep learning-based approach for malaria detection using red blood cell smears. Scientific Reports, 14(1). https://doi.org/10.1038/s41598-024-63831-0

[26] Modiya, P., & Vahora, S. (2022, May 27). Brain Tumor Detection Using Transfer Learning with Dimensionality Reduction Method. https://www.ijisae.org/index.php/IJISAE/article/view/1310

[27] Lister Hill National Center for Biomedical Communications. Malaria datasheet. National Library of Medicine. Retrieved August 25, 2025, from https://lhncbc.nlm.nih.gov/LHC-research/LHC-projects/image-processing/malaria-datasheet.html

[28] Isozumi, R., Fukui, M., Kaneko, A., Chan, C. W., Kawamoto, F., & Kimura, M. (2014). Improved detection of malaria cases in island settings of Vanuatu and Kenya by PCR that targets the Plasmodium mitochondrial cytochrome c oxidase III (cox3) gene. Parasitology International, 64(3), 304–308. https://doi.org/10.1016/j.parint.2014.09.006

[29] Chattopadhay, A., Sarkar, A., Howlader, P., & Balasubramanian, V. N. (2018). Grad-CAM++: Generalized Gradient-Based Visual Explanations for Deep Convolutional Networks. https://doi.org/10.1109/wacv.2018.00097.