Archives of Business Research – Vol. 13, No. 03
Publication Date: March 25, 2025
DOI:10.14738/abr.1303.18433.
Coutinho, G. S., Medeiros, E. C., Fônseca, J. C. B., Moser, P. C., de Andrade, R. C. D., de Carvalho, F. F., de Souza Leão Júnior, F. P., & de Oliveira Domingues, M. A. (2025). Classification of Malaria-Infected Cells Using Convolutional Neural Networks: An Image-Based Microscopic Approach to Aid Diagnosis. Archives of Business Research, 13(03). 152-170.

Classification of Malaria-Infected Cells Using Convolutional Neural Networks: An Image-Based Microscopic Approach to Aid Diagnosis

Guilherme Silveira Coutinho
University of Pernambuco, Caruaru-PE, Brazil

Erika Carlos Medeiros
ORCID: 0000-0003-2506-7116
University of Pernambuco, Caruaru-PE, Brazil

Jorge Cavalcanti Barbosa Fônseca
University of Pernambuco, Caruaru-PE, Brazil

Patrícia Cristina Moser
University of Pernambuco, Caruaru-PE, Brazil

Rômulo César Dias de Andrade
University of Pernambuco, Caruaru-PE, Brazil

Fernando Ferreira de Carvalho
University of Pernambuco, Caruaru-PE, Brazil

Fernando Pontual de Souza Leão Júnior
University of Pernambuco, Caruaru-PE, Brazil

Marco Antônio de Oliveira Domingues
Federal Institute of Science and Technology of Pernambuco, Recife-PE, Brazil

ABSTRACT

This study investigates the use of convolutional neural networks to automatically detect malaria-infected cells in blood smear images, offering an alternative to manual diagnosis, which depends on specialized professionals and adequate infrastructure. Manual diagnosis is time-consuming and prone to human errors, especially in malaria-endemic regions with limited resources. Automated approaches based on convolutional neural networks provide a promising solution to optimize the diagnostic process and improve access to rapid treatment in remote areas. The research evaluates the performance of different convolutional neural network architectures for malaria diagnosis, following the Cross Industry Standard Process for Data Mining methodology to structure preprocessing, modeling, and model evaluation. Preprocessing involved normalization and data augmentation techniques to enhance sample quality and diversity. Two architectures were compared: a customized convolutional neural network designed to balance computational efficiency and accuracy, and an adapted VGG16, recognized for its

advanced image feature extraction capabilities. Both models were trained and evaluated using a robust dataset of blood cell images. The customized network achieved 95.6% accuracy, outperforming similar models in the literature and demonstrating its practicality for low-resource settings. It also exhibited high precision, recall, and F1-score, ensuring reliable and balanced classification of infected and healthy cells. This approach reduces reliance on specialists and advanced equipment, making malaria diagnosis more accessible and efficient. The findings position the customized convolutional neural network as a viable solution for automated malaria diagnosis, combining simplicity with high performance and offering significant potential for improving healthcare in endemic regions.

Keywords: Malaria Diagnosis, Convolutional Neural Network, Machine Learning, Image Processing

INTRODUCTION

Malaria remains a significant public health threat, especially in tropical and subtropical regions, where its high incidence and mortality rates directly impact the quality of life of affected populations. According to the World Health Organization (WHO) report of 2023, the disease affects millions of people every year, with more than 247 million cases reported in 2021 and an average of 619,000 annual deaths. In Brazil, most cases occur in the Amazon region, which accounts for approximately 99.9% of infections, reflecting the favorable environmental conditions for the proliferation of the malaria vector, the *Anopheles* mosquito. This scenario highlights the need for effective strategies for the diagnosis and control of the disease, particularly considering the high transmissibility and risk of recurrence in endemic areas [1].

The early diagnosis of malaria is essential to interrupt its transmission and minimize its severe consequences. According to the Malaria Treatment Guide in Brazil, the rapid identification of *Plasmodium*, the parasite that causes the disease and is transmitted by the female *Anopheles* mosquito, is crucial to prevent the worsening of symptoms and reduce the risk of severe complications. The traditional diagnostic method, thick blood smear microscopy, is widely recognized for its ability to detect and identify different species of *Plasmodium*. However, this technique requires highly qualified professionals to ensure precision and reduce the margin of error, which is challenging in areas with limited healthcare infrastructure. This scenario is even more complex in remote and hard-to-reach regions, where the availability of equipment and specialized professionals is restricted, underscoring the need to strengthen diagnostic strategies for effective malaria control in Brazil [2].

Beyond diagnostic challenges, malaria poses a significant social and economic burden, deeply affecting communities. In endemic regions, the disease undermines productivity, imposes high costs on healthcare systems, and creates barriers to economic development. According to the Epidemiological Bulletin published by the Brazilian Ministry of Health, without effective and timely diagnosis, the ability to respond to malaria is compromised, which can intensify the transmission cycle and increase the number of severe cases and deaths [3]. In this context, implementing solutions that promote early infection identification is fundamental for combating malaria, reducing mortality, and enabling more assertive interventions against the disease [4]. Additionally, malaria negatively impacts the Human Development Index (HDI) of

URL: http://doi.org/10.14738/abr.1303.18433

endemic areas, as observed in the Brazilian Amazon region, where the disease burden is associated with socioeconomic inequalities and structural limitations in healthcare systems [1]. The motivation to address the challenge of early malaria diagnosis lies in the urgent need to reduce mortality rates and improve the quality of life in endemic regions, where the disease continues to be a leading cause of morbidity and healthcare system overload. Investing in effective and accessible diagnostic methods is essential to promote sustainable development in these regions and align with the United Nations Sustainable Development Goals (SDGs), which aim to eradicate communicable diseases and promote global health [1].

Visual evaluation by specialists is one of the main challenges in malaria diagnosis, especially in endemic locations with limited resources. The traditional method of microscopic inspection of Giemsa-stained blood slides requires technicians to manually count parasites in red blood cells, which is not only time-consuming but also prone to variations among specialists, introducing the potential for inconsistencies and human errors in diagnosis. As noted by Davidson et al. [5], the accuracy of this manual approach can vary depending on the experience and working conditions of microscopists, highlighting a significant limitation of the traditional inspection method.

Another aggravating factor is sensory fatigue, which directly affects diagnostic accuracy in high-workload environments. The repetitive and detailed nature of manual parasite counting requires continuous visual effort and concentration from specialists, leading to visual fatigue and, consequently, an increased likelihood of errors, such as false negatives or positives. According to Siłka et al. [6], these problems are exacerbated in resource-limited environments, where fatigue and time pressure affect both the consistency and effectiveness of manual diagnosis. This reality reinforces the need for automated solutions that can complement and enhance the reliability of analysis, providing faster diagnoses with fewer errors.

Given this reality, there is a clear motivation to develop an improved solution that not only maintains high accuracy but also provides robustness to advance AI-assisted malaria diagnosis. This can be achieved by exploring convolutional neural network (CNN) configurations [7] capable of integrating into practical systems and enhancing the efficiency of infection identification.

Given this context, the central research question guiding this study is: Is it possible to explore CNN architectures capable of detecting malaria presence based on Giemsa-stained blood smear images?

The general objective of this work is to explore CNN architectures capable of detecting malaria presence based on Giemsa-stained blood smear images.

To achieve the general objective, four specific objectives were outlined:

1. Preprocess the image dataset downloaded from the repository of the National Library of Medicine (NLM) and the Lister Hill National Center for Biomedical Communications (LHNCBC) [8], where images were collected as described in the studies by Rajaraman et al. [9, 10]. This includes normalization, resizing, and data augmentation techniques to optimize feature extraction and improve model accuracy.

2. Explore and train two proposed CNN architectures using the preprocessed images from the previous step.
3. Identify and compare the proposed CNN architectures based on performance metrics such as accuracy [11], precision [12], recall [13], F1-score [14], confusion matrix [15], and the area under the Receiver Operating Characteristic Curve (ROC Curve) [16]. This will enable an understanding of each model's effectiveness.
4. Select the model with the highest performance metrics as the final model for detecting malaria presence based on Giemsa-stained blood smear images.

This work is structured into five sections. The present introductory section provides the context, motivation, research question, and objectives. The second section will cover related works, presenting recent research on machine learning for identifying malaria-infected cells using CNNs. The third section will describe the adopted methodology, detailing data preprocessing, model training, and performance evaluation stages. The fourth section will present the results and analysis of the models, while the fifth and final section will consolidate the conclusions of the study, highlighting contributions, limitations, and suggestions for future research.

In the next section, related works will be presented, providing a solid foundation for understanding existing approaches and identifying gaps to be explored.

RELATED WORK

In this section, relevant research and studies that provide the foundation and context for this work will be presented, offering a comprehensive overview of the state of the art in the field of study. Key concepts, methodologies, results from previous research, and gaps that serve as the theoretical and practical basis for the development of this work will be explored.

Jusman et al. [17] presented a model for detecting *Plasmodium* parasites in the schizont stage using CNNs and the pre-trained models GoogleNet and VGG-19, achieving up to 98.53% accuracy in the classification of three malaria species (*Plasmodium falciparum*, *P. vivax*, and *P. malariae*). This work stands out for its application of image augmentation and enhancement techniques, aiming for greater accuracy and reduced dependency on human specialists, especially in regions with limited medical infrastructure. However, limitations were observed in the extended training time for VGG-19 and the restricted variability of the dataset. Additionally, the use of an NVIDIA GeForce RTX 2060 GPU for training indicates a reliance on specialized hardware, which may limit applicability in locations where such resources are unavailable. These issues highlight the need to explore faster architectures and expand the diversity of images to achieve more robust generalization and a more accessible implementation of the model.

Janardhan et al. [18] highlighted a model for the automatic detection of *Plasmodium* parasites using CNNs and the VGG16 model, achieving an accuracy of 94.7% in distinguishing between infected and uninfected cells. The study highlights the use of segmentation and data enhancement techniques to reduce dependency on specialists and improve diagnostic accuracy, especially in regions with limited infrastructure. However, the study has limitations, including reliance on images from a specific stage of the parasite and the absence of a multi-

URL: http://doi.org/10.14738/abr.1303.18433

stage approach. For future advancements, it is suggested to apply the model to different stages of the parasite's lifecycle to enhance the generalization of the results.

Gupta et al. [19] investigated a sequential CNN model for the detection of malaria parasites in thin blood smears, achieving an accuracy of 94.6% and a precision of 97.5%. The study stands out for the development of an efficient and automated system that can operate without the need for specialized professionals, contributing to the reduction of human errors and accelerating diagnosis in resource-limited environments. However, the research identifies gaps, such as the need for testing on data from different malaria strains and adaptation for portable devices, suggesting directions for improvement and practical application in real-world contexts.

Siłka et al. [6] explored a deep neural network architecture for malaria detection, achieving an accuracy of 99.68% in distinguishing between infected and healthy cells. This work stands out for its use of a semantic segmentation architecture, enabling pixel-by-pixel detection of parasites in blood smears, which enhances the accuracy and feasibility of the solution for resource-limited environments. However, limitations include reliance on high-quality data to achieve this level of accuracy and the need for robust hardware for real-time processing. The study suggests that future research should focus on optimizations to enable the technique's use in lower-cost and more portable devices.

Raj et al. [20] proposed a deep CNN model for malaria parasite detection in thin blood smears, achieving an accuracy of 93.47% using different optimizers. The study stands out for implementing a custom CNN architecture for feature extraction and cell classification, aiming to reduce the need for human intervention and specialized equipment in low-infrastructure regions. However, limitations include reliance on high-quality images and the need for experimentation with learning rates to further improve accuracy, indicating potential areas for future advancements.

The following section describes the methodology used.

METHODOLOGY

The methodology used for the development of this work is based on the Cross Industry Standard Process for Data Mining (CRISP-DM) model, widely recognized in the field of data mining according to Wirth and Hipp [21]. The process will be structured into five of the six phases of this model: Business Understanding, Data Understanding, Data Preparation, Modeling, and Evaluation. The Implementation phase, although an integral part of CRISP-DM, will not be addressed in this work as it is outside the defined scope. This methodological approach will allow for a systematic and robust analysis of the data, ensuring clarity and coherence in the results achieved. The stages of the methodology are described below:

Business Understanding

The business understanding phase was conducted through a literature review, presented in Section 2, where the main impacts of AI in the field of malaria-infected cell classification using microscopic images were discussed. In this phase, the state of the art of existing solutions, the advancements brought by CNNs, and the gaps still present in the literature were identified. This

analysis enabled the contextualization of the relevance of applying AI in the healthcare field, highlighting its ability to assist in the diagnosis and treatment of diseases such as malaria.

Data Understanding

The data understanding phase began with a detailed analysis of the dataset used in this work. The dataset consists of a total of 27,560 cell image records, divided into two main classes: healthy cells and malaria-infected cells. Of the 27,560 images, 13,780 correspond to healthy cells, while 13,780 are images of infected cells, as illustrated in Figure 1. The dataset is originally balanced. These images were prepared for modeling with CNNs, enabling the classification of healthy and infected cells.

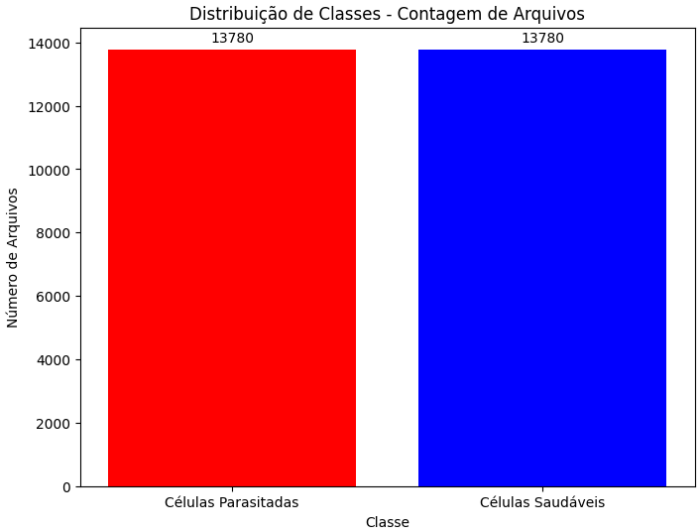

Figure 1: Class distribution in the dataset. (a) Distribution of healthy cell images, totaling 13,780 samples. (b) Distribution of malaria-infected cell images, also totaling 13,780 samples.

The images used in this work were collected by Rajaraman et al. [9, 10] as part of an international collaboration between the LHNCBC, part of the NLM, and the Mahidol Oxford Tropical Medicine Research Unit (MORU). Blood samples were prepared at the Chittagong Medical College Hospital in Bangladesh, where Giemsa-stained slides were obtained from 150 patients infected with *Plasmodium falciparum* and 50 healthy patients. Using an Android® smartphone attached to an optical microscope, images of different microscopic fields of view were captured and manually annotated by MORU specialists, ensuring precision in the identification of infected cells. The images were then de-identified and stored at the NLM, where segmentation algorithms were applied to automatically identify cells on the slides, facilitating the creation of a robust database for the development and testing of deep neural network models aimed at malaria detection [9, 10].

Data Preparation

First, a thorough check of the image folder contents was performed to ensure that all files present were valid images. Additionally, any corrupted images or non-image files that could compromise the model's performance were identified and removed. Next, the images were resized to a 160x160 format, suitable for feeding into the CNN used in the model. The choice of the 160x160 format was based on the need for a balance between visual quality and

URL: http://doi.org/10.14738/abr.1303.18433

computational efficiency, as original sizes such as 169x133 and 141x121 presented incompatible proportions and could introduce bias and unwanted noise. To prevent overfitting, the data augmentation technique was applied, using rotations, zooms, and flips to increase the diversity of the training data.

Modeling

For the modeling phase, the dataset was divided into three subsets: training (70%), validation (15%), and testing (15%), ensuring that the model was evaluated at different stages of training and tested with unseen data. Two CNN models were proposed: a custom CNN and an adapted VGG16 model, with their architectures illustrated in Figure 2 and Figure 3, respectively. The models were designed to be trained over 50 epochs; however, Early Stopping was applied at the 24th epoch to prevent overfitting.

The architecture of the custom CNN consists of multiple convolutional layers, each followed by a ReLU activation layer. The ReLU activation introduces non-linearity to the model, enabling it to learn complex patterns in the images. The number of filters in the convolutional layers starts with 32 in the first layer, followed by 64 in the second, 128 in the third, and so on, doubling the number of filters at each stage to capture increasingly detailed and complex features.

After each convolutional layer, a pooling operation is applied, specifically max-pooling with a filter size of 2x2. Max-pooling is a technique used to reduce the dimensionality of the extracted features without losing relevant information by selecting only the maximum value in each filter region. This results in a compact representation that emphasizes the most important features, helping to minimize computational requirements and reduce the risk of overfitting.

Following the convolutional layers, the network includes a fully connected layer with 128 neurons. This layer integrates the extracted features and generates a final representation of the images. The output layer consists of a single neuron with a sigmoid activation function, which is ideal for binary classification tasks, returning the probability of the image belonging to one of the classes. The model is optimized using the Adam optimizer with an initial learning rate of 0.001 and binary cross-entropy as the loss function, suitable for binary problems.

The second model is an adaptation of the VGG16 model, referred to as the "Adapted VGG16," tailored for a binary classification task. It begins with the base VGG16 network, a pre-trained convolutional neural network that takes input images in the 160x160x3 format (height, width, and color channels). This base performs deep feature extraction, leveraging pre-trained knowledge from large datasets such as ImageNet. The output of the adapted VGG16 is a 3D activation matrix with dimensions (5, 5, 512), where 512 represents the number of filters applied in the convolutional layers.

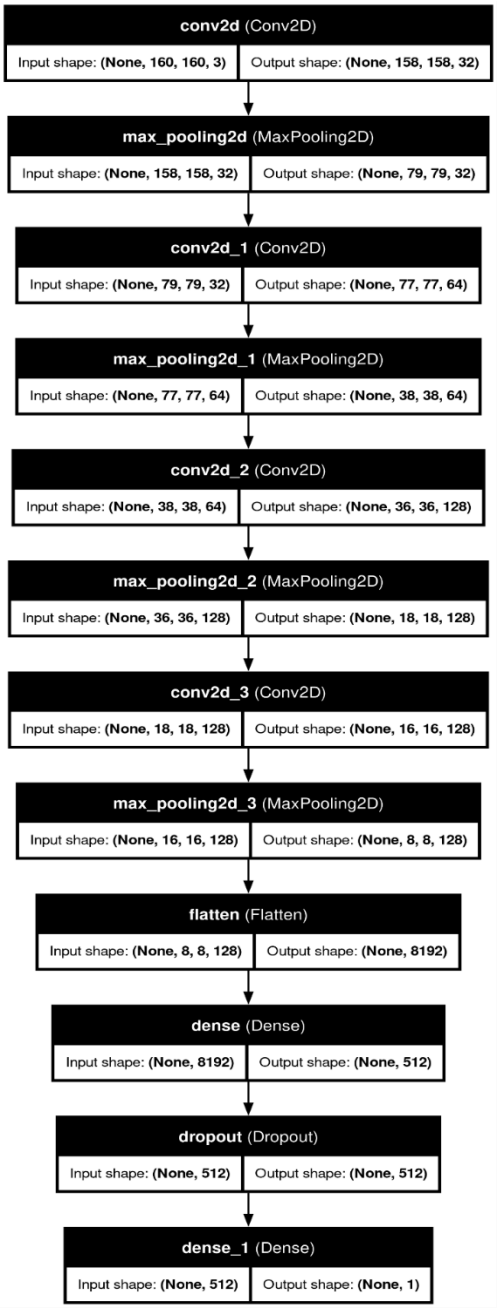

Figure 2: Architecture of the custom CNN.

Next, a Flatten layer transforms this 3D matrix into a one-dimensional vector with 12,800 units, preparing the output for subsequent dense layers. After the Flatten layer, a Dense layer with 512 neurons is added, allowing the model to learn combinations of features extracted by the adapted VGG16. Following this, a Dropout layer is included to reduce overfitting by randomly deactivating a fraction of the units during training, helping the model to generalize better.

Finally, a last Dense layer with a single neuron and a sigmoid activation function produces the model's output, which is suitable for binary classification tasks. This architecture combines the feature extraction capabilities of the adapted VGG16 with additional dense layers, allowing for

URL: http://doi.org/10.14738/abr.1303.18433

adaptation to the specific problem while retaining the generalization power learned by the base model.

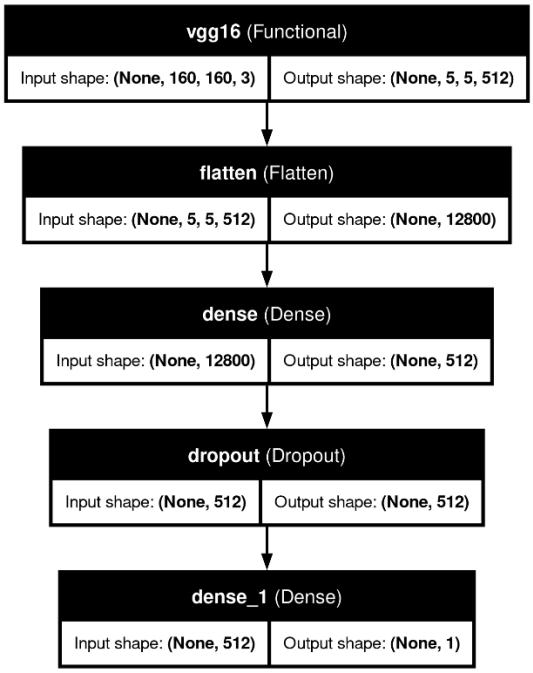

Figure 3: Architecture of the adapted VGG16 model.

Evaluation

In the evaluation phase, the performance of the models was analyzed using a diverse set of metrics. The models were assessed for accuracy, precision, recall, and F1-score, providing a detailed view of their effectiveness in various aspects of classification. Additionally, the confusion matrix was utilized to visualize the model's correct and incorrect predictions for each class. For a more comprehensive evaluation, the ROC curve was generated, representing the relationship between the true positive rate and the false positive rate, illustrating the model's ability to distinguish between classes. This thorough evaluation approach ensured an in-depth analysis of the models' performance.

In the next section, the results obtained from the evaluations will be discussed. The performance of the models will be analyzed based on the applied metrics, highlighting the strengths and limitations of each approach. This analysis will provide a better understanding of the models' effectiveness and help identify potential improvements for future work.

RESULTS

This section presents the performance evaluation results of the machine learning models applied to the predictive task of classifying malaria-infected cells. The evaluated models include the custom CNN and the adapted VGG16.

The analysis of the custom CNN and adapted VGG16 models for the task of classifying malaria-infected cells revealed significant differences in their performance. The custom CNN outperformed the adapted VGG16 across all evaluated metrics. As shown in Table 1, during

training, the custom CNN achieved an accuracy of 0.938, slightly higher than the adapted VGG16, which achieved 0.912. This difference indicates that the custom CNN was better able to capture the features of the training data, likely due to its architecture being more suitable for the microscopic cell image dataset analyzed.

During the validation phase, the custom CNN maintained its performance, achieving an accuracy of 0.956, as seen in Table 1, demonstrating its ability to generalize to new data. It again outperformed the adapted VGG16, which achieved an accuracy of 0.937. This result suggests that the custom CNN not only effectively learned the patterns from the training set but also avoided the issue of overfitting, successfully applying this knowledge to unseen data.

Table 1: Comparison of Model Performance.

Model	Training Accuracy	Validation Accuracy	Test Accuracy
Adapted VGG16	0.912	0.937	0.944
Custom CNN	0.938	0.956	0.956

In the final test, the custom CNN achieved an accuracy of 0.956, while the adapted VGG16 lagged slightly behind with an accuracy of 0.944, as seen in Table 1. This confirms that the custom CNN is the most effective model for predicting malaria-infected cells on completely new data.

The confusion matrix of the custom CNN, as shown in Figure 4, revealed 1,957 correct predictions for the "Infected" class and 1,997 for the "Healthy" class, with only 110 false positives and 70 false negatives. This highlights the custom CNN's strong ability to accurately distinguish between infected and healthy cells.

The errors observed in the confusion matrix of the custom CNN model, where 110 false positives (infected cells classified as healthy) and 70 false negatives (healthy cells classified as infected) were identified, can be attributed to several factors. One reason is the visual similarity in the images. In some cases, malaria-infected cells exhibited visual characteristics very similar to healthy cells, especially in early stages or mild infections.

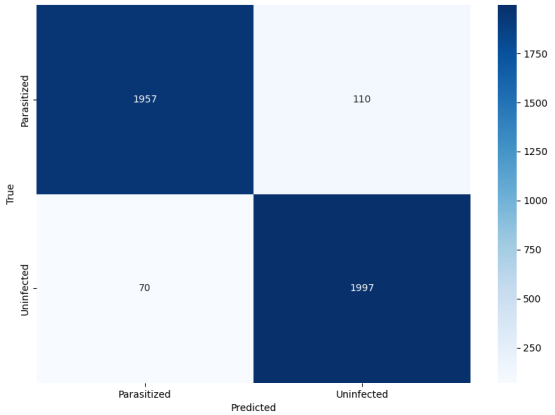

Figure 4: Confusion matrix of the custom CNN.

The texture, coloration, and patterns in microscopic images are subtle and challenging to distinguish, even for a trained model. Additionally, healthy cells showed artifacts or natural variations that made them appear infected, leading to incorrect classifications by the model.

The quality of the images also directly impacted the model's performance. The dataset contains some low-resolution images with noise, inconsistent lighting, or slight blurriness, which hindered the model's ability to learn distinctive features. These variations confused the model during the classification process, resulting in the observed false positives and false negatives. The confusion matrix of the adapted VGG16 showed 1,905 correct predictions for the "Infected" class and 2,000 for the "Healthy" class, with 162 false positives and 67 false negatives, indicating a slight tendency to misclassify infected cells as healthy.

The classification report of the custom CNN, shown in Table 2, revealed a precision and recall of 0.97 and 0.95 for the "Infected" class and 0.95 and 0.97 for the "Healthy" class, resulting in an F1-Score of 0.96 for both classes. These values outperform those presented in the classification report of the adapted VGG16, shown in Table 3, which achieved a precision and recall of 0.97 and 0.92 for the "Infected" class and 0.93 and 0.97 for the "Healthy" class. The lower recall rate for the "Infected" class in the adapted VGG16 suggests a higher false negative rate, which could be problematic in a clinical context.

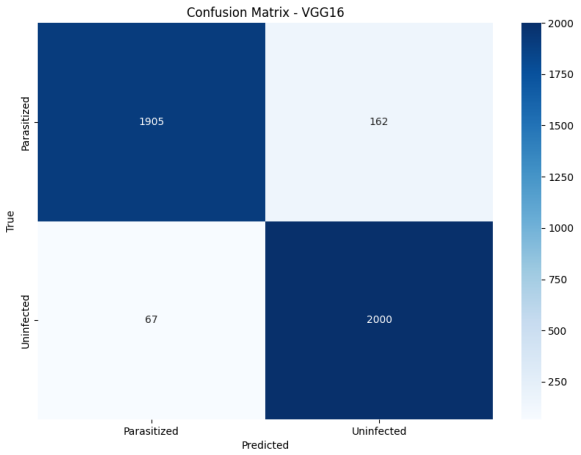

Figure 5: Confusion matrix of the adapted VGG16 model.

Table 2: Classification Report of the Custom CNN.

	Precision	Recall	F1-Score
Infected	0.97	0.95	0.96
Healthy	0.95	0.97	0.96
Accuracy			0.96
Macro Average	0.96	0.96	0.96
Weighted Average	0.96	0.96	0.96

Both models exhibited a high area under the ROC curve, as shown in Figure 6, with the custom CNN achieving a value of 0.99 and the adapted VGG16 reaching 0.98.

The learning curves of the custom CNN, shown in Figure 7, demonstrate a rapid increase in training accuracy, stabilizing around the 5th epoch, while the model's loss consistently decreased over time.

The validation accuracy remained close to the training accuracy, indicating that the model did not suffer from overfitting. On the other hand, the learning curves of the adapted VGG16, shown in Figure 8, displayed greater variability in validation accuracy, suggesting that the model struggled to find a stable point during training.

Table 3: Classification Report of the Adapted VGG16.

	Precision	Recall	F1-Score
Infected	0.97	0.92	0.94
Healthy	0.93	0.97	0.95
Accuracy			0.94
Macro Average	0.95	0.94	0.94
Weighted Average	0.95	0.94	0.94

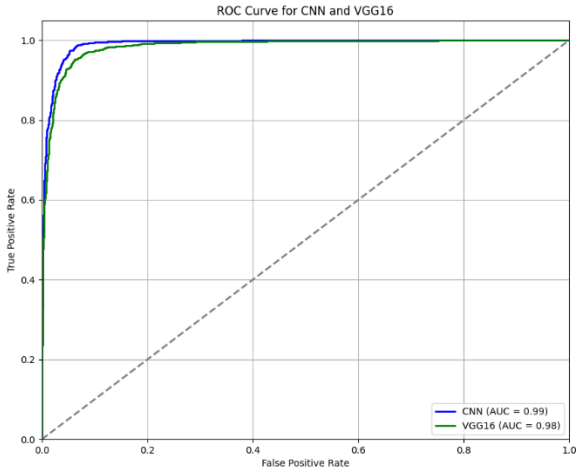

Figure 6: Area under the ROC curve for both models.

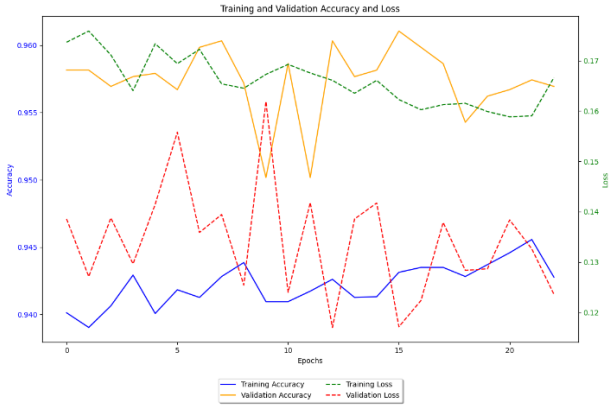

Figure 7: Learning Curves of the Custom CNN.

Based on these analyses, the selection of the custom CNN model is justified by its superiority in essential metrics: the training, validation, and test accuracies were higher compared to the adapted VGG16. The confusion matrix indicates a lower rate of critical errors (false negatives), precision and recall demonstrate a reliable ability to correctly identify infected cells, and the F1-Score reflects a proper balance between these measures.

Additionally, the ROC curve reinforces the model's effectiveness in class separation, while the learning curves demonstrate that the custom CNN learned efficiently without overfitting. These combined factors make the custom CNN the ideal choice for the task of classifying malaria-infected cells, providing safer and more reliable performance for application in clinical scenarios.

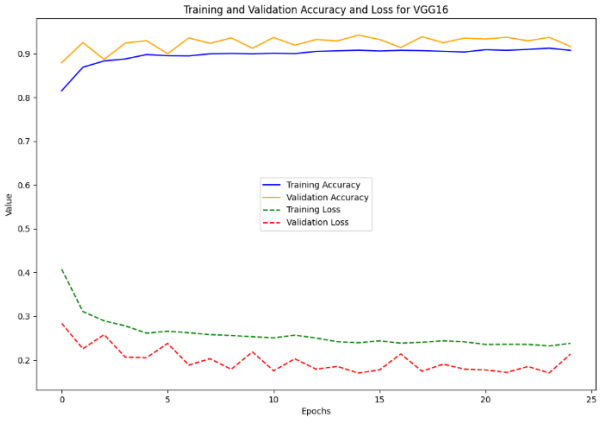

Figure 8: Learning Curves of the Adapted VGG16.

In the next section, the conclusion of the study will be presented, consolidating the main findings and discussing the contributions of the developed model. An analysis of the study's limitations will be conducted, highlighting areas for improvement that can be explored in future research. This section will also provide an overview of this study in relation to related approaches, aiding the understanding of its impact and suggesting pathways for advancements in automated malaria detection.

CONCLUSION

This study addresses the urgent need for efficient and automated diagnostic methods for malaria, which overcome the limitations of traditional microscopy-based processes. These methods require highly skilled professionals and are prone to human errors, especially in endemic regions with limited healthcare infrastructure. As an alternative, this work explores the use of CNNs for automated detection of malaria infection in blood smear images, aiming to improve diagnostic accuracy and accessibility in resource-constrained settings.

To achieve the ultimate goal of developing an automated and efficient method for malaria detection, four specific objectives were progressively established and addressed.

First, the preprocessing of the blood smear image dataset stained with Giemsa, downloaded from the NLM repository, was carried out. This step involved normalization, resizing, and data augmentation techniques to optimize the extraction of visual features and improve model

accuracy. Upon completing this preparation, the images were properly formatted to feed the selected CNNs for the study.

Next, two CNN architectures were trained using the preprocessed images. During this stage, various hyperparameter configurations were applied to fine-tune the learning process and ensure that both networks efficiently captured the distinctive patterns of infected cells compared to healthy ones.

Following the training, the next objective was to compare the two CNN architectures based on several performance metrics, such as accuracy, precision, recall, F1-Score, confusion matrix, and ROC curve. This analysis allowed for an evaluation of the effectiveness of each model and the identification of the neural network with the best classification capability, highlighting the most robust architecture for the task.

Finally, the model with the best performance metrics was selected as the final solution for automated malaria detection. The chosen model proved to be more reliable for clinical applications, demonstrating that it successfully fulfilled the overall objective of developing an effective and accessible diagnostic solution. By addressing all the specific objectives, the ultimate goal of this study was achieved.

In this study, two CNN models were explored: a custom CNN and an adaptation of the VGG16 model. The custom CNN was developed with multiple convolutional layers and max-pooling for feature extraction and reduction, followed by a dense layer to integrate this information and a binary output to classify images as either infected or healthy cells. Meanwhile, the VGG16 model, a pre-trained network, was adapted for the task, leveraging its robust architecture for deep feature extraction from the images.

The results showed that the custom CNN outperformed the adapted VGG16 model across various performance metrics. With higher training, validation, and test accuracies, as well as a more balanced F1-Score, the custom CNN demonstrated greater precision in identifying infected cells while maintaining the ability to generalize well on unseen data. The confusion matrix revealed a lower false-negative rate in the custom model, a critical factor for minimizing the risk of incorrect diagnoses.

Thus, the final proposed solution employs the custom CNN, which proved to be more suitable for automated malaria detection, offering a reliable and efficient alternative to traditional diagnosis, especially in resource-limited areas.

Building on the studies presented in Section 2, this work aimed to develop a specific solution for malaria detection that is both accurate and feasible in resource-limited contexts. Unlike models such as VGG-19 and GoogleNet, explored by Jusman et al. [17], which achieved a high accuracy of 98.53% but require robust hardware infrastructure, this study focused on a custom CNN better suited for environments with resource and processing time constraints. The custom CNN demonstrated high accuracy and reliable performance, achieving 95.6% accuracy without relying on advanced GPUs, making it more accessible for malaria-endemic regions with limited infrastructure.

Additionally, unlike the work of Janardhan et al. [18], which used segmentation in a single stage of the parasite's lifecycle, this study adopts a broader approach by classifying between infected and healthy cells, aiming to optimize diagnosis even at varying stages of infection. While Janardhan et al. achieved an accuracy of 94.7% with the VGG16 model, the custom CNN developed in this study surpassed this result, reaching an accuracy of 95.6%. This improvement adds value to its application in real-world scenarios, where cells may exhibit variations across different stages of infection.

In comparison to the model proposed by Gupta et al. [19], which achieved high precision but faced limitations in dataset variability and reliance on specialized hardware, the solution developed in this study incorporates data augmentation techniques to diversify the image set. This enhances the generalization capability of the custom model and reduces the need for advanced hardware, such as that used by Gupta et al., whose model achieved 94.6% accuracy, below the 95.6% reached by our custom CNN. This improvement allows for a more practical implementation.

Considering the work of Raj et al. [20], whose custom CNN achieved an accuracy of 93.47%, this study presented an optimized CNN that outperformed that performance, reaching 95.6%. This improvement in accuracy provides a more precise and efficient solution, particularly advantageous for resource-limited contexts where specialized hardware may be scarce. The approach in this study reinforces the feasibility of a more reliable diagnosis in endemic regions, maintaining high accuracy without the need for advanced infrastructure.

Finally, compared to the model by Siłka et al. [6], which employs a pixel-by-pixel semantic segmentation architecture for high precision but requires high-quality data and specialized hardware, this study stands out for the simplicity and efficiency of its custom CNN. The model achieved a competitive accuracy of 95.6% compared to the 99.68% precision of Siłka et al.'s model, yet with lower hardware requirements and reduced complexity. This approach decreases diagnostic costs and time, making it a viable and scalable alternative for field use.

Table 4 presents an analysis of malaria detection studies, highlighting the contributions, methodologies, and gaps of each approach. Among the models analyzed, the custom CNN developed in this study offers a balanced solution between accuracy and feasibility for resource-limited environments, surpassing the accuracy of studies such as those by Raj et al. [20] and Gupta et al. [19]. Unlike models like VGG-19 and GoogleNet, which, despite their high precision, require robust infrastructure, the approach in this study is more accessible for malaria-endemic regions. Table 4, therefore, illustrates how the custom CNN stands out in terms of efficiency and precision compared to pre-trained models, particularly for practical implementation contexts.

Thus, this study presents a distinct contribution in terms of balancing performance and practical feasibility, focusing on an automated and accessible solution for malaria detection that can be applied in various infrastructure conditions, making the diagnosis more comprehensive and accessible.

The study does present some important limitations. First, it relies on high-quality blood smear images for accurate classification, meaning that images with low resolution, noise, or lighting variations may negatively impact the model's performance. Additionally, the custom CNN was trained on a specific dataset, which may limit its generalization to other malaria image datasets or sources. In such cases, adjustments in preprocessing or even retraining may be required. Another limitation is that, while the model automates the diagnosis, it does not replace the detailed clinical analysis of a specialist, as it does not capture clinical nuances that a professional might identify. The use of data augmentation to balance the classes also introduces the risk of overfitting, especially if the model is overly adjusted to the training set, potentially reducing its performance on completely new data. Finally, the model was designed for binary classification (infected vs. healthy cells), and to identify different *Plasmodium* species or stages of infection, further adaptations would be necessary. These limitations suggest areas for improvement and point to potential approaches to make the model more robust and applicable in various malaria diagnostic contexts.

Table 4: Analysis of Studies for Malaria Detection.

Study	Main Contributions	Methodological Process	Gaps
This Study	Exploration of a custom CNN and an adapted VGG16 for malaria detection in blood smear images; applicable in resource-limited areas. Achieved an accuracy of 95.6%.	Follows the CRISP-DM methodology, focusing on preprocessing (normalization and data augmentation) and comparing the custom CNN with the adapted VGG16.	Dependence on high-quality images and limitation to binary classification (infected vs. non-infected cells).
Jusman et al. [17]	Classification of malaria parasites in the schizont stage using GoogleNet and VGG-19, with an accuracy of up to 98.53% to detect multiple malaria species.	Uses pre-trained GoogleNet and VGG-19 models for classification; cross-validation methods (10-Fold Cross Validation) and fine-tuning were applied to improve accuracy and reduce execution time.	Focus on a specific stage of the parasite lifecycle; high time consumption for processing, limiting its use in regions with low infrastructure and no support for different stages of the parasite lifecycle.
Janardhan et al. [18]	Detection of *Plasmodium* parasites using CNN and VGG models, achieving 94.7% accuracy, focusing on automation to reduce human errors.	Uses open-source data, applying data augmentation to improve model performance, with preprocessing focused on segmentation and noise reduction.	High computational cost and the need for robust hardware for deployment; the impact of variations in image quality is still a concern.
Gupta et al. [19]	Implementation of a sequential CNN model for parasite detection with 94.6% accuracy, promoting greater	Uses a sequential CNN model with an 80:20 data split for training and testing, along with	Tested only with limited data; needs validation with different malaria strains and specialized

URL: http://doi.org/10.14738/abr.1303.18433

	automation in malaria diagnosis.	preprocessing to improve accuracy.	hardware for deployment in large scale.
Siłka et al. [6]	Semantic segmentation network for high precision (99.68%) in distinguishing infected and healthy cells; viable for resource-limited environments.	Uses an advanced CNN architecture for pixel-by-pixel semantic segmentation, allowing for greater precision in environments where image quality varies.	Requires high-resolution images and robust hardware; complex to implement in areas with limited infrastructure; segmentation methods limited to certain malaria subtypes.
Raj et al. [20]	Custom CNN to detect malaria parasites in thin blood smears, with 93.47% accuracy, applicable in resource-limited regions.	Uses feature extraction techniques and cell classification with a custom CNN. The methodology includes validation with three different optimizers: Adam, RMSprop, and SGD.	Need for experimentation with different learning rates and lack of multi-stage approach, limiting the generalization of results.

For future work, several improvements and extensions can be explored. First, a promising step would be to adapt the model for multiclass classification, enabling the identification of different *Plasmodium* species or specific stages of infection, which could expand its applicability in more detailed clinical diagnoses. Additionally, using advanced image preprocessing techniques to improve the quality and visual consistency of samples, such as noise removal and automatic lighting adjustments, could help increase the robustness of the model in the face of image variations.

Another interesting avenue for further exploration would be the development of transfer learning strategies, allowing the model to benefit from data from other sources or domains, thus expanding its generalization capability to different contexts and datasets. Future work could also investigate the use of model interpretability techniques, such as Grad-CAM, to provide a visual interpretation of the image regions most relevant to parasite detection, helping doctors and specialists understand the model's decisions.

Finally, conducting tests in real clinical environments, with data collected directly from laboratories in endemic regions, would be essential to validate the model in practical scenarios and assess its effectiveness as an auxiliary tool in malaria diagnosis. These future endeavors have the potential to strengthen the developed model and expand its impact on malaria diagnosis and control in endemic areas.

References

[1]. World Health Organization. World malaria report 2023. Geneva: WHO, 2023.

[2]. BRASIL. Ministério da Saúde. Guia de tratamento da malária no Brasil. Brasília, DF: Ministério da Saúde, Secretaria de Vigilância em Saúde, Departamento de Imunização e Doenças Transmissíveis, 2020.

[3]. BRASIL. Ministério da Saúde. Dia da malária nas Américas – um panorama da malária no Brasil em 2022 e no primeiro semestre de 2023. Boletim Epidemiológico da Secretaria de Vigilância em Saúde e Ambiente, 55, Jan 2024.

[4]. Carvalho, C.C.; Oliveira, G.L.; Antunes, Y.R. Malária e a eficácia diagnóstica para o controle da doença. Brazilian Journal of Development, 9(5), 16680-16698, May 2023. DOI: 10.34117/bjdv9n5-145.

[5]. Davidson, M.S.; Andradi-Brown, C.; Yahiya, S.; Chmielewski, J.; O'Donnell, A.J., et al. Automated detection and staging of malaria parasites from cytological smears using convolutional neural networks. Biological Imaging, 1, e2, 2021. DOI: https://doi.org/10.1017/S2633903X21000015.

[6]. Siłka, W.; Wieczorek, M.; Siłka, J.; Woźniak, M. Malaria Detection Using Advanced Deep Learning Architecture. Sensors, 23(3), 1501, 2023. DOI: https://doi.org/10.3390/s23031501.

[7]. O'Shea, K.; Nash, R. An Introduction to Convolutional Neural Networks. 2015. Available at: https://doi.org/10.48550/arXiv.1511.08458. Accessed Nov 11, 2024.

[8]. National Library of Medicine. Malaria Datasets. Bethesda, MD: National Institutes of Health, Lister Hill National Center for Biomedical Communications, 2024. Available at: https://lhncbc.nlm.nih.gov/LHC-downloads/downloads.html#malaria-datasets.

[9]. Rajaraman, S.; Antani, S.; Pocock, J. et al. Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images. Heliyon, 4(10), e00885, 2018. Available at: https://www.sciencedirect.com/science/article/pii/S193152441730333X.

[10]. Rajaraman, S.; Silamut, K.; Hossain, M.A.; Ersoy, I.; Maude, R.J.; Jaeger, S.; Thoma, G.R.; Antani, S.K. Understanding the learned behavior of customized convolutional neural networks toward malaria parasite detection in thin blood smear images. Computers in Biology and Medicine, 101, 102-112, 2018. Available at: https://pmc.ncbi.nlm.nih.gov/articles/PMC6050500/#sec2.

[11]. Monico, J.F.G.; Dal Póz, A.P.; Galo, M.; Santos, M.C.; Oliveira, L.C. Acurácia e precisão: revendo os conceitos de forma acurada. Boletim de Ciências Geodésicas, 15(3), 469-483, Jul-Sep 2009. Available at: http://www.redalyc.org/articulo.oa?id=393937709010. Accessed Nov 11, 2024.

[12]. Mariano, D. Métricas de avaliação em Machine Learning: acurácia, sensibilidade, precisão, especificidade e F-score. BIOINFO - Revista Brasileira de Bioinformática e Biologia Computacional, 2021.

[13]. Torgo, L.; Ribeiro, R. Precision and recall for regression. In: Discovery Science: 12th International Conference, DS 2009, Porto, Portugal, Oct 3-5, 2009. Proceedings. Berlin Heidelberg: Springer, 2009, 332-346.

[14]. Zhang, D.; Wang, J.; Zhao, X. Estimating the uncertainty of average F1 scores. In: International Conference on the Theory of Information Retrieval, 2015, 317-320.

[15]. Liang, J. Confusion matrix: Machine learning. POGIL Activity Clearinghouse, 3(4), 2022.

[16]. Nakas, C.T.; Bantis, L.E.; Gatsonis, C.A. ROC analysis for classification and prediction in practice. Chapman and Hall/CRC, 2023.

[17]. Jusman, Y.; Aftal, A.A.; Tyassari, W.; Hayati, N.; Kanafiah, S.N.A.M.; Mohamed, Z. Classification of Parasite Malaria in Schizon Stage with GoogleNet and VGG-19 Pre-Trained Models. In: 2023 10th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE), IEEE, 2023, 219-223.

[18]. Janardhan, G.; Shriya, D.; Babu, D.A.; Kumar, S.A.; Kishore, R.S. Automatic Detection of Plasmodium Parasite Using Convolution Neural Network. In: Proceedings of the International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS 2023), IEEE, 2023, 145-150.

[19]. Gupta, M.; Naga Sabarish, Y.S.; Dungarwal, Y.; Anil Kumar, K. A Novel Approach Towards Detecting Malaria Parasite in Thin Blood Smears Using a Sequential Convolutional Neural Network (CNN) Model. In: 2023 First International Conference on the Advancements of Artificial Intelligence in African Context (AAIAC), IEEE, 2023, 145-150.

[20]. Raj, M.; Sharma, R.; Sain, D. A Deep Convolutional Neural Network for Detection of Malaria Parasite in Thin Blood Smear Images. In: 10th IEEE International Conference on Communication Systems and Network Technologies (CSNT), IEEE, 2021, 510-514. DOI: 10.1109/CSNT.2021.88.

[21]. Wirth, R.; Hipp, J. CRISP-DM: Towards a Standard Process Model for Data Mining. In: Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining, 2000.