The First International Conference on
Machine Learning and Knowledge Discovery (MLKD 2024)
Amirkabir University of Technology, December 18-19, 2024

From Images to Insights: Advanced CNN Architectures for Accurate Malaria Cell Classification

Tahereh Majidi* Samira Qasemi† Saeedeh Kamjoo‡ Mehrdad Lakestani§

Abstract

This paper presents a comparative analysis of different approaches for classifying malaria-infected cells, including pre-trained models, mutual information techniques, and a custom-designed Convolutional Neural Network (CNN). Malaria, a life-threatening disease, necessitates precise and swift diagnosis. The study uses a publicly available dataset of malaria cell images, applying preprocessing and data augmentation to enhance the models' performance. The proposed CNN architecture was evaluated using 5-fold cross-validation and compared against DenseNet121 and mutual information techniques. The proposed model achieved the highest accuracy of 96.15% on the test dataset, outperforming the others. This work demonstrates that the custom CNN model provides a superior solution for automated malaria detection, especially in low-resource environments.

Keywords: 5fold cross-validation, Mutual information, and Pretrained models

1 Introduction

With over 400,000 deaths annually due to malaria, the need for scalable, accurate, and cost-effective diagnostic tools is more critical than ever. Recent advancements in AI offer transformative potential, yet their application in low-resource settings remains underexplored.

Accurate and rapid disease diagnosis is a cornerstone of effective medical management and treatment. Malaria, as a life-threatening disease, remains a significant challenge in many tropical and subtropical regions. Despite notable advancements in laboratory diagnostics, conventional methods such as microscopy often face limitations in resource-constrained environments due to the need for specialized equipment and trained personnel.

The emergence of artificial intelligence and deep learning has opened new horizons for automating diagnostic processes and reducing reliance on traditional techniques. Convolutional Neural Networks (CNNs), with their exceptional ability to analyze complex visual data, have become highly effective tools for identifying patterns and anomalies in medical imaging. These approaches have the potential to enhance diagnostic accuracy and speed, offering novel opportunities for improving efficiency in resource-limited settings.

This study focuses on designing and evaluating a deep learning-based model for the detection of malaria from medical images. Leveraging the hierarchical learning capabilities of CNNs, we aim to streamline the diagnostic process and reduce dependence on manual expertise. The primary goal of this research is to provide a practical and accurate solution for automated diagnostic systems that can be implemented effectively in real-world, resource-constrained environments.

While traditional methods and pre-trained models like DenseNet have achieved significant results in medical image classification, our study proposes a novel CNN architecture specifically optimized for malaria detection in resource-constrained environments. This architecture incorporates advanced feature extraction techniques and tailored preprocessing methods to address challenges like low image resolution and data imbalance, setting it apart from existing approaches.

2 Literature Review

Malaria is a life-threatening disease transmitted by female *Anopheles* mosquitoes. Symptoms include fever, vomiting, headaches, and fatigue, with severe cases leading to coma or death [**?**]. The disease is caused by protozoa from the genus *Plasmodium*, with *P. falciparum* being the most lethal [8]. Malaria is widespread in tropical and subtropical regions, particularly in Sub-Saharan Africa, Asia, and Latin America. Globally, malaria caused around 731,000 deaths in 2016, with 90% occurring in Africa.

Microscopy remains the gold standard for malaria diagnosis, involving the placement of a blood drop on a glass slide, followed by staining and examination for par-

*Department of Applied Mathematics, Faculty of Mathematical Sciences, University of Tabriz, Tabriz, Iran, t.majidi@tabrizu.ac.ir
†Department of Artificial Intelligence, East Golestan Higher Education Institute, Golestan, Iran, samira.qasemi67@gmail.com
‡Department of Applied Mathematics, Faculty of Mathematical Sciences, University of Tabriz, Tabriz, Iran, s.kamjoo@tabrizu.ac.ir
§Department of Applied Mathematics, Faculty of Mathematical Sciences, University of Tabriz, Tabriz, Iran, lakestani@tabrizu.ac.ir

asites [31]. This method, however, poses challenges, especially in resource-limited African countries where there is a scarcity of equipment, materials, and skilled personnel necessary for accurate diagnoses [**?**]. The overlap of malaria symptoms with other diseases can complicate treatment, potentially increasing the risk of antibiotic and drug resistance when treatments are based solely on symptoms [5**?**]. Leaving malaria untreated is critically dangerous and can be fatal.

Malaria diagnosis is often initiated based on clinical symptoms; however, the Centers for Disease Control and Prevention (CDC) underscores the necessity of laboratory confirmation [9]. Various laboratory techniques are utilized to confirm malaria infection, including polymerase chain reaction (PCR), which is pivotal for identifying the specific *Plasmodium* species responsible for malaria in confirmed cases [14]. Antigen detection assays serve as rapid diagnostic tests that detect *Plasmodium*-derived antigens, facilitating timely diagnosis [17, 28]. Serological testing, such as enzyme-linked immunosorbent assays (ELISA), is employed to identify antibodies against malaria parasites, offering another layer of confirmation [24].

Despite their effectiveness, these laboratory methods are often impractical in low-resource settings due to their high costs, the need for specialized equipment, and the requirement for trained personnel [9]. As a result, light microscopy of thin or thick blood smears stained with Giemsa is the predominant method for diagnosing malaria in these contexts, providing confirmation of *Plasmodium* presence [7].

The severity of malaria infection is frequently assessed by calculating the percentage of red blood cells infected with malaria parasites, commonly referred to as percent parasitemia or parasitemia burden. However, the diagnostic accuracy of Giemsa-stained blood smears is heavily reliant on the technician's skill, as manual classification and enumeration of infected cells are required. This dependence on manual processes can lead to substantial inter-observer variability, particularly in low-resource environments where technicians may face multiple responsibilities and receive insufficient training specific to malaria diagnostics [4, 6]. For instance, a study in Nigeria revealed significant concerns regarding the reliability of malaria test results, attributing discrepancies to technician incompetence [13]. Moreover, research conducted in primary health care facilities in Tanzania reported a sensitivity of 74.5% and specificity of 59.0% for microscopy-based malaria diagnoses, highlighting deficiencies in technician training [27]. Similar findings were reported in Angola, where inadequate training for technicians engaged in microscopy-based malaria diagnostics was identified as a critical issue [25].

To address these challenges, automated algorithms leveraging image processing, computer vision, and artificial intelligence are continuously evolving . These technologies can enhance diagnostic reliability and standardization, particularly in low-resource settings, and enable researchers to conduct evaluations swiftly without the need for costly laboratory equipment. The adoption of machine learning techniques, especially neural networks, is expanding rapidly across various clinical domains. Primarily, these methods find applications in segmentation and classification tasks involving clinical images [2, 22, 32] as well as histological specimens [16, 35]. Among these applications, the deployment of machine learning for malaria diagnosis has gained particular attention, leading to the development of numerous classification models aimed at distinguishing between infected and uninfected red blood cells.

In situations where these red blood cell images lack sufficient resolution, the Fast Super-Resolution CNN (FSRCNN) model has been applied to upscale low-resolution images from 32x32 pixels to 128x128 pixels [12]. This enhancement is particularly useful when low-end cameras compromise image quality during the acquisition of thin blood smear images. Subsequently, a variant of the VGG16 CNN has been utilized to classify each red blood cell as either infected or uninfected. This sequential approach establishes an efficient mechanism whereby our screening platform processes thin blood smear images, providing healthcare practitioners with quantifiable data on the number of infected red blood cells and the associated parasitemia burden within a sample.

This study focuses on the design and development of a deep learning model tailored for the detection of malaria in medical images. We utilize Convolutional Neural Networks (CNNs), which have shown exceptional performance in image classification tasks across various domains. The inherent capability of CNNs to automatically learn hierarchical features renders them particularly effective for analyzing complex visual data, making them well-suited for identifying patterns and anomalies in medical imaging. Our selection of CNNs stems from their robustness in processing high-dimensional data and their proficiency in learning from extensive datasets—qualities that are vital given the variability inherent in medical images. By training our model on a diverse dataset comprising both malaria-infected and non-infected samples, we aim to develop a system that not only improves diagnostic accuracy but also streamlines the detection workflow.

Malaria detection from medical images presents challenges due to the complexity of the disease and reliance on manual methods. Malaria, caused by *Plasmodium* parasites, remains a major global health issue. Accurate and rapid diagnosis is critical for effective disease management, but traditional methods, like manually examining blood smears, are time-consuming and prone to

error.

Overcoming the challenges of manual malaria detection requires more efficient, automated solutions. Deep learning techniques, especially Convolutional Neural Networks (CNNs) [19], have shown promising results in detecting malaria from medical images. These models can reduce diagnostic errors and accelerate the process, though challenges remain in improving accuracy and adapting to diverse imaging conditions. This section of the literature review addresses research conducted on the identification and classification of malaria parasites using image processing techniques and deep learning methods.

We structured our paper as follows: Section 1 introduces the problem of malaria detection from medical images and provides a literature review on existing algorithms utilizing Convolutional Neural Networks (CNNs) for image classification. Section 2 describes the architecture of our proposed CNN model and outlines the training methodology, including data preprocessing and augmentation techniques. Section 3 presents the performance evaluation of our model, comparing it to traditional diagnostic methods, and discusses the quantitative results achieved in distinguishing between malaria-positive and malaria-negative cases. Finally, Section 4 concludes the paper by highlighting the implications of our findings and suggesting directions for future research in malaria detection.

2.1 Detection of Malarial Parasites in Blood using Image Processing

Malaria, caused by the Plasmodium parasite, is a highly infectious disease and a significant global health concern. Traditional microscopy, the "gold standard" for detection, is often inconsistent and time-consuming. To enhance detection, a system was developed using image processing techniques for rapid and reliable identification of malaria through stained thin blood smear images. This model analyzes datasets of both infected and healthy erythrocytes, extracting features to determine if a sample is infected [26, 34].

2.2 Malaria Cell Image Classification Using Deep Learning

Timely detection of malaria is crucial for effective patient treatment and preventing its spread through mosquitoes. Researchers advocate for treating malaria as a medical emergency and applying machine learning (ML) techniques to analyze microscopic red blood smear images. They utilize Convolutional Neural Networks (CNNs), a type of deep learning (DL) model, known for their scalability and efficiency in end-to-end feature extraction and classification. This approach serves as a valuable diagnostic tool for automated malaria detection. The study evaluates the performance of pre-trained CNNs as feature extractors for classifying parasitized and uninfected cells, identifying optimal model layers for statistical validation of results, thereby demonstrating the effectiveness of pre-trained CNNs in feature extraction [3, 23].

3 Experimental Design and Methodology

3.1 Preparing Dataset

The NIH Malaria Dataset, available from the Lister Hill National Center for Biomedical Communications (LHNCBC) at the National Library of Medicine (NLM), can be accessed at `https://lhncbc.nlm.nih.gov/publication/pub9932`. This dataset consists of 27,588 labeled and segmented cell images, all obtained from Giemsa-stained thin blood smear slides. It contains an equal distribution of uninfected red blood cells and *Plasmodium falciparum*-infected red blood cells, sourced from 150 infected individuals and 50 uninfected individuals.

The images in the dataset were manually annotated by an expert slide reader from the Mahidol-Oxford Tropical Medicine Research Unit. The dataset includes 13,779 parasitized red blood cell images and 13,779 uninfected red blood cell images. Positive samples contain *Plasmodium*, while negative samples do not contain *Plasmodium* but may include other objects such as staining artifacts or impurities. The segmented red blood cell patches are in RGB format, with sizes ranging from 110 to 150 pixels, which were later re-sampled to a uniform size of 128 x 128 pixels. The images have a depth of 3 channels and are stored in 32-bit floating point precision (FP32) to meet the input requirements of the classification algorithms used in this study. A selection of sample images from the dataset is shown in Figure 1.

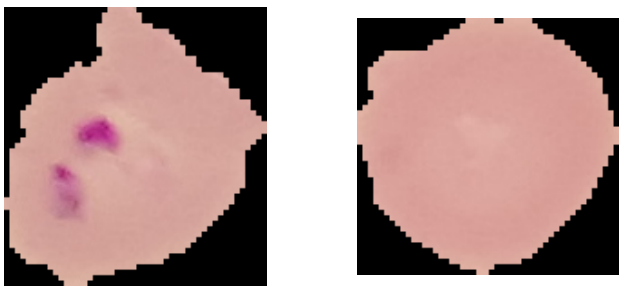


Figure 1: Comparison of parasitized (left) and uninfected (right) cell images

3.1.1 Data Preprocessing

To facilitate the training of deep learning models, a comprehensive series of preprocessing steps was methodically implemented for both the **Infected** and **Uninfected** directories. This careful and thorough preprocessing of the dataset guaranteed that the images were standardized and suitably prepared for applications in deep learning. Such an approach improves the model's capacity to effectively process the data, leading to dependable results in disease detection. Consequently, the resulting algorithm will establish a solid framework for disease identification through an extensive analysis of both infected and healthy images. Below are the specific preprocessing steps undertaken:

First, all image names were gathered and organized from the Infected and Uninfected folders to facilitate efficient processing. A filter was applied to retain only .png files, ensuring the dataset's integrity by excluding non-image files.

Next, light denoising was applied to the images, and all were resized to a standardized size (defined by the variable SIZE) to maintain uniformity, which is essential for improving model accuracy and performance. The resized images were then converted into arrays for efficient computational processing, enabling rapid data handling for the deep learning model. Finally, the processed images were added to the dataset, preparing it for model training.

3.1.2 Data Splitting

The dataset was first split into two subsets: training and test sets, with 80% of the data allocated to training and 20% to testing. Subsequently, the training data was further divided into 80% for training and 20% for validation. This approach ensures an effective division of data for training, validating, and testing the deep learning models, allowing for a comprehensive assessment of their performance on unseen data.

3.1.3 Data Augmentation

In this study, data augmentation was employed to enhance the training dataset's diversity and improve the robustness of the deep learning model. The augmentation process was facilitated using the `ImageDataGenerator` class from the TensorFlow Keras library, which allows for real-time data transformation during model training.

The relevant augmentation techniques are:

1. **Rescale**: Rescaling was applied by dividing the pixel values of the images by 255, adjusting them to a [0, 1] range. This normalization ensures consistent input data and enhances the model's training efficiency.
2. **Rotation**: Images were randomly rotated within a range of 20 degrees, allowing the model to learn from different orientations of the input data.
3. **Width and Height Shifts**: A shift of up to 20% of the total width and height was applied to the images. This horizontal and vertical translation enables the model to become invariant to the position of objects within the images.
4. **Shearing**: A shear transformation was applied within a range of 20%. This technique alters the perspective of the images, helping the model generalize better to variations in shape and orientation.
5. **Zoom**: Random zooming of up to 20% was applied to images, simulating varying distances of the subject from the camera.
6. **Horizontal Flipping**: Images were randomly flipped horizontally. This transformation is particularly useful for symmetrical objects, allowing the model to learn features from both orientations.
7. **Fill Mode**: The fill mode was set to 'nearest', ensuring that any newly created pixel values from the transformations are filled with the nearest pixel value, preventing any artifacts in the augmented images.

3.1.4 Evaluation Metrics

To comprehensively evaluate the robustness and generalization ability of the proposed CNN, 5-fold cross-validation was employed. In this technique, the dataset was divided into five equal folds. Each fold was used as a validation set once, while the remaining four folds served as the training set. This ensures that every data point is used for both training and validation, providing a reliable estimate of the model's performance. Averaging the performance metrics across all folds mitigates the effects of overfitting and provides a robust evaluation framework.

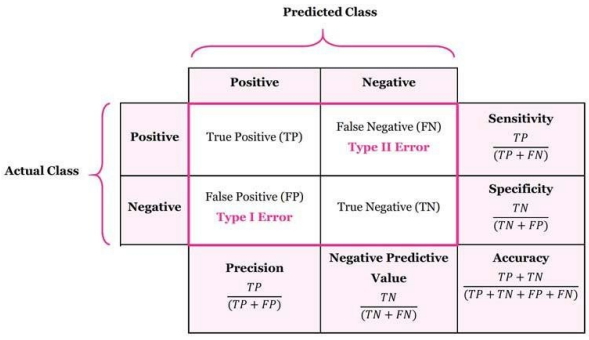


Figure 2: Performance measurement [33].

The evaluation metrics used to assess the model's performance include accuracy, precision, recall, and F1-score 2.

Here, the terms are defined as follows:

- True Negative (TN): These are cases where the model correctly predicts the negative class. In other words, the instances where the actual class is negative, and the model predicted it as negative.
- True Positive (TP): These are cases where the model correctly predicts the positive class. It refers to instances where the actual class is positive, and the model correctly predicted it as positive.
- False Negative (FN): These occur when the model predicts a negative outcome when it should have predicted a positive one. In other words, the instances where the actual class is positive, but the model predicted it as negative.
- False Positive (FP): These occur when the model predicts a positive outcome when it should have predicted a negative one. It refers to instances where the actual class is negative, but the model incorrectly predicted it as positive.

In our experimental framework, accuracy was designated as the primary metric for optimization, while the other metrics were used to validate the model's balanced performance across different evaluation criteria.

3.2 Methods

3.2.1 Mutual Information

Mutual information, when applied in neural networks, helps us understand how much useful information is passed from the input to the network's layers and, ultimately, to the output. It measures the connection between the input features (like image pixels) and what the network is learning. By focusing on the most important features, mutual information helps the network become more efficient, making better decisions with less noise or irrelevant data. This method ensures the network is learning from the most meaningful parts of the data, leading to improved accuracy in tasks like image classification.

3.2.2 Pretrained Models

Pretrained models are neural networks that have been previously trained on large datasets and can be fine-tuned for new tasks. By leveraging their learned features, these models allow for faster training and improved performance, especially when labeled data is limited. In this paper, we used DenseNet model. DenseNet's architecture allows efficient feature propagation by connecting each layer to every other layer, ensuring strong gradient flow during training. Its ability to capture intricate patterns in the data made it a suitable choice for our classification task, providing a solid baseline for comparison with our custom network.

3.2.3 Grad-CAM Visualization

To enhance the interpretability of the proposed model, Grad-CAM was employed to visualize the regions of the input images that contribute the most to the network's predictions. Heatmaps were generated using the final convolutional layer, as it encodes high-level task-specific features. Grad-CAM was applied to test samples to assess the generalizability and decision-making process of the model.

3.3 Proposed Architecture

The proposed CNN starts with convolutional layers using kernel sizes optimized based on dataset characteristics. Batch normalization and dropout layers were carefully tuned to prevent overfitting, while the Adam optimizer was selected due to its adaptive learning rate capabilities, which ensure stability across variable data distributions.

Convolutional Neural Networks (CNNs) are a specialized class of feedforward neural networks that have proven highly effective in extracting features from data through convolutional operations. Unlike traditional feature extraction methods, which often rely on manual engineering [1, 10, 20], CNNs autonomously learn hierarchical features directly from the input data. The architecture of CNNs is inspired by the mechanisms of human visual perception, where artificial neurons simulate biological ones, convolutional kernels act as receptors detecting specific patterns, and activation functions resemble the selective transmission of signals that exceed a defined threshold [18]. In addition, loss functions and optimizers are carefully designed to guide the network towards achieving task-specific learning objectives.

CNNs offer distinct advantages over traditional artificial neural networks. One of the key benefits is the use of local connections, where each neuron is connected only to a small subset of neurons in the previous layer, reducing the number of parameters and accelerating model convergence [20]. Furthermore, weight sharing allows multiple connections to share the same weights, simplifying the network structure and minimizing the risk of overfitting [15]. Dimensionality reduction through down-sampling, achieved by pooling layers, not only reduces computational complexity but also helps retain critical features while discarding irrelevant information [10]. These advantages contribute to CNNs' high efficiency in processing complex data, especially in tasks such as medical image analysis.

These capabilities make CNNs particularly suited for processing complex visual data, an essential requirement for recognizing patterns and detecting abnormalities in malaria-related medical images. Through the use of CNNs, researchers have made significant strides in automating the detection and classification of medical anomalies, demonstrating their transformative potential in healthcare applications [19, 29].

The proposed model is a custom Convolutional Neural Network (CNN) designed to classify malaria cell images with high accuracy. The architecture of the network is carefully crafted to extract relevant features while ensuring robustness and efficiency during training.

The network begins with an initial series of convolutional layers, each employing 32 filters of size 3x3, followed by ReLU activation functions to introduce nonlinearity. These layers are responsible for capturing low-level features such as edges and textures from the input images. The convolutional outputs are then passed through a MaxPooling layer with a pool size of 2x2, which reduces the spatial dimensions of the feature maps, preserving only the most salient information.

To enhance training stability and speed up convergence, batch normalization is applied after the convolutional layers. This technique normalizes the activations of each layer, helping maintain an optimal distribution of values throughout the network. Additionally, dropout is incorporated into the architecture to mitigate the risk of overfitting, randomly setting a portion of the inputs to zero during training and forcing the model to generalize better.

As the network progresses, more convolutional layers with 64 and 128 filters are introduced to capture increasingly complex and abstract features. These deeper layers are also followed by MaxPooling operations, which further reduce the spatial dimensions of the data while retaining essential features.

After the convolutional and pooling layers, the network flattens the output feature maps into a one-dimensional vector, which is then passed through a fully connected layer. This layer, consisting of 256 neurons, learns to combine the extracted features into higher-level representations that are essential for classification.

Finally, the output layer consists of a single neuron with a sigmoid activation function, which is appropriate for the binary classification task of distinguishing between malaria-infected and non-infected cells.

The model is trained using the Adam optimizer, with a learning rate of 0.001, and the categorical cross-entropy loss function, which is well-suited for binary classification tasks. During training, the model demonstrates robust performance, achieving high classification accuracy and outperforming several baseline models.

This custom CNN architecture was evaluated on a comprehensive dataset containing both malaria-infected and non-infected samples, demonstrating its efficacy in classifying medical images. The network's design, focusing on efficient feature extraction, regularization, and deep learning techniques, positions it as an effective solution for automated malaria detection.

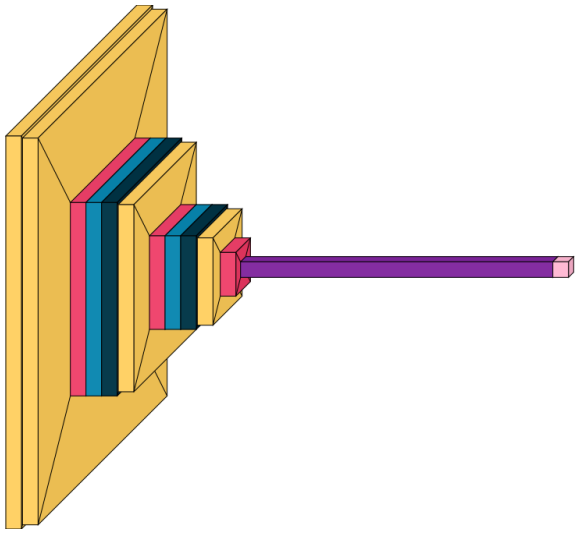


Figure 3: Proposed Architecture: Yellow - Convolutional, Red - Maxpooling,Blue - Batch Normalization, Green - Dropout, Purple - Dense, and last Layer - Classification.

See Figure 3 for a visual representation of the proposed network architecture.

3.4 Experimental Setup

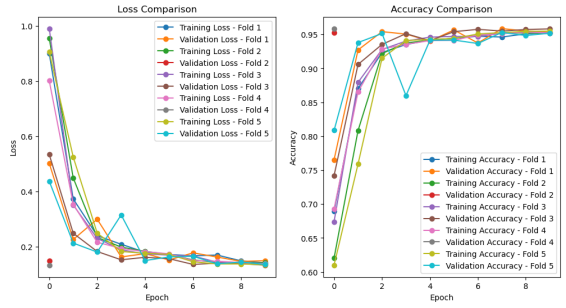
The computational setup comprised a robust infrastructure featuring 32 GB of RAM, an NVIDIA RTX 3060 6 GB graphics card, and a high-performance Intel Core i7 processor. The foundation of our methodology is built upon the utilization of Convolutional Neural Networks (CNNs), which are well-known for their superior performance in image classification tasks across various domains.

4 Results

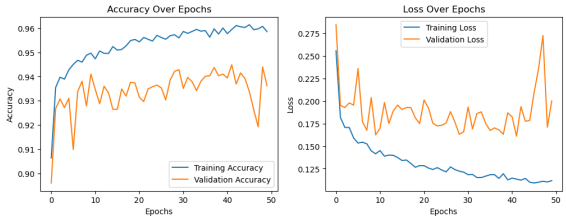
we conducted a comparative analysis using four distinct methods: 5-fold cross-validation, DenseNet121, mutual information, and the proposed custom architecture.

The comparative analysis shows that the proposed model outperformed the other approaches in terms of accuracy and efficiency, making it the most suitable solution for the malaria cell classification task.

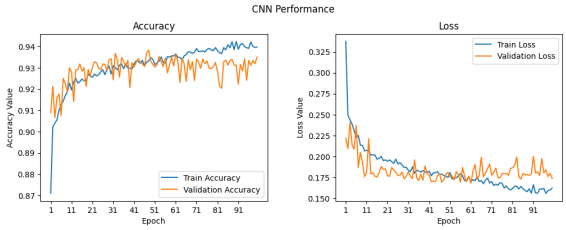
The proposed model achieved a Mean F1-score of 0.97 ± 0.01 at a 95% confidence interval, indicating that the model's performance is highly stable and consistent



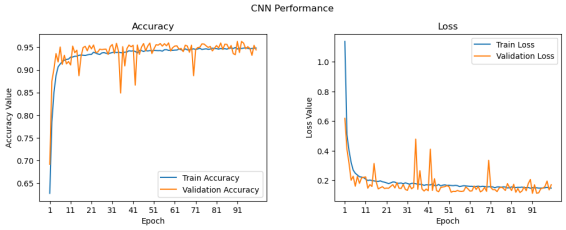
(a) 5-fold cross-validation



(b) DenseNet121



(c) Mutual information



(d) The proposed custom network

Figure 4: Comparison accuracy and loss values

across different folds of the dataset.

The narrow confidence interval ($\pm$ 0.01) of the Mean F1-score further suggests that the performance of the proposed model is stable across different splits of the dataset. This demonstrates the model's robustness and reliability for malaria cell classification tasks.

Furthermore, to statistically validate the superiority of the proposed model compared to DenseNet121, a t-test was performed. The p-value obtained ($p = 0.00047$) demonstrates that the difference in F1-scores between the proposed model and DenseNet121 is statistically significant ($p < 0.05$). This result confirms that the observed performance improvement is unlikely to have occurred by chance, reinforcing the robustness and reliability of the proposed approach for malaria cell classification.

The p-value ($p = 0.00047$) indicates that there is less than a 0.05% probability that the observed performance improvement occurred due to chance, reinforcing the statistical reliability of the proposed model.

In the classification report for our results, the proposed model achieved high precision and recall scores for both classes, leading to a strong F1-score. This indicates that the model not only predicts malaria-infected cells accurately but also detects most true positives, with a minimal false positive and false negative rate. These results further confirm that our proposed architecture is effective for this classification task, outperforming other models used in the study.

Additionally, to further evaluate the model's performance on the test dataset, a confusion matrix was generated. The confusion matrix provides a detailed breakdown of true positives, true negatives, false positives, and false negatives for both classes (Parasitized and Uninfected). This visualization offers deeper insight into the strengths and weaknesses of the proposed architecture in terms of classification accuracy. Figure 5 presents the confusion matrix, illustrating the model's robust performance with minimal misclassifications.

The Grad-CAM visualizations demonstrate how the proposed model focuses on specific regions of the input image to make predictions. Figure 6 illustrates an example of Grad-CAM applied to a test sample, showing the original input image, the generated heatmap highlighting regions of activation, and the overlay of the heatmap on the input image for better interpretability, respectively. These visualizations validate the proposed model's ability to focus on relevant areas, confirming the effectiveness of the proposed architecture.

Compared to DenseNet121, our proposed model achieved a statistically significant improvement (p ¡ 0.001) in F1-score, demonstrating its ability to generalize effectively on imbalanced datasets. Furthermore, the confusion matrix highlights that our architecture minimized false negatives, a critical metric in

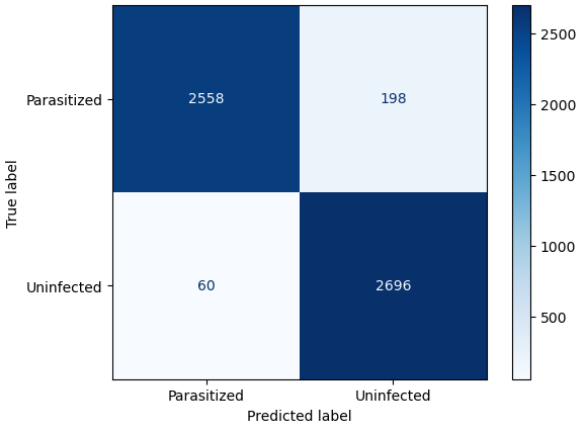
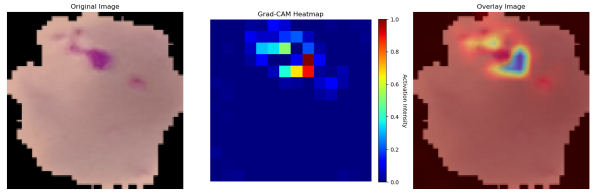
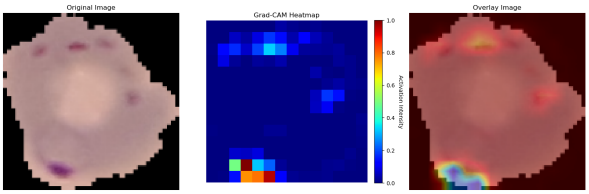


Figure 5: Confusion matrix of the proposed model on the test dataset, showing the classification results for the Parasitized and Uninfected classes.



(a)



(b)

Figure 6: Grad-CAM visualization applied on a (a) train and (b) test sample from the Parasitized class using the final convolutional layer of the neural network. Left to Right: Original image, Grad-CAM heatmap highlighting the regions of activation, and Overlay for Feature Activation Analysis.

malaria detection, where under-diagnosis could lead to life-threatening consequences.

Table 1: Classification Results for Parasitized and Uninfected Classes. Prec.: Precision, Rec.: Recall, F1: F1-score.

Method	Parasitized			Uninfected		
	Prec.	Rec.	F1	Prec.	Rec.	F1
ResNet50	0.82	0.52	0.64	0.65	0.89	0.75
DenseNet121	0.97	0.93	0.95	0.93	0.97	0.95
Mutual Info	0.93	0.93	0.93	0.94	0.93	0.93
Proposed	0.98	0.93	0.94	0.93	0.98	0.95

In summary, the proposed custom CNN demonstrates superior performance in terms of accuracy, F1-score, and statistical reliability compared to baseline models. The combination of a robust architecture and thorough preprocessing techniques ensures its effectiveness for automated malaria cell classification.

5 Conclusions

The proposed CNN architecture developed in this study outperforms other models, including 5-fold cross-validation, mutual information, and pretrained models such as DenseNet121 and ResNet50, achieving the highest accuracy of 96.15% on the test dataset. Statistical analysis, including confidence intervals and significance testing ($p = 0.00047$), confirms the model's reliability and makes it a viable solution for real-world deployment.

By leveraging a custom architecture tailored for malaria classification, the proposed model demonstrates a significant improvement over traditional and pretrained methods. This confirms its effectiveness in distinguishing between parasitized and uninfected cells with high accuracy and reliability.

The success of this approach underscores its potential for practical deployment in low-resource settings, addressing critical gaps in automated malaria diagnosis. Future research can explore integrating this model with lightweight architectures or real-time diagnostic tools to enhance its scalability and usability.

Future research could explore integrating the proposed CNN with lightweight architectures for real-time deployment on mobile devices, particularly in remote areas. Additionally, incorporating unsupervised learning techniques to handle unlabeled data could further enhance the model's robustness.

References

[1] T. Ahonen, A. Hadid, and M. Pietikainen. Face description with local binary patterns: Application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(12):2037–2041, 2006.

[2] S. Anwar and others. A review of deep learning in medical image analysis. *Medical Image Analysis*, 45:1–17, 2018.

[3] D. Bibin, M. S. Nair, and P. Punitha. Malaria Parasite Detection From Peripheral Blood Smear Images Using Deep Belief Networks. IEEE Access, 2017.

[4] N. Billo and others. Inter-observer variability in malaria diagnosis. *Journal of Clinical Microbiology*, 51(6):1865–1870, 2013.

[5] P. B. Bloland and World Health Organization. Drug Resistance in Malaria. *Technical Report*, World Health Organization, Geneva, Switzerland, 2001.

[6] K. Bowers and others. Diagnostic accuracy of microscopy-based malaria testing. *Malaria Journal*, 8:74, 2009.

[7] E. Charpentier and others. The role of light microscopy in malaria diagnostics. *Tropical Medicine and International Health*, 25(6):705–711, 2020.

[8] H. Caraballo and K. King. Emergency department management of mosquito-borne illness: malaria, dengue, and West Nile virus. *Emergency Medicine Practice*, 16(5):1–23, 2014.

[9] Centers for Disease Control and Prevention (CDC). Diagnosis of Malaria. 2020. `https://www.cdc.gov/malaria/diagnosis_treatment/diagnosis.html`

[10] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 886–893, 2005.

[11] T. Davenport and R. Kalakota. The potential for artificial intelligence in healthcare. *Future Healthcare Journal*, 6:94, 2019.

[12] C. Dong, C. C. Loy, and X. Tang. Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(2):295–307, 2016.

[13] U. Ezeoke and others. Reliability of malaria test results in Nigeria. *Nigerian Journal of Parasitology*, 33(1):15–22, 2012.
. P. Gopakumar, M. Swetha, G. S. Siva, and G. R. K. S. Subrahmanyam. Convolutional neural network-based malaria diagnosis from focus stack of blood smear images acquired using custom-built slide scanner. *Journal of biophotonics*, 11(3), 2018.

[14] K. Hong and others. The role of PCR in malaria diagnosis. *American Journal of Tropical Medicine and Hygiene*, 88(4):706–712, 2013.

[15] W. T. N. Hubel and D. H. Hubel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, 160(1):106–154, 1962.

[16] Y. Kan and others. Deep learning for histopathology: A review. *Journal of Pathology Informatics*, 8:38, 2017.

[17] M. Khan and others. Antigen detection kits for malaria. *Clinical Microbiology Reviews*, 23(2):203–227, 2010.

[18] A. Krizhevsky, I. Sutskever, and G. E. Hinton. ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25:1097–1105, 2012.

[19] G. Liang, L. Zheng, and Y. Zhang. Combining convolutional neural network with autoencoder for COVID-19 diagnosis on chest X-ray images. *Journal of Visual Communication and Image Representation*, 74:102929, 2020.

[20] T. Lindeberg. Scale invariant feature transform. In *2012*.

[21] W. Liu and others. SSD: Single Shot Multibox Detector. *Proceedings of the European Conference on Computer Vision (ECCV)*, 21–37, 2015.

[22] G. Litjens and others. A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42:60–88, 2017.

[23] Z. R. Mohd, A. Areeb, C. Firoz, and M. A. Shaikh. Malaria cell image classification using deep learning. *International Research Journal of Engineering and Technology (IRJET)*, 7(4):6259, 2020. e-ISSN: 2395-0056.

[24] L. Murungi and others. Serological tests for malaria detection. *PLOS ONE*, 14(3):e0214240, 2019.

[25] A. Nazar-Pembele, C. Rojas, and L. Ángel Nez. Training challenges for malaria microscopy technicians in Angola. *Malaria Research and Treatment*, 2016:1–6.

[26] C. P. Neha, K. Pallavi, P. R. Prapthi, A. Ahil, and Mrs. K. Ankitha. Detection of malarial parasites in blood using image processing. *International Journal of Engineering Research and Technology (IJERT)*, 2019. ISSN: 2278-0181.

[27] B. Ngasala and others. Evaluating microscopy-based malaria diagnostics in Tanzania. *Malaria Journal*, 11:302, 2012.

[28] S. Polpanich and others. Antigen detection methods for malaria. *Parasitology Research*, 101(4):941–947, 2007.

[29] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, et al. CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning. *arXiv preprint arXiv:1711.05225*, 2017.

[30] G. Rong, A. Mendez, E.B. Assi, B. Zhao, and M. Sawan. Artificial intelligence in healthcare: Review and prediction case studies. *Engineering*, 6:291–301, 2020.

[31] N. Tangpukdee, C. Duangdee, P. Wilairatana, and S. Krudsood. Malaria diagnosis: A brief review. *Korean Journal of Parasitology*, 47:93, 2009.

[32] D. Shen, G. Wu, and H. I. Suk. Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, 19:221–248, 2017.

[33] N. Shajihan. Classification of stages of Diabetic Retinopathy using Deep Learning. *Preprint*, 2020.

[34] M. S. Suryawanshi and P. V. V. Dixit. Comparative Study of Malaria Parasite Detection using Euclidean Distance Classifier SVM. *vol. 2, no. 11, pp. 2994–2997*, 2013.

[35] S. Wang and others. Deep learning for histopathology: A comprehensive review. *Artificial Intelligence in Medicine*, 97:1–9, 2019.