Improving Malaria detection using enhanced-efficientnet deep neural network approach

Vipin Kataria[1], Nitin Kumar[2*], Parth Patel[3]

[1,2]*University of Illinois Urbana Champaign, USA.*
[3]*Southern Methodist University Dallas, USA.*

Corresponding author: Nitin Kumar (*Email: kumar.nitin0710@gmail.com*)

Abstract

Malaria detection traditionally relies on microscopic examination of blood smears, a process that is labor-intensive and prone to human error. This study aims to introduce a robust automated detection method using deep learning, designed to enhance diagnostic accuracy and reduce human effort. The research presents an innovative Enhanced-EfficientNet (EEN) deep neural network approach comprising three distinct phases: image preprocessing, feature extraction using the Enhanced-EfficientNet model, and classification using a Deep Neural Network (DNN). The proposed methodology was validated using a dataset containing 27,558 labeled blood cell images equally divided between "infected" and "uninfected" samples. The proposed EEN approach achieved superior diagnostic performance, with a maximum classification accuracy of 97.71%, precision of 97.71%, recall of 97.72%, and an F1 score of 97.71% on the test dataset. Comparative evaluation with established models, including VGG16, Xception, ResNet152, EfficientNetB3, and InceptionV3, confirmed significant performance improvements offered by the proposed method. The Enhanced-EfficientNet model effectively addresses the accuracy and reliability challenges associated with traditional malaria diagnostics, presenting a robust deep learning alternative with improved diagnostic outcomes. The study underscores deep learning's practical value as a supportive diagnostic tool, facilitating quicker, more reliable detection of malaria infections. Clinicians can leverage this technology to enhance patient care, significantly reduce diagnostic errors, and improve survival outcomes for patients.

Keywords: Deep neural network, Machine learning, Malaria detection.

Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: All authors contributed equally to the conception and design of the study. All authors have read and agreed to the published version of the manuscript.
Transparency: The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.

1. Introduction

The 2023 WHO World Malaria Report highlights a concerning increase in global malaria cases in 2022, surpassing pre-COVID-19 pandemic levels from 2019, and underscores several threats to the global malaria response, including climate change [1]. Malaria remains a significant global health challenge, with an estimated 249 million cases and 608,000 deaths reported in 2022, predominantly affecting the WHO African Region, which accounts for 95% of cases and a substantial number of childhood deaths [2]. Despite efforts to control malaria, the disease continues to pose a major threat due to factors such as weak healthcare systems, resistance to drugs and insecticides, and inadequate funding [3]. The report also emphasizes the importance of accurate diagnosis and effective treatment strategies, noting the need for new technologies and integrated approaches to combat the resurgence of malaria [4].

Malaria detection faces numerous challenges, primarily due to the limitations of current diagnostic methods and the complexities of implementing these technologies in resource-constrained settings. Traditional methods like microscopy and rapid diagnostic tests (RDTs) are widely used but have significant drawbacks. Microscopy, while considered the gold standard, requires trained personnel and is time-consuming, making it less feasible in many endemic regions [5, 6]. RDTs, although more accessible and quicker, suffer from issues such as low sensitivity at low parasite densities, cross-reactivity, and the prozone effect, which can lead to false negatives [7, 8]. Molecular techniques such as PCR and LAMP offer higher sensitivity and specificity but are hindered by the need for specialized equipment and expertise, limiting their use in field conditions [6, 7]. The emergence of drug-resistant Plasmodium strains further complicates the diagnostic landscape, necessitating more precise and reliable detection methods to guide treatment decisions [9, 10]. Recent advancements in biomarker research and the integration of artificial intelligence hold promise for overcoming these challenges by providing more accurate, rapid, and non-invasive diagnostic options [11]. However, the adoption of these innovations is hampered by infrastructural and economic barriers, particularly in low-income regions where malaria is most prevalent [8, 10]. Additionally, the diagnostic pathway from innovation to adoption is complex, involving multiple stakeholders and requiring robust evidence generation and data sharing to ensure alignment with user needs and regulatory requirements [12]. Addressing these challenges requires a multifaceted approach that includes strengthening healthcare infrastructure, enhancing stakeholder collaboration, and developing cost-effective, scalable diagnostic tools that can be deployed in diverse settings [9, 13].

Machine learning (ML) has emerged as a transformative tool in the detection and prediction of malaria, offering significant improvements over traditional diagnostic methods. Traditional methods such as rapid diagnostic tests (RDTs) and light microscopy, while effective, are labor-intensive and prone to human error, necessitating the development of automated systems for more efficient malaria diagnosis [14]. Machine learning techniques, including logistic regression, random forest, support vector machines (SVM), and artificial neural networks (ANN), have been employed to predict malaria prevalence and incidence by analyzing various factors such as climatic, epidemiological, and socioeconomic data [15, 16]. For instance, in Rwanda, logistic regression models have been used to predict malaria prevalence by considering factors like region, wealth index, and preventive measures, thereby aiding in resource allocation and intervention planning [15]. In Indonesia, ML models, particularly random forests, have demonstrated strong predictive performance by integrating multidimensional data sources, highlighting the potential of ML in enhancing malaria surveillance and control efforts [16]. Furthermore, advanced techniques such as convolutional neural networks (CNNs) and deep learning have been applied to automate the detection of malaria parasites in blood smear images, achieving high accuracy and efficiency. These AI-driven approaches, including models like VGG19, InceptionV3, and Xception, have shown promising results, with some achieving accuracy rates as high as 97% [17, 18]. The use of digital twins and gradient boosting techniques has also been explored, achieving remarkable accuracy in predicting malaria incidence, thus confirming their practical applicability for prompt epidemic responses [19]. Additionally, the development of web-based and mobile applications for malaria detection addresses resource and expertise limitations in developing countries, further demonstrating the scalability and accessibility of ML-based systems [20]. Overall, the integration of machine learning in malaria detection and prediction not only enhances diagnostic accuracy and efficiency but also supports effective public health decision-making and resource management, ultimately contributing to the global fight against malaria [21, 22].

Recent advancements in machine learning for malaria detection have significantly enhanced the accuracy and efficiency of diagnostic methods, primarily through the application of deep learning and convolutional neural networks (CNNs). These technologies have been pivotal in automating the analysis of blood smear images, traditionally a manual and error-prone process. For instance, the EfficientNetB3 model has been refined to achieve a 97% accuracy in classifying blood smears into parasitized and uninfected categories, demonstrating its potential in resource-limited settings [23]. Similarly, the MosquitoNet CNN architecture, utilizing data augmentation and GPU acceleration, has achieved over 96.5% accuracy, highlighting the robustness of CNNs in malaria detection [24]. Other models, such as those employing ResNet-50, MobileNet-v2, and Inception-v3, have also demonstrated high efficacy, with some achieving perfect accuracy scores across various statistical matrices [18]. The integration of transfer learning and adaptive boosting further enhances these models' adaptability and accuracy, as seen in studies using VGG16 and other CNN architectures [25]. Additionally, the use of Support Vector Machines (SVM) in conjunction with CNNs has been explored, achieving up to 96% accuracy, which alleviates some of the time complexity issues associated with deep learning methods [26]. Beyond image-based diagnostics, advancements in biomarker research, such as the identification of volatile organic compounds (VOCs) in breath, offer promising non-invasive diagnostic tools, potentially revolutionizing malaria detection [11]. These machine learning advancements not only improve diagnostic accuracy but also facilitate the development of scalable, web-based systems that can be accessed remotely, thus broadening the reach of malaria diagnostics in underserved areas [17]. Overall, the integration of machine

learning in malaria detection is poised to significantly impact global health by enabling timely and accurate diagnosis, crucial for effective disease management and control.

The purpose of this research is to propose a novel approach using Deep Learning techniques aimed at improving the identification of malaria parasites in blood smear microscopic images. The proposed approach makes the following contributions to the field of malaria detection.

1. Developed a novel deep learning technique that extracts features from images of parasitized and uninfected cells using the Enhanced-EfficientNetB7 (EEN) network, then feeds those features into a Deep Neural Network for classification to categorize each image.
2. Data Preprocessing enhances the machine learning model's ability to generalize within the proposed approach.
3. Implemented Batch Normalization, Dropout, L2 Regularization, and Layer Normalization which improves model robustness and accuracy.
4. Performed extensive analysis of model metrics such as accuracy, F1 Score, Precision, Recall, Confusion Matrix and AUC ROC Curve to evaluate the performance of the model.
5. Comparison of the model's performance against other models. A high classification accuracy of 97.71% has been obtained.

The rest of the discussion is organized as follows: In the section, Material and Method the dataset used for models is discussed, and the proposed approach is outlined. In the section, Results, a detailed description of experimental results and model evaluation is provided. In the section, Discussion, we compared our proposed model results with other current research work results and other state-of-the-art models' results.

2. Materials and Methods

The dataset used in the current study is from a publicly available dataset [27]. The dataset contains 27,558 images (13,779 images each for "infected" and "uninfected") of blood cell images. The dataset is substantial enough to train a deep learning model, with each category containing hundreds of images. The dataset is divided into 80% for training, 10% for testing, and 10% for validation.

In the current research, we have developed a new technique for Malaria detection. Below is a detailed workflow of the proposed technique to detect Malaria when given a cell image, as shown in the Figure 1.
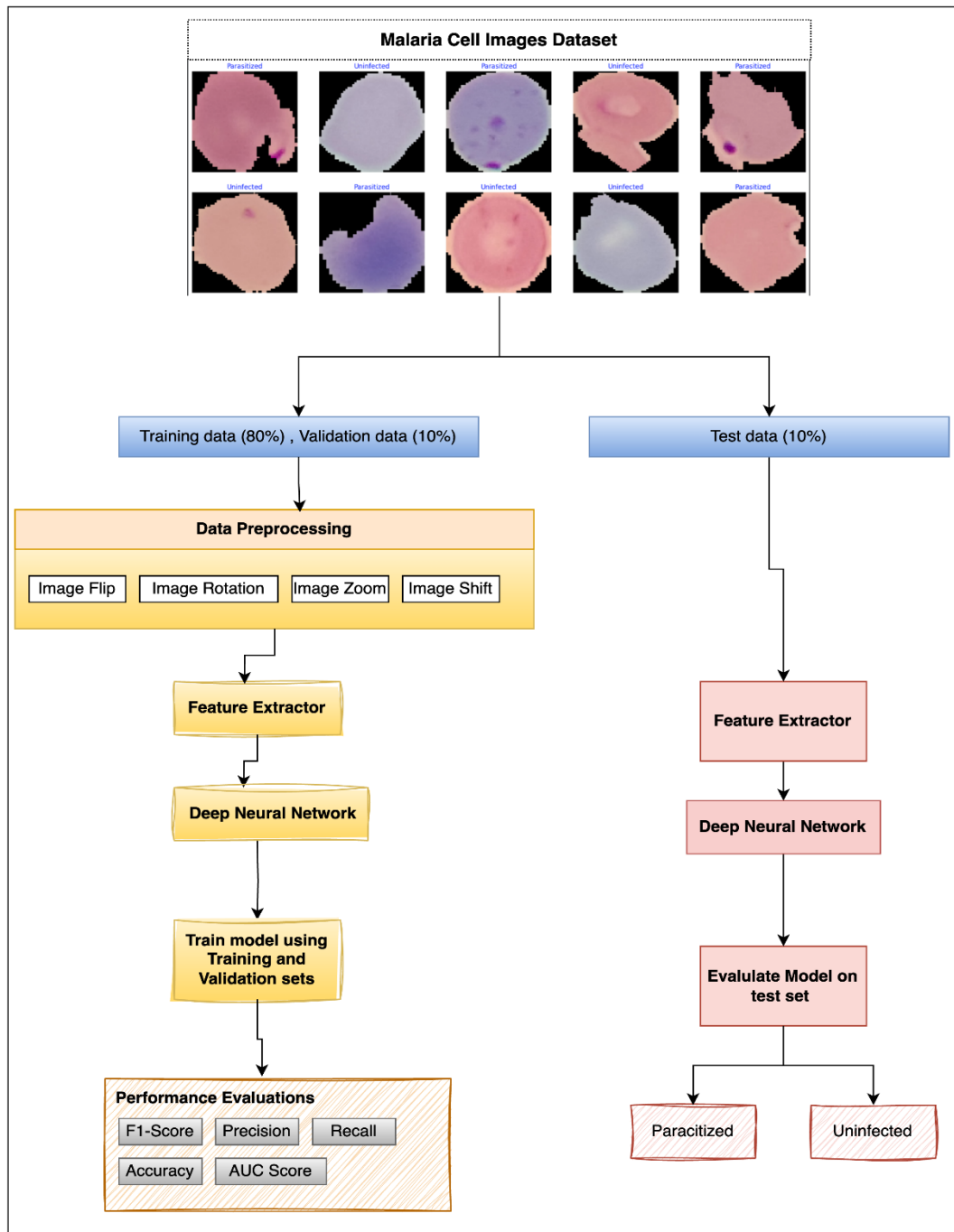
Figure 1.
Workflow of the proposed approach.

The proposed method performs Malaria detection through the following steps:
1. Data Preprocessing
2. Feature Extractor using EEN Network
3. Classification based on Deep Neural Network

   According to the Figure 1 , our proposed method for malaria detection consists of several key stages. The process begins with preprocessing malaria cell images through data augmentation techniques to enhance the generalizability of our approach. These preprocessed images are then passed to the feature extractor: EEN Network. The extracted features, which form a comprehensive feature representation, are fed into a deep neural network for final classification. The following sections provide detailed explanations of each stage in this pipeline.

2.1. Data Preprocessing
   The step-1 of the proposed method involves preprocessing the images of parasitized cells and images of uninfected cells, which are equally balanced, to enhance the generalization capability. Multiple image augmentation strategies such as rotation, flipping, zooming, and shifting, are applied. Computer vision models require extensive datasets covering all visual aspects of the target image. In the real world, data collection requires manual capturing and annotations of images, and it is impossible

to capture every single scenario. Image augmentation generates new, transformed versions of existing images to increase dataset diversity, reduce overfitting, and save time by avoiding the need to capture an exhaustive set of real-world variations manually. These variations mirror real-world conditions in cell imaging, including differences in illumination intensity, cell density, morphological features, background noise, focal planes, staining intensity, cellular orientation, microscope settings, image resolution, sample preparation methods, and environmental conditions.

2.2. Feature Extractor

In the step-2 of the proposed method, the preprocessed images are fed to the EEN network which is based on deep neural network. The detail is depicted in the Figure 2.

EEN network is made of EfficientNetB7 [28] network which is connected to additional convolutional layers of 5 x 5 x 64 and 5 x 5 x 32, followed by $3 \times 3 \times 32$ max pooling layers. The output of the max pooling layer is flattened to represent the features in a single dimension.
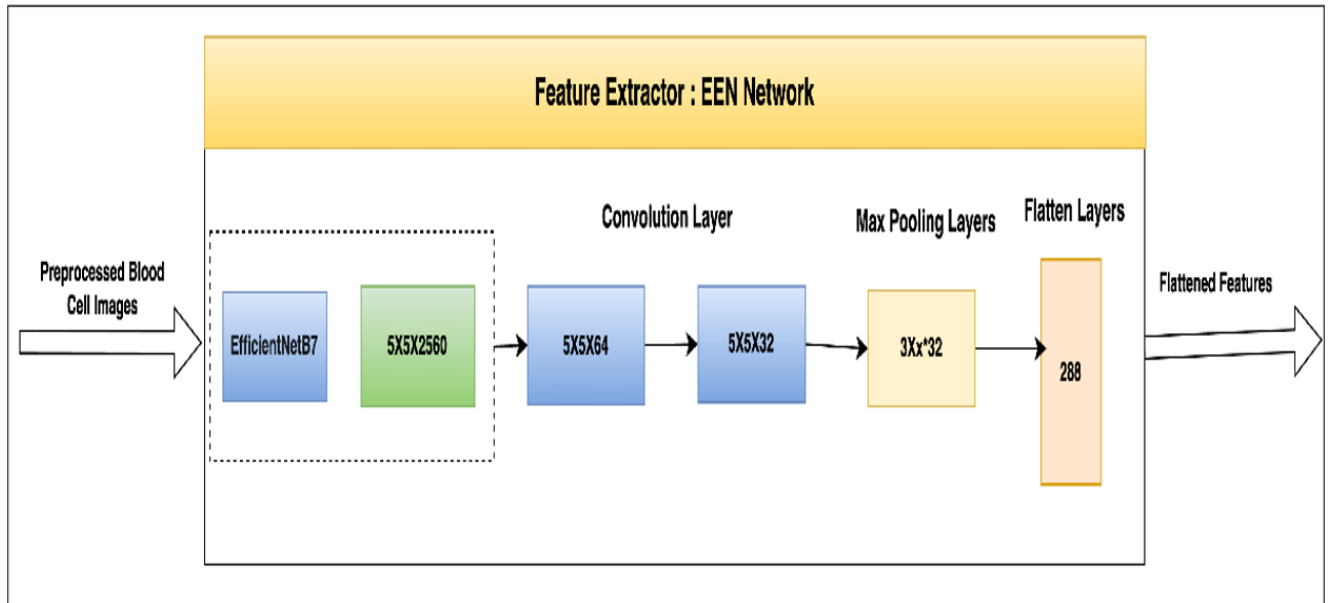

Figure 2.
Feature Extractor, EEN Network.

EfficientNetB7 is another model from the EfficientNet family of convolutional neural networks. It achieves a top-1 accuracy of 84.3% to 84.4% and a top-5 accuracy of 97.0% to 97.1% on the ImageNet dataset, making it one of the top-performing models in its class. It uses a compound scaling method which uniformly scales the network across depth, width, and resolution. This makes the model optimize its computational resources. The architecture is based on Mobile Inverted Bottleneck (MBConv) layers and incorporates Squeeze-and-Excitation (SE) optimization to improve the model's performance. Table 1 details out the layer type, output of each layer and parameters used in the EEN network.

Table 1.
The number of layers, types, output shape and its parameters for the proposed EEN Network.

Number of Layers	Layer Type	Output Shape	Parameters
1	EfficientNetB7	(None, 5, 5, 2560)	64097687
2	Conv2D	(None, 5, 5, 64)	8028224
3	Conv2D	(None, 5, 5, 32)	100384
4	Max Pooling	(None, 3, 3, 32)	0
5	Flatten	(None, 288)	0

2.3. Classification

In the step-3 of the proposed method, features are fed into a deep neural network. Table 2 details out the layer type, output of each layer, and parameters used in the network. The architecture starts with a Batch Normalization layer, which stabilizes the network by normalizing input data and thus reducing the covariate shift, helping in accelerating the training process. The network contains two Dense layers. The first one contains 1024 neurons, and the second one contains 512 neurons. Each employee uses ReLU activation, which helps in reducing overfitting. After every dense layer, Batch Normalization is used along with Dropout layers, which randomly deactivate neurons during the training phase. The dropout layers force the network to develop more robust features rather than becoming strongly dependent on a limited set of neurons for prediction. The dropout rates follow a decreasing pattern (0.5, 0.4) as the architecture narrows, which enhances the model's robustness and its capability to handle unseen data effectively.

Table 2.
The number of layers, layer type, output shape and parameters of the Deep neural network.

Number of Layers	Layer Type	Output Shape	Parameters
1	Dense	(None, 1024)	295936
2	Batch Normalization	(None, 1024)	4096
3	Dropout	(None, 1024)	0
4	Dense	(None, 512)	524800
5	Batch Normalization	(None, 512)	2048
6	Dropout	(None, 512)	0
7	Dense	(None, 2)	1026

2.4. Deep Neural Networks

The study compares the proposed approach against other deep neural networks (VGG16, Xception, ResNet152, InceptionV3, EfficientNetB3). The details of each compared network are mentioned below.

VGG16 [29] is a deep neural network model used for image classification. The model was developed by the Visual Geometry Group at Oxford University. The model is composed of 16 layers of artificial neurons, including 13 convolutional layers and three fully connected layers. The model uses small 3×3 filters to improve performance over larger filters. VGG16 achieves a top-5 test accuracy of 92.7% on the ImageNet dataset. It is widely used for transfer learning due to its pre-trained model availability.

Xception [30] model is another major step in CNN evolution. Building upon the architecture of Inception Architecture [20] the model consists of 36 layers which are organized into three different flows: entry, middle, and exit flows. Each flow serves a specific purpose in feature extraction. The entry flow prepares the model for complex patterns, the middle flow predefines the understanding of the data, and the exit flow combines the processed data, which is further used for classification. It decouples spatial and cross-channel correlation more effectively than Inception, which helps in improving efficiency as well as performance. It uses residual connections to prevent vanishing gradients, overfitting, and thus improves training stability. Pretrained on ImageNet, Xception provides a strong base for transfer learning in all computer vision tasks.

ResNet152 [31] is a deep convolutional neural network with 152 layers. The strength lies in utilizing residual connections to mitigate the vanishing gradient problem. Trained on the ImageNet dataset, ResNet152 outperforms ResNet50 in terms of accuracy but requires more computational resources and longer training times. The model features a bottleneck architecture with multiple 3x3 convolutional layers and employs ReLU activations.

In 2015, Google introduced InceptionV3 [32] deep convolutional neural network. The network has better accuracy compared to its predecessors, InceptionV1 and InceptionV2. The architecture has several innovative features to enhance computational efficiency and classification accuracy. The InceptionV3 network has multiple inception modules that process input through parallel paths, and each module includes several convolutional layers and pooling operations, which help the network capture a wide range of features. The network also employs factorized convolutions, which involve breaking down larger convolutions into smaller ones, along with using asymmetric convolutions such as replacing a 7x7 filter with a combination of 1x7 and 7x1 filters, thus helping reduce parameters. The model also uses regularizers to improve training stability and prevent overfitting, along with batch normalization to stabilize training and label smoothing to prevent overconfidence in predictions.

EfficientNetB5 [28] is another model from the EfficientNet family of convolutional neural networks. The family is known for its efficient scaling method that balances depth, width, and resolution. The architecture of EfficientNetB5 uses Mobile Inverted Bottleneck (MBConv) layers and incorporates Squeeze-and-Excitation (SE) optimization. Like other EfficientNet models, it follows the compound scaling principle, where network dimensions (depth, width, and resolution) are scaled up uniformly using a compound coefficient. This balanced scaling approach helps maintain optimal performance while managing computational costs. The MBConv layers first expand the number of channels, perform efficient depthwise convolutions, and then compress the features back, reducing computational complexity while preserving important information. The Squeeze-and-Excitation blocks enhance the model's performance by dynamically recalibrating channel-wise feature responses. This allows the network to emphasize informative features and suppress less useful ones. EfficientNetB5 represents a middle-range variant in the EfficientNet family, offering a good balance between model size and accuracy. When trained on the ImageNet dataset, EfficientNetB5 achieves impressive results with a top-1 accuracy of 83.6% and top-5 accuracy of 96.7%. The model's architecture and scaling principles have made it popular for transfer learning tasks, where it can be fine-tuned for specific image classification applications while maintaining its efficient performance characteristics.

2.5. Evaluation Metrics

The performance of the proposed technique is evaluated using standard classification metrics, including Precision, Recall, Accuracy, F1-Score, and the Confusion matrix. In classification tasks involving images, the terms TP (True Positive), TN (True Negative), FP (False Positive), and FN (False Negative) are used to assess the classifier's performance. The terms "True" and "False" indicate whether the classifier's prediction aligns with the actual classification, while "Positive" and "Negative" refer to the classifier's prediction. The calculation methods for these metrics are detailed in formulas (1) through (4).

Precision is the ratio of the correctly classified actual positives to everything classified as positive.

$$precision = \frac{TP}{TP+FP} \qquad (1)$$

Recall is the proportion of all actual positives that were classified correctly as positives.

$$Recall = \frac{TP}{TP + FN} \qquad (2)$$

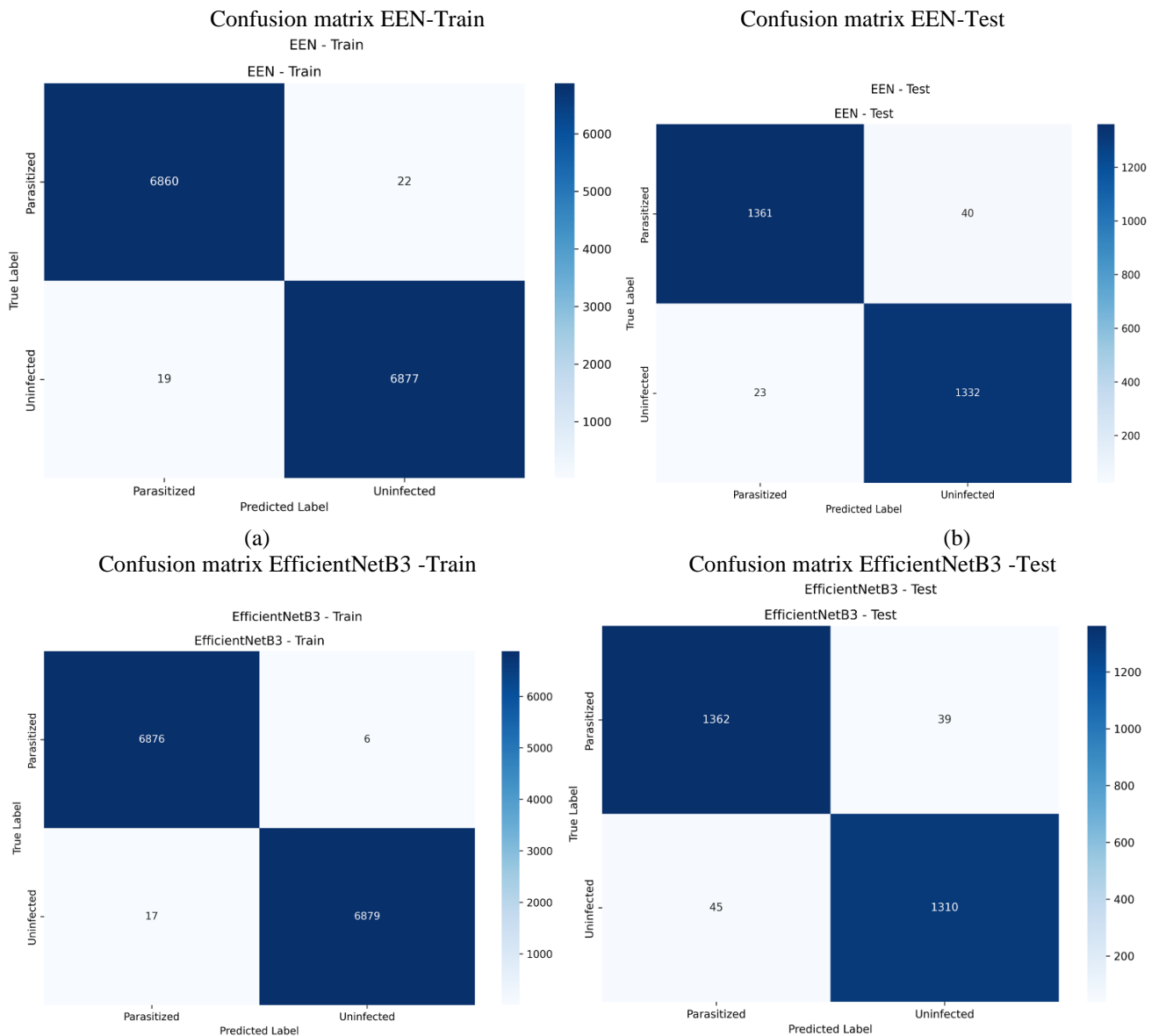Accuracy is the proportion of all classifications that were correct, whether positive or negative.

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

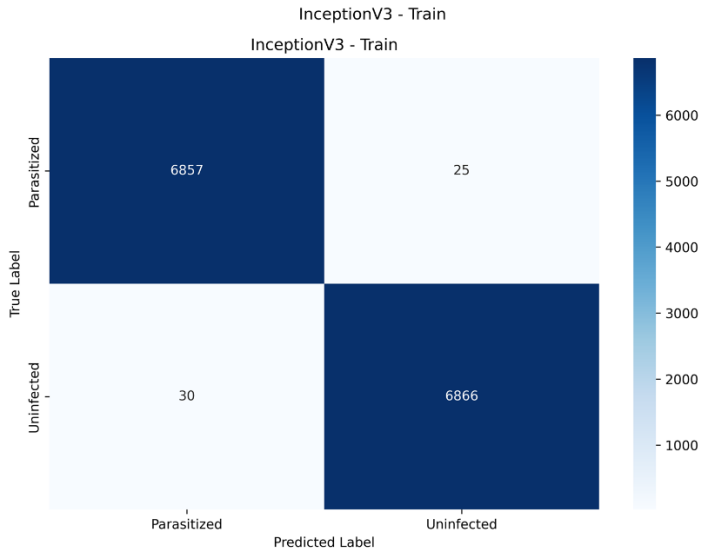'F1 Score' or 'F-measure' is a measure that combines precision and recall is the harmonic mean of precision and recall.

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \qquad (4)$$

3. Results and Discussion

Our novel approach is as shown in Figure 3 better predictions on the unseen test data. We can see the test metric for the EEN approach has correctly predicted 1361 instances of the Parasitized class and 1332 instances of the uninfected class. However, it predicted 23 false positives and 40 false negatives. On the training data, it performs better than on the test data, which is expected; it correctly predicted 6860 instances of the Parasitized class and 6877 instances of the uninfected class. However, it predicted 19 false positives and 22 false negatives.

Confusion matrix EEN-Train

Confusion matrix EEN-Test

(a)

(b)

Confusion matrix EfficientNetB3 -Train

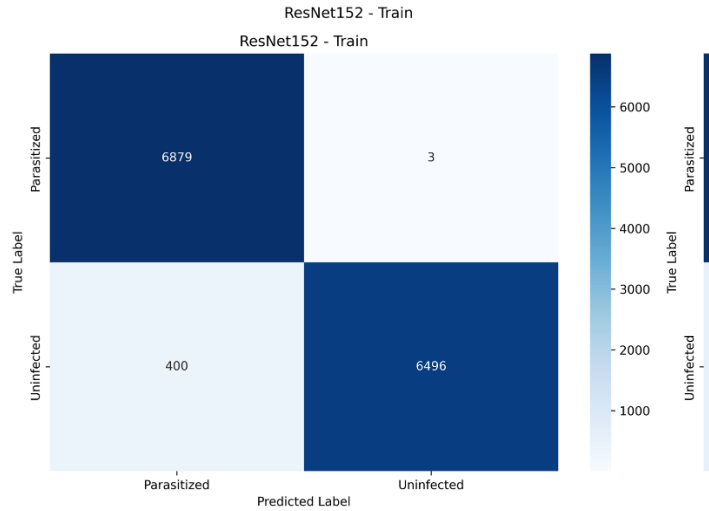
Confusion matrix EfficientNetB3 -Test
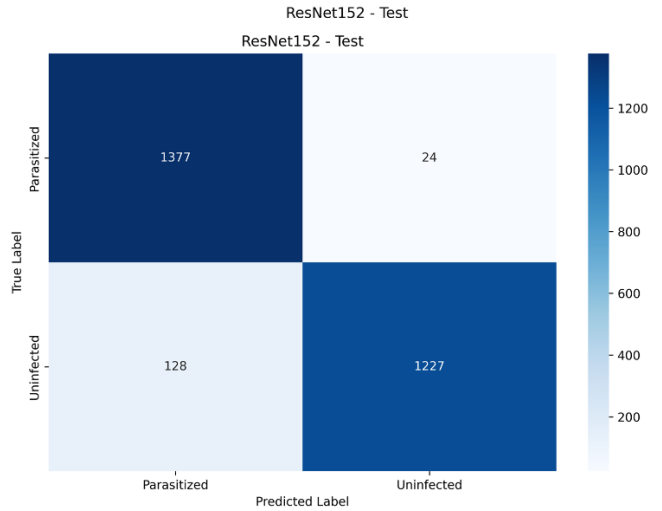
(c)
Confusion matrix InceptionV3 -Train

(d)
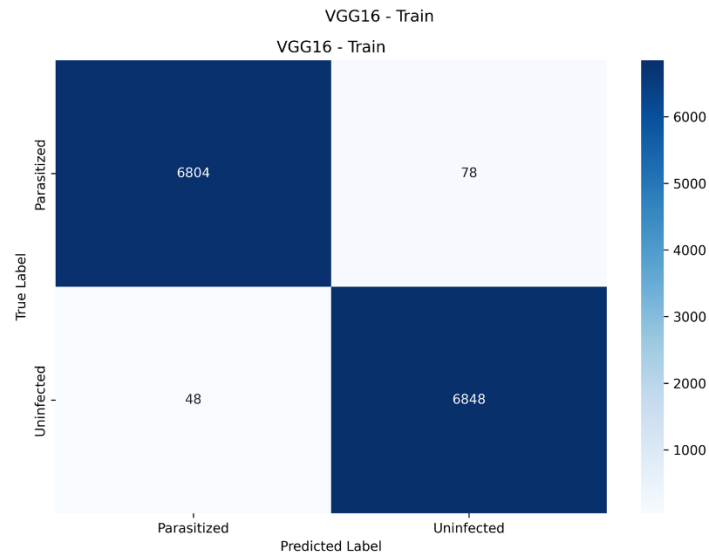Confusion matrix InceptionV3 -Test
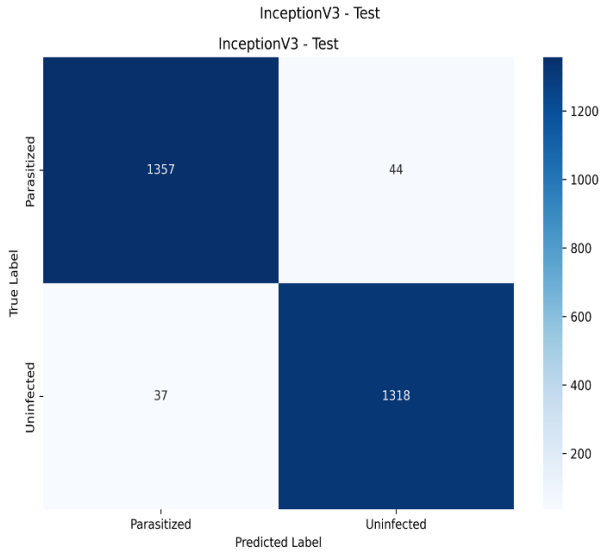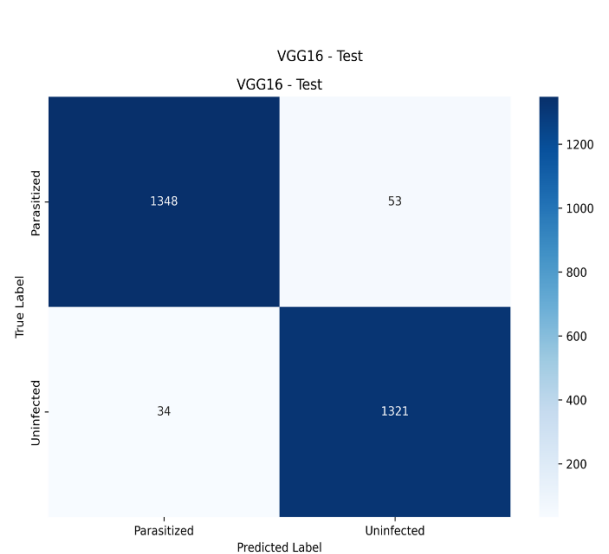

(e)
Confusion matrix Resnet152 -Train

(f)
Confusion matrix Resnet152 -Test

(g)
Confusion matrix VGG16 -Train

(h)
Confusion matrix VGG16 -Test

(i)
Confusion matrix Xception -Train
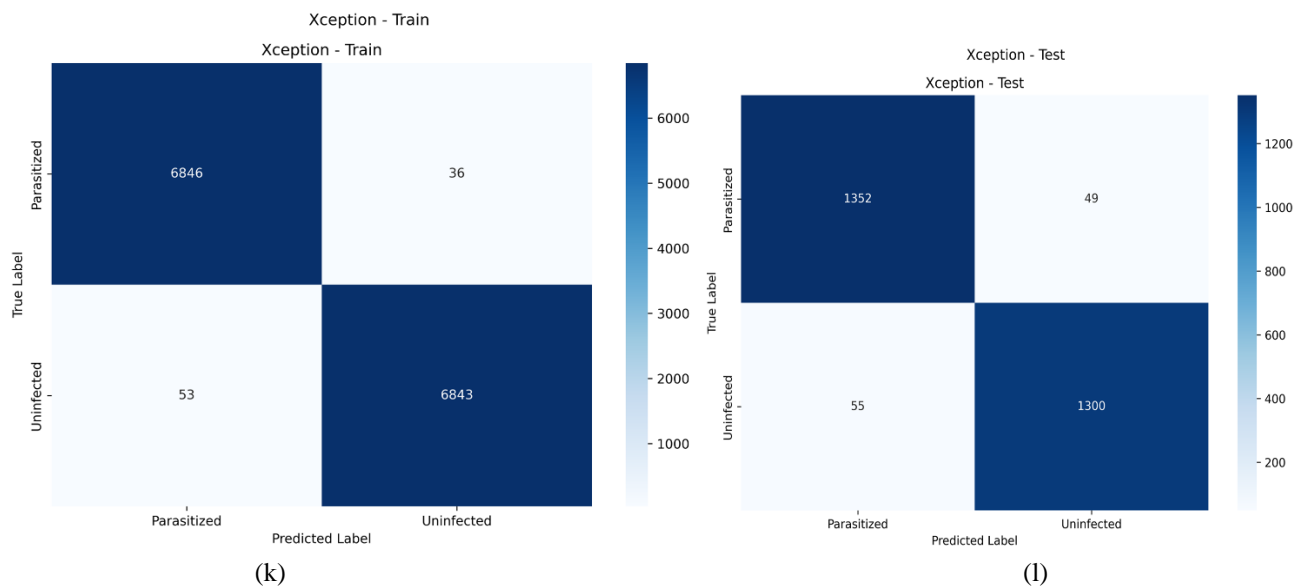
(j)
Confusion matrix Xception -Test

(k)　　　　　　　　　　　　　　　　　　(l)

Figure 3.
(a) Confusion matrix EEN-Train; (b) Confusion matrix EEN-Test, (c) Confusion matrix Efficient. NetB3 -Train; (d) Confusion matrix EfficientNetB3 -Test,(e) Confusion matrix InceptionV3 -Train; (f) Confusion matrix InceptionV3 -Test;(g) Confusion matrix Resnet152 -Train; (h) Confusion matrix Resnet152 -Test; (i) Confusion matrix VGG16 -Train; (j) Confusion matrix VGG16 -Test;(k) Confusion matrix Xception -Train; (l) Confusion matrix Xception -Test

In Figure 4, shows the EEN model; the first image presents the train metric scores by class for an EEN model. Here, Precision is 99.72% for the Parasitized class and 99.68% for the Uninfected class, while Recall is 99.68% for the Parasitized class and 99.72% for the Uninfected class. The F1 score is 99.70% for both classes, and Accuracy is 99.68% for the Parasitized class and 99.72% for the Uninfected class. The second image presents the test metric scores by class for the same model, highlighting Precision, Recall, F1, and Accuracy. Precision is 98.34% for the Parasitized class and 97.08% for the Uninfected class, while Recall is 97.14% for the Parasitized class and 98.30% for the Uninfected class. The F1 score stands at 97.74% for the Parasitized class and 97.69% for the Uninfected class, with Accuracy recorded at 97.14% for the Parasitized class and 98.30% for the Uninfected class. Overall, the model exhibits high performance across both training and testing datasets, with near-perfect scores in most metrics. The model remains highly accurate and precise in its predictions.

Precision, Recall, F1 Score and Accuracy: EEN- Train

(a)

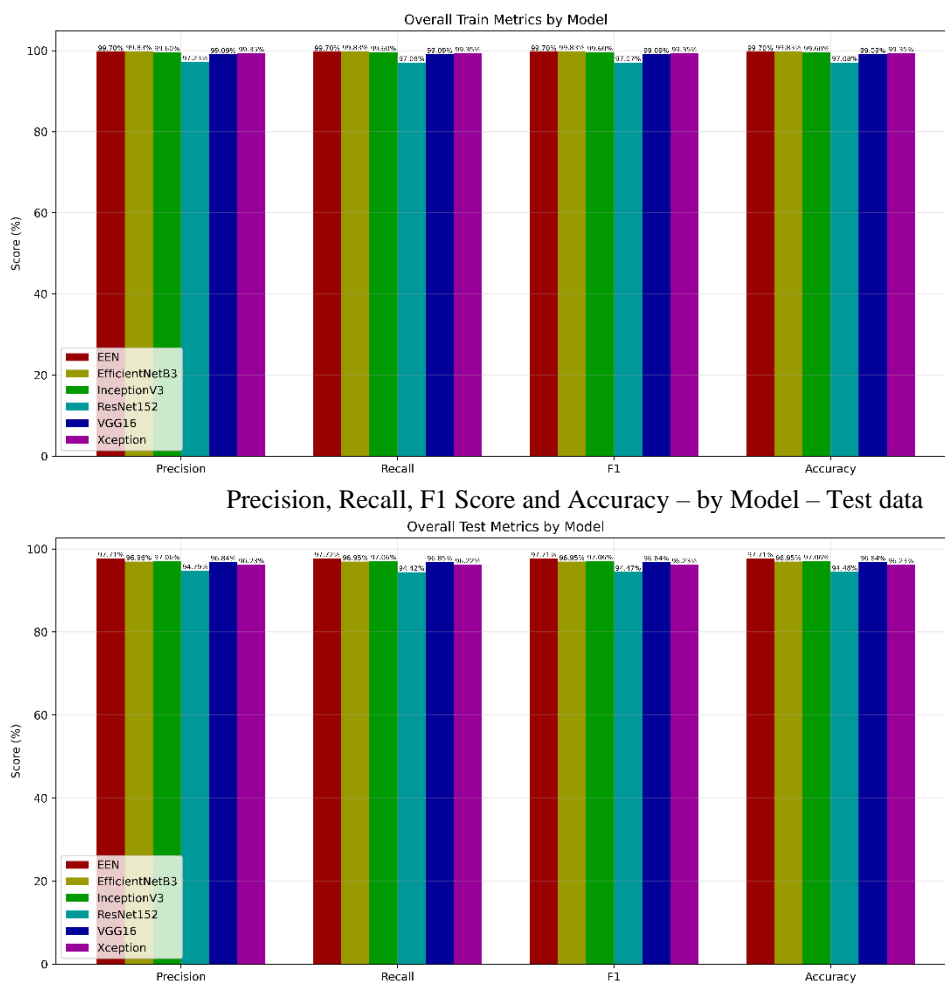Precision, Recall, F1 Score and Accuracy: EEN- Test
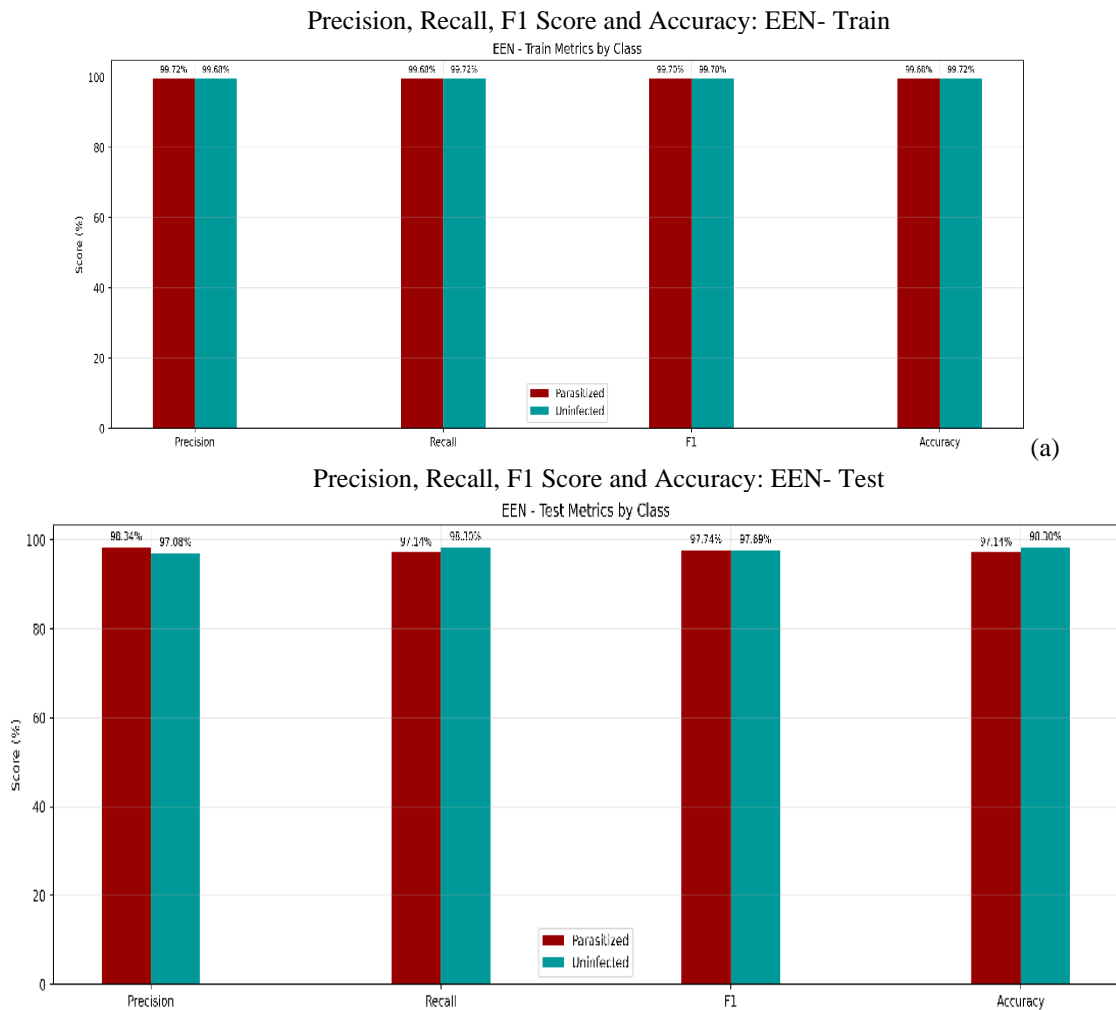
Figure 4.
(a) Precision, Recall, F1 Score and Accuracy – EEN- Train; (b) Precision, Recall, F1 Score and Accuracy – EEN- Test.

In our comprehensive evaluation of various deep learning architectures, shown in Figure 5, we observed distinct performance patterns across both training and testing phases. During the training phase, all models demonstrated exceptional performance with metrics consistently above 97%. The EEN architecture exhibited the highest training performance across all metrics, achieving 99.70% in precision, recall, F1-score, and accuracy. This was closely followed by EfficientNetB3, which maintained a robust 99.83% across all metrics. InceptionV3 showed training performance at 99.60%, while ResNet152 achieved slightly lower but still impressive results at 97.23%. Both VGG16 and Xception maintained consistent performance at 99.09% and 99.35%, respectively, across all training metrics.

When examining the test results, we observed a slight but expected degradation in performance across all architectures. The EEN architecture maintained its superior performance, achieving 97.71% across precision, recall, F1-score, and accuracy. EfficientNetB3 demonstrated robust generalization with test metrics at approximately 96.95%. InceptionV3 performed admirably with consistent test metrics around 97.06%. ResNet152 showed the most significant drop in performance during testing, with metrics averaging around 94.54%. VGG16 and Xception maintained relatively stable performance in testing, with metrics hovering around 96.84% and 96.23%, respectively.

EEN demonstrated the most stable performance transition from training to testing, with only about a 2-percentage point decrease, indicating superior generalization capabilities.

Precision, Recall, F1 Score and Accuracy – by Model - Train data

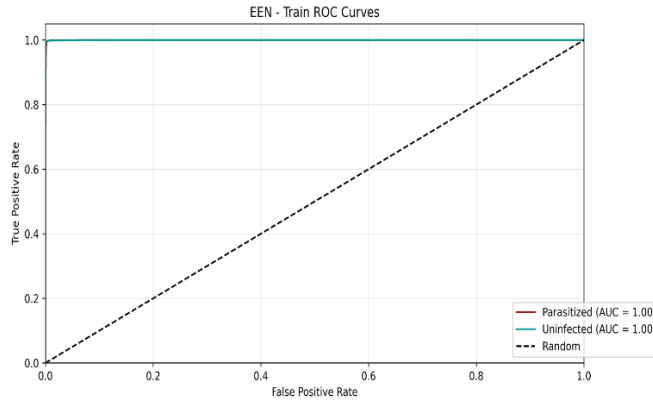Precision, Recall, F1 Score and Accuracy – by Model – Test data

Figure 5.
(a) Precision, Recall, F1 Score and Accuracy – Train; (b) Precision, Recall, F1 Score and Accuracy – Test.

In Figure 6, the ROC curves for the EEN (Enhanced Efficient Network) model show exceptional performance during both training and testing phases in distinguishing between parasitized and uninfected samples. During the training phase, the model demonstrated perfect classification capabilities, achieving an Area Under the Curve (AUC) of 1.00 for both parasitized and uninfected classes. The training ROC curves for both classes completely overlapped and maintained maximum separation from the random classifier baseline (represented by the dashed diagonal line), indicating optimal learning of the discriminative features for both conditions.

When evaluating the model's performance on the test set, we observed only a minimal decrease in performance, with the AUC slightly reducing to 0.99 for both classes. The test ROC curves exhibited remarkably similar characteristics to the training curves, with near-perfect separation from the random classifier baseline. The curves demonstrate a steep rise to high true positive rates at very low false positive rates, indicating the model's strong ability to correctly classify both parasitized and uninfected samples with high confidence. The nearly identical performance between the two classes suggests that the model maintains balanced classification capabilities without bias toward either category.
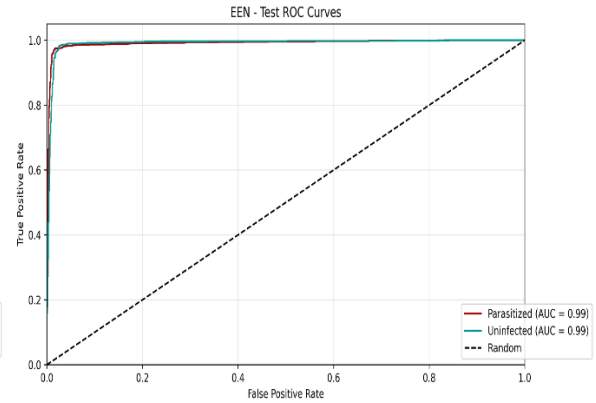
The marginal difference between training (AUC = 1.00) and testing (AUC = 0.99) performance indicates excellent generalization capabilities of the EEN model, with minimal overfitting. This is particularly noteworthy as it suggests the model has learned robust and generalizable features for distinguishing between parasitized and uninfected samples, rather than merely memorizing the training data. The consistent performance across both classes in both training and testing phases further validates the model's reliability for practical deployment in screening applications.
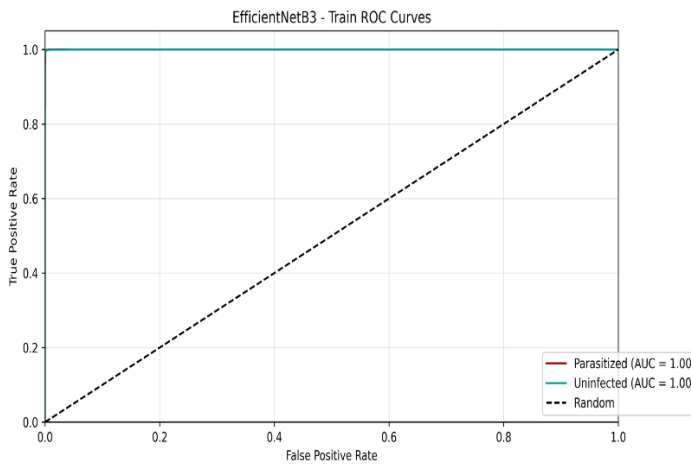
AUC-ROC EEN-Train

EEN - Train ROC Curves
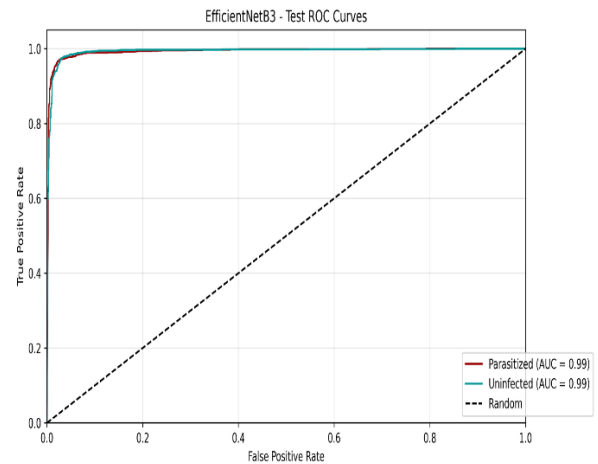
AUC-ROC EEN-Test

EEN - Test ROC Curves

(a)

(b)

AUC-ROC EfficientNetB3 -Train

EfficientNetB3 - Train ROC Curves

AUC-ROC EfficientNetB3 -Test

EfficientNetB3 - Test ROC Curves

(c)

(d)

AUC-ROC InceptionV3 -Train

InceptionV3 - Train ROC Curves

AUC-ROC InceptionV3 -Test

InceptionV3 - Test ROC Curves

(e)

(f)

AUC-ROC Resnet152 -Train

ResNet152 - Train ROC Curves

AUC-ROC Resnet152 -Test

ResNet152 - Test ROC Curves

(g)

(h)

AUC-ROC VGG16 -Train

AUC-ROC VGG16 -Test

Figure 6.
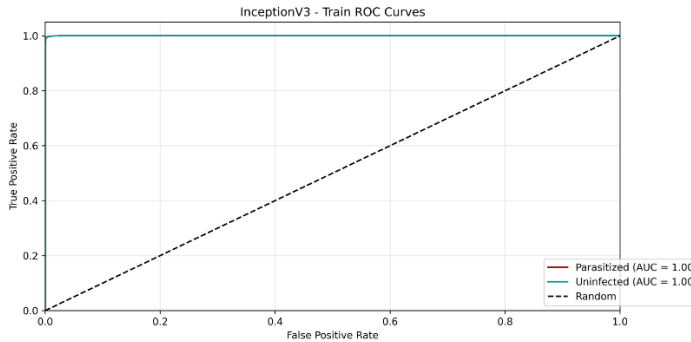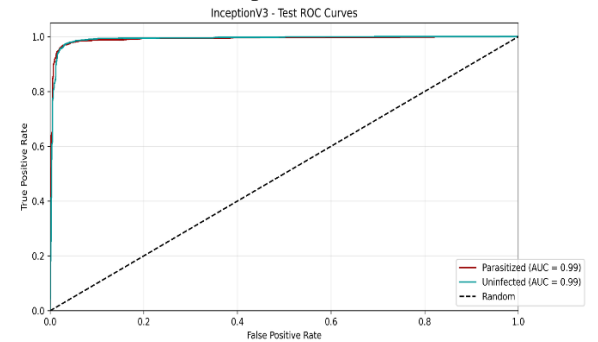(a) AUC-ROC EEN-Train; (b) AUC-ROC EEN-Test; (c) AUC-ROC EfficientNetB3 –Train; (d) AUC-ROC EfficientNetB3 –Test; (e) AUC-ROC InceptionV3 –Train; (f)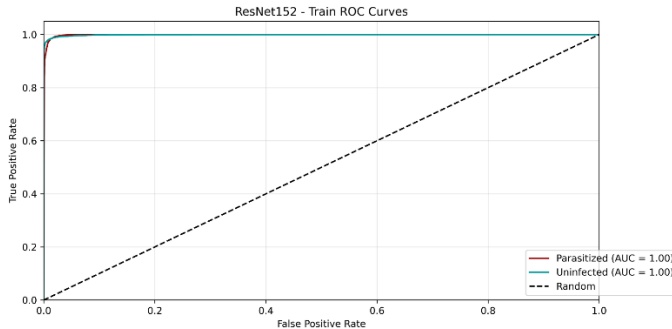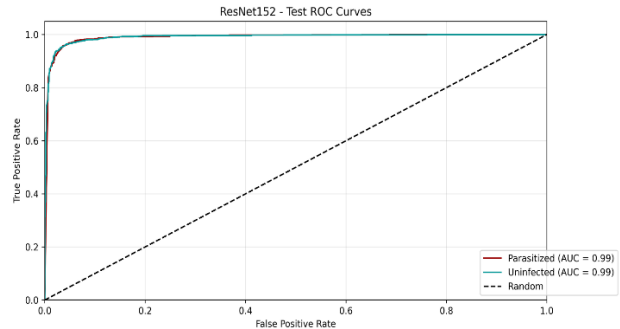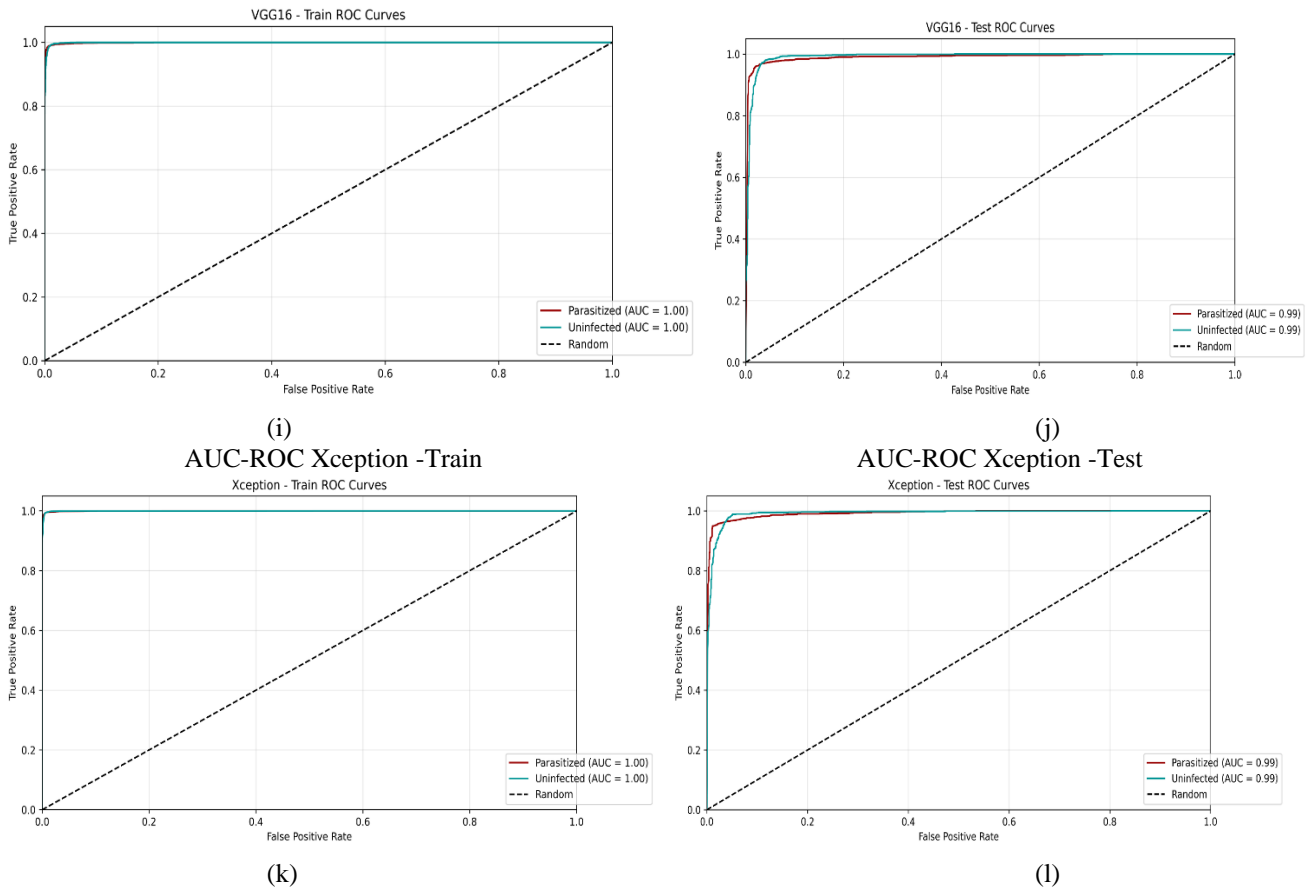 AUC-ROC InceptionV3 –Test; (g) AUC-ROC Resnet152 -Train; (h) AUC-ROC Resnet152 –Test; (i) AUC-ROC VGG16 –Train; (j) AUC-ROC VGG16 –Test; (k) AUC-ROC Xception –Train; (l) AUC-ROC Xception -Test.

In Figure 7, we present the accuracy and loss curves for the EEN (Enhanced Efficient Network) model over the course of training and evaluation. The left subfigure illustrates the accuracy trends for both the training and test datasets, while the right subfigure displays the corresponding loss curves.

The accuracy plot reveals a rapid increase in training accuracy during the initial epochs, reaching near-perfect classification performance beyond epoch 10. The test accuracy follows a similar trajectory but exhibits a relatively smaller increase, stabilizing around 97.7%. Notably, the best-performing epoch, marked by a red dot, aligns with a point where test accuracy reaches its peak, highlighting the model's optimal generalization capability.

The loss curves provide additional insights into the model's learning dynamics. The training loss experiences a sharp decline within the first few epochs, converging to near-zero values as training progresses. The test loss follows a similar decreasing trend but remains slightly higher than the training loss, suggesting a minor generalization gap. The best epoch, indicated in red, aligns with a point where test loss is at its lowest, reinforcing the selection of this epoch for the final model.

Overall, these results suggest that the EEN model has effectively learned the underlying patterns in the data while maintaining a strong generalization capability. The minimal difference between training and test accuracy, coupled with the stabilization of test loss, indicates that the model avoids significant overfitting, making it suitable for deployment in real-world applications.

EEN-Accuracy and Loss Over Epochs
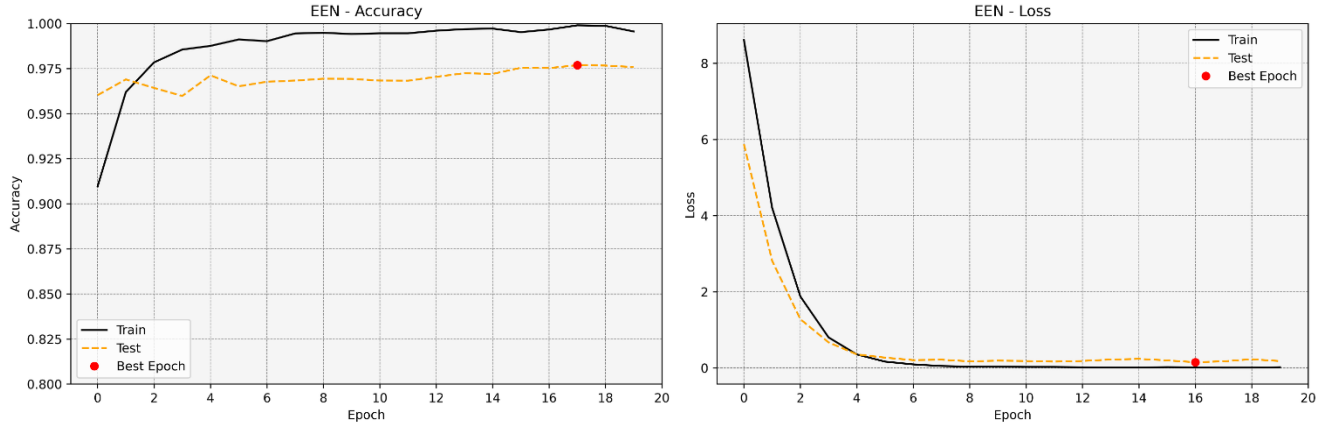
(a)

EfficientNetB3 -Accuracy and Loss Over Epochs
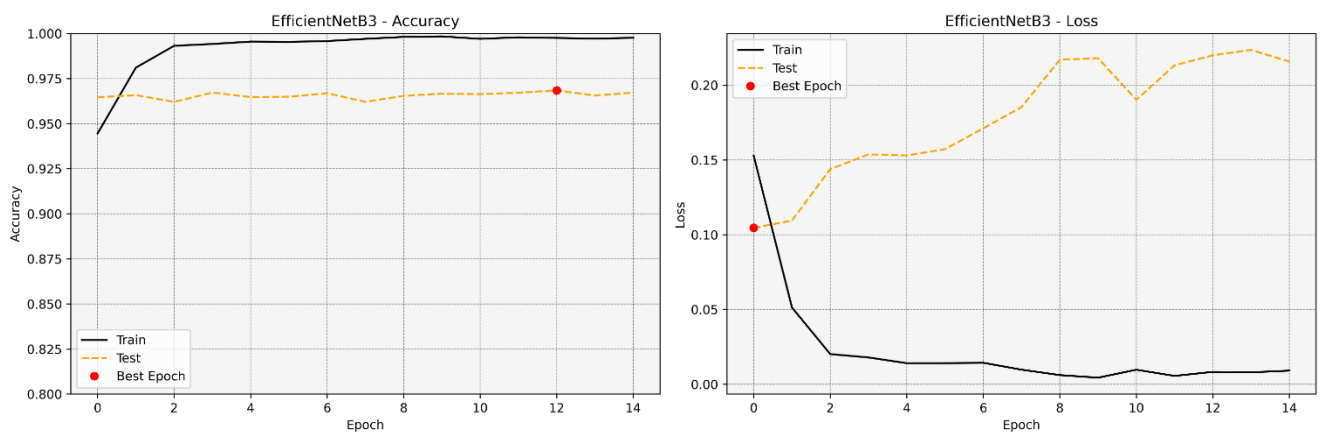
(b)

InceptionV3 -Accuracy and Loss Over Epochs

(c)

Resnet152 -Accuracy and Loss Over Epochs

(d)

VGG16 -Accuracy and Loss Over Epochs

(e)

Xception -Accuracy and Loss Over Epochs

(f)

Figure 7.
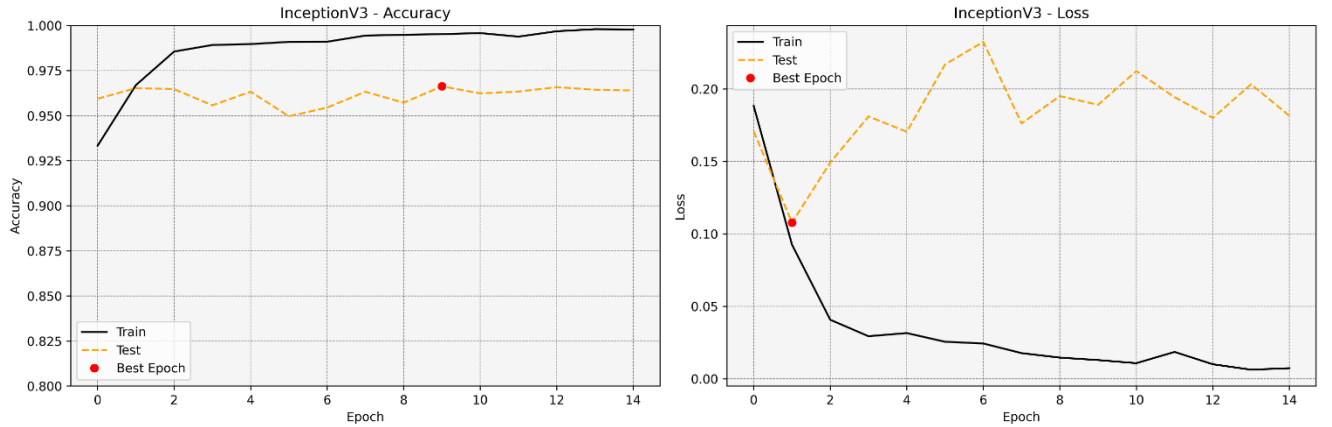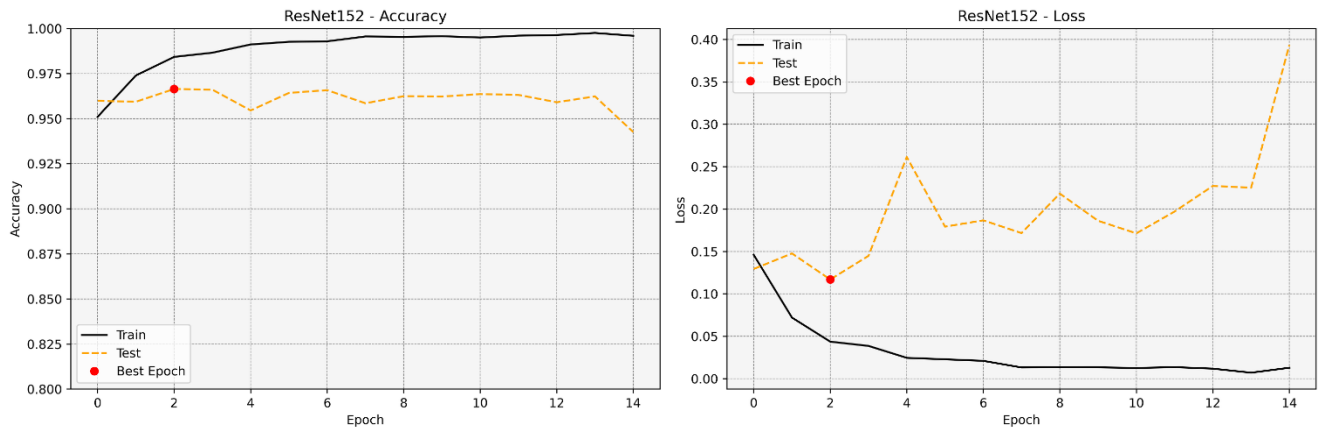(a) EEN-Accuracy and Loss Over Epochs; (b) EfficientNetB3 -Accuracy and Loss Over Epochs; (c) InceptionV3 -Accuracy and Loss Over Epochs; (d) Resnet152 -Accuracy and Loss Over Epochs; (e) VGG16 -Accuracy and Loss Over Epochs; (f) Xception -Accuracy and Loss Over Epochs.
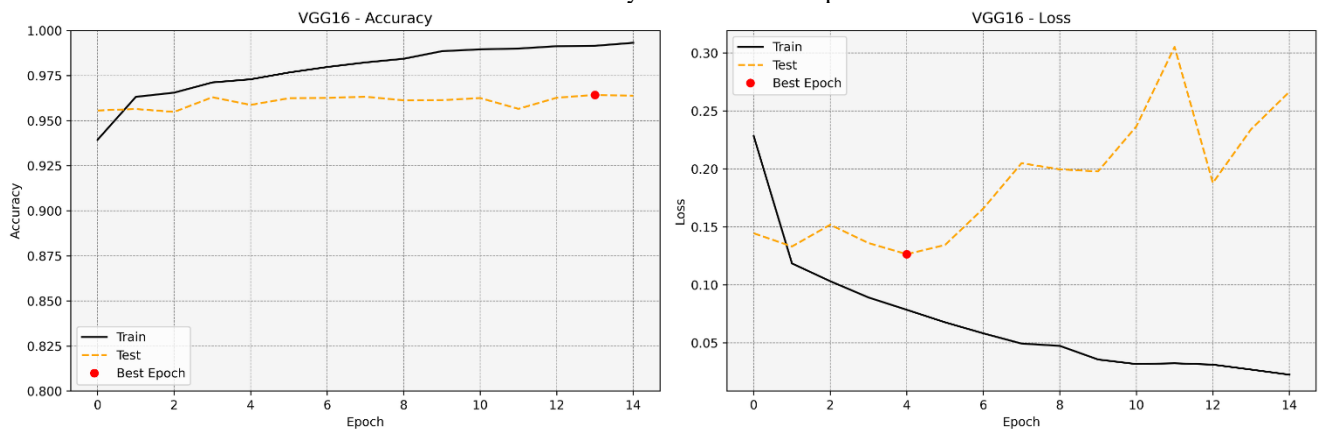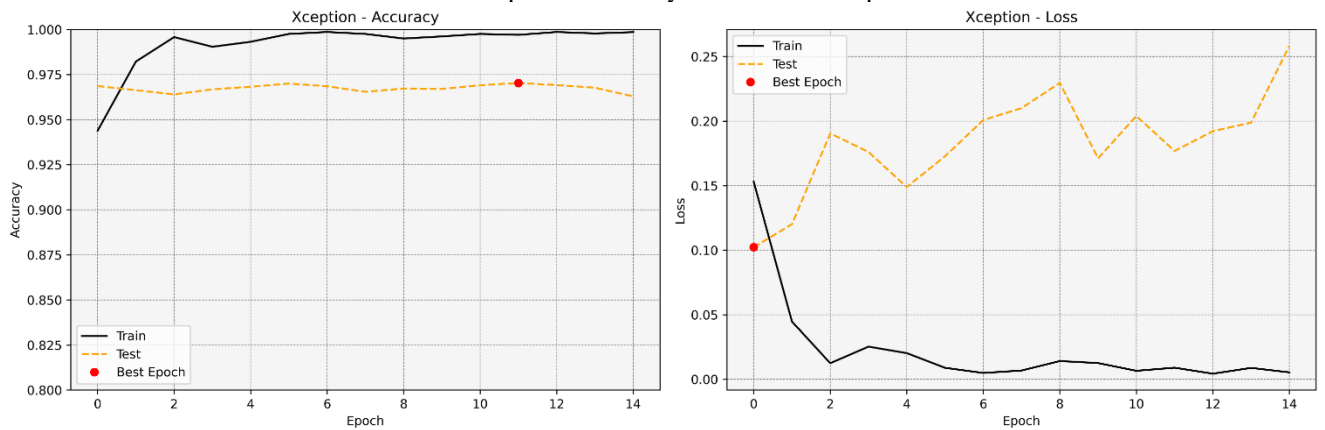
4. Discussion

Table 3.
Evaluating the accuracy of the proposed model against other recent research.

Reference	Method	Dataset	Accuracy
Vijayalakshmi [33]	VGG-19 + SVM	NIH Dataset	93%
Liang et al. [34]	CNN	NIH Dataset	96.54%
Dong et al. [35]	CNN	NIH Dataset	95.28%
Rajaraman et al. [36]	Ensemble of pre-trained CNNs	NIH Dataset	95.9%
Zhou et al. [37]	FCNN	ALL-IDB1	85%
Khandekar et al. [38]	Thresholding	ALL-IDB1 and CNMC 2019	95%
Chola et al. [39]	BCNet	HPBC	98.51%
Dese et al. [40]	Deep learning methods	Private dataset	95%
Ansari et al. [41]	Type 2 fuzzy + CNN	Private dataset	98%
Abhishek et al. [42]	VGG 16	Private dataset	85%
Awais et al. [43]	CNNs	Private dataset	99.15%
EEN (Proposed approach)	EfficientNet + Deep neural network	NIH Dataset	97.71%

Several studies have leveraged the NIH Dataset and other datasets, as shown in Table 3, a widely used dataset for medical image classification. Vijayalakshmi [33] employed a combination of VGG-19 and SVM, achieving an accuracy of 93%, while Liang et al. [34] implemented a CNN-based approach that yielded a higher accuracy of 96.54%. Similarly, Dong et al. [35] and Rajaraman et al. [36] reported accuracies of 95.28% and 95.9%, respectively, using CNN-based and ensemble CNN techniques.

For the ALL-IDB1 dataset, Zhou et al. [37] utilized a fully convolutional neural network (FCNN) and achieved an accuracy of 85%, while Khandekar et al. [38] employed thresholding techniques across both the ALL-IDB1 and CNMC 2019 datasets, attaining 95% accuracy.

Other studies utilized private or specialized datasets. Chola et al. [39] introduced BCNet, achieving an impressive 98.51% accuracy on the HPBC dataset. Dese et al. [40], Awais et al. [43] and Abhishek et al. [42] employed deep learning-based approaches, obtaining accuracies of 95% and 99.15%, respectively, on private datasets. Additionally, Ansari et al. [41] combined Type 2 fuzzy logic with CNNs, achieving a classification accuracy of 98%, while Abhishek et al. [42] utilized VGG-16 on a private dataset, yielding 85% accuracy. However, the highest accuracy was achieved by Awais et al. [43] using CNN on a private dataset.

Our proposed model is performing better compared to other studies, as shown in Table 1. Our model shows an accuracy of 97.17% on the test dataset. It will be difficult to compare the results of our proposed models with those of other studies, as some of the other studies are conducted on different datasets. We have also compared our proposed model with other state-of-the-art models such as EfficientNetB3, InceptionV3, ResNet152, VGG16, and Xception. Our proposed model performs better than these models, as shown in Table 4.

Table 4.
Evaluating the accuracy of proposed model against other state of the art models.

Model	Precision (%)	Recall (%)	F1 Score (%)	Accuracy (%)
EEN (Proposed Model)	97.71	97.72	97.71	97.71
EfficientNetB3	96.96	96.95	96.95	96.95
InceptionV3	97.06	97.06	97.06	97.06
ResNet152	94.79	94.42	94.47	94.48
VGG16	96.84	96.85	96.84	96.84
Xception	96.23	96.22	96.23	96.23

Abbreviations
The following abbreviations are used in this manuscript:
EEN   Enhanced EfficientNet

References
[1]   P. Venkatesan, "The 2023 WHO World malaria report," *The Lancet Microbe,* vol. 5, no. 3, p. e214, 2024. https://doi.org/10.1016/s2666-5247(24)00016-8
[2]   T. L. I. Diseases, "A new dawn for malaria prevention," *The Lancet Infectious Diseases,* 2024. https://doi.org/10.1016/s1473-3099(24)00012-4
[3]   M. Thellier, A. A. J. Gemegah, and I. Tantaoui, "Global Fight against Malaria: Goals and Achievements 1900–2022," *Journal of Clinical Medicine,* vol. 13, no. 19, p. 5680, 2024. https://doi.org/10.3390/jcm13195680
[4]   B. Amoah *et al.*, "Global trends in the burden of malaria: Contemporary diagnostic approaches, and treatment strategies," *World Journal of Advanced Research and Reviews,* 2023. https://doi.org/10.30574/wjarr.2023.20.1.2038

[5]   V. Balakrishnan, S. Padmanabhan, and P. Ravi, "Advancements in malaria diagnosis and treatment: A comprehensive review of current strategies and emerging innovations," 2024. https://doi.org/10.58532/v3bfms9p1ch6

[6]   Malaria Diagnostics, "Infectious diseases," 2023. https://doi.org/10.5772/intechopen.106631

[7]   V. Chauhan, R. Anand, and A. P. Singh, "Advancements and challenges in malaria diagnostics," 2024. https://doi.org/10.20944/preprints202411.0937.v1

[8]   U. Beisel, R. Umlauf, E. Hutchinson, and C. I. Chandler, "The complexities of simple technologies: Re-imagining the role of rapid diagnostic tests in malaria control efforts," *Malaria Journal,* vol. 15, pp. 1-9, 2016. https://doi.org/10.1186/S12936-016-1083-2

[9]   H. P. Ndayishimiye, "Challenges in malaria control and prevention in East Africa," *Newport International Journal of Biological and Applied Sciences,* 2024. https://doi.org/10.59298/nijbas/2024/5.2.24271

[10]   O. A. Onile-ere, J. Openibo, and G. I. Olasehinde, "Malaria diagnosis: Current approaches and future prospects," in *Proceedings of the 10th International Malaria Conference, Lagos, Nigeria, July 2016*, 2016.

[11]   E. I. Obeagu *et al.*, "Revolution in malaria detection: unveiling current breakthroughs and tomorrow's possibilities in biomarker innovation," *Annals of Medicine and Surgery,* vol. 86, no. 10, pp. 5859-5876, 2024.

[12]   B. Simmons *et al.*, "Defining a malaria diagnostic pathway from innovation to adoption: Stakeholder perspectives on data and evidence gaps," *PLOS Global Public Health,* vol. 4, no. 5, p. e0002957, 2024. https://doi.org/10.1371/journal.pgph.0002957

[13]   E. I. Obeagu and G. U. Obeagu, "Emerging public health strategies in malaria control: Innovations and implications," *Annals of Medicine and Surgery,* vol. 86, no. 11, pp. 6576-6584, 2024. https://doi.org/10.1097/ms9.0000000000002578

[14]   A. Mehta, T. Singh, S. K. Sharma, and N. Garg, "A relational study on artificial intelligence methods for identification of Malaria parasites: A review," *Challenges in Information, Communication and Computing Technology,* pp. 415-420, 2025. https://doi.org/10.1201/9781003559092-71

[15]   E. Mukeshimana and J. Nzabanita, "Using machine learning techniques to predict malaria prevalence in Rwanda," *Scientific Reports,* vol. 14, no. 1, p. 12345, 2024. https://doi.org/10.21203/rs.3.rs-5318706/v1

[16]   R. Syabrina and G. Wang, "Predictive analysis of malaria cases in Indonesia using machine learning," *Syntax Literate: Jurnal Ilmiah Indonesia,* vol. 9, no. 9, p. 16714, 2024. https://doi.org/10.36418/syntax-literate.v9i9.16714

[17]   A. G. Taye, S. Yemane, E. Negash, Y. Minwuyelet, M. Abebe, and M. H. Asmare, "Automated web-based malaria detection system with machine learning and deep learning techniques," *arXiv preprint arXiv:2407.00120,* 2024. https://doi.org/10.48550/arxiv.2407.00120

[18]   M. J. Hoque, M. S. Islam, M. Khaliluzzaman, A. Al Muntasir, and M. A. Mohsin, "Revolutionizing malaria diagnosis: Deep learning-powered detection of parasite-infected red blood cells," *International Journal of Electrical & Computer Engineering (2088-8708),* vol. 14, no. 4, pp. 4518-4530, 2024.

[19]   L. Vedula, A. Pilly, and S. Doss, *Revolutionizing malaria prediction using digital twins and advanced gradient boosting techniques* (Exploring the Advancements and Future Directions of Digital Twins in Healthcare 6.0). IGI Global. https://doi.org/10.4018/979-8-3693-5893-1.ch013, 2024.

[20]   F. Grignaffini, P. Simeoni, A. Alisi, and F. Frezza, "Computer-aided diagnosis systems for automatic malaria parasite detection and classification: A systematic review," *Electronics,* vol. 13, no. 16, p. 3174, 2024. https://doi.org/10.3390/electronics13163174

[21]   K. Yahuza, A. M. Umar, B. u. Salisu, M. L. Gambo, and B. Abdulkadir, "Recent advancements in detection and quantification of malaria using artificial intelligence," *UMYU Journal of Microbiology Research,* vol. 9, no. 2, pp. 1-21, 2024. https://doi.org/10.47430/ujmr.2492.001

[22]   S. Yanik *et al.*, "Application of machine learning in a rodent malaria model for rapid, accurate, and consistent parasite counts," *The American Journal of Tropical Medicine and Hygiene,* vol. 111, no. 5, p. 967, 2024. https://doi.org/10.4269/ajtmh.24-0135

[23]   G. Verma, "Automated malaria detection using EfficientNetB3 in deep learning frameworks," presented at the 2024 8th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC), 2024.

[24]   S. Kumar and S. Hirthick, *Deep learning advancements in malaria diagnosis* (Artificial Intelligence Transformations for Healthcare Applications: Medical Diagnosis, Treatment, and Patient Care). IGI Global. https://doi.org/10.4018/979-8-3693-7462-7.ch010, 2024.

[25]   H. Irrigisetty, K. R. Madhavi, V. S. Prasad, H. Jonna, M. Kurlapalli, and H. R. Gangadasari, "Enhancing high-resolution malaria parasite detection in blood smears using deep learning," in *Proceedings of the International Conference on Emerging Technologies (INCET)*, 2024: IEEE, pp. 1-6.

[26]   S. Shashikiran and H. Sunitha, "Malaria cell detection using advanced SVM techniques," presented at the 2024 Second International Conference on Networks, Multimedia and Information Technology (NMITCON), 2024.

[27]   L. H. N. C. B. National Library of Medicine, *Malaria data sheet*. Bethesda, MD: National Institutes of Health, 2024.

[28]   M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," presented at the International Conference on Machine Learning, 2019.

[29]   K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556,* 2014. https://doi.org/10.48550/arXiv.1409.1556

[30]   F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251-1258.

[31]   K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770-778.

[32]   C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 2818-2826.

[33]   A. Vijayalakshmi, "Deep learning approach to detect malaria from microscopic images," *Multimedia Tools and Applications,* vol. 79, no. 21, pp. 15297-15317, 2020.

[34]   Z. Liang *et al.*, "CNN-based image analysis for malaria diagnosis," in *Proceedings of the 2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE*, 2016: IEEE, pp. 493-496.

[35]   Y. Dong *et al.*, "Evaluations of deep convolutional neural networks for automatic identification of malaria infected cells," presented at the 2017 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI), 2017.

[36]   S. Rajaraman *et al.*, "Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images," *PeerJ,* vol. 6, p. e4568, 2018. https://doi.org/10.7717/peerj.4568

[37]   M. Zhou *et al.*, "Development and evaluation of a leukemia diagnosis system using deep learning in real clinical scenarios," *Frontiers in Pediatrics,* vol. 9, p. 693676, 2021. https://doi.org/10.3389/fped.2021.693676

[38]   R. Khandekar, P. Shastry, S. Jaishankar, O. Faust, and N. Sampathila, "Automated blast cell detection for Acute Lymphoblastic Leukemia diagnosis," *Biomedical Signal Processing and Control,* vol. 68, p. 102690, 2021. https://doi.org/10.1016/j.bspc.2021.102690

[39]   C. Chola *et al.*, "Bcnet: A deep learning computer-aided diagnosis framework for human peripheral blood cell identification," *Diagnostics,* vol. 12, no. 11, p. 2815, 2022. https://doi.org/10.3390/diagnostics12112815

[40]   K. Dese *et al.*, "Accurate machine-learning-based classification of leukemia from blood smear images," *Clinical Lymphoma Myeloma and Leukemia,* vol. 21, no. 11, pp. e903-e914, 2021. https://doi.org/10.1016/j.clml.2021.06.025

[41]   S. Ansari, A. H. Navin, A. Babazadeh Sangar, J. Vaez Gharamaleki, and S. Danishvar, "Acute Leukemia diagnosis based on images of lymphocytes and monocytes using type-II fuzzy deep network," *Electronics,* vol. 12, no. 5, p. 1116, 2023. https://doi.org/10.3390/electronics12051116

[42]   A. Abhishek, R. K. Jha, R. Sinha, and K. Jha, "Automated detection and classification of leukemia on a subject-independent test dataset using deep transfer learning supported by Grad-CAM visualization," *Biomedical Signal Processing and Control,* vol. 83, p. 104722, 2023. https://doi.org/10.1016/j.bspc.2023.104722

[43]   M. Awais, R. Ahmad, N. Kausar, A. I. Alzahrani, N. Alalwan, and A. Masood, "ALL classification using neural ensemble and memetic deep feature optimization," *Frontiers in Artificial Intelligence,* vol. 7, p. 1351942, 2024. https://doi.org/10.3389/frai.2024.1351942