Available online at sjuoz.uoz.edu.krd

journals.uoz.edu.krd

Science Journal of University of Zakho

Vol. 13, No.3 pp342–347July-September,2025

p-ISSN: 2663-628X
e-ISSN: 2663-6298

A HYBRID APPROACH FOR MALARIA CLASSIFICATION USING CNN-BASED FEATURE EXTRACTION AND TRADITIONAL MACHINE LEARNING CLASSIFIERS

Walat A. Ahmed[1], Bafreen N. Mohammed[2], Asaad Kh. Ibrahim[3], and Omar M. Ahmed[4,*]

[1] Computer Information System Department, Technical College of Zakho, Duhok Polytechnic University, Zakho, KRG, Iraq

*Corresponding author email: omar.alzakholi@dpu.edu.krd

Received: 22 Feb. 2025 Accepted: 26 Apr. 2025 Published:02 Jul. 2025 https://doi.org/10.25271/sjuoz.2025.13.3.1489

ABSTRACT:

Malaria is a major global health threat, and timely and correct diagnosis is essential for effective treatment. Traditional diagnostic methods, such as the microscopic examination of blood smears, are time-consuming and require expert personnel. The study presents a mix of machine learning methods for automatic diagnosis of malaria by using the feature extraction capability of Convolutional Neural Networks (CNNs) along with the efficient classification performance of traditional machine learning classifiers. For our study, we utilize VGG16 CNN with a weight pre-trained on ImageNet to extract the features from non-infected and infected blood cell images from malaria. Five classical machine learning algorithms, such as Random Forest, Logistic Regression, K-Nearest Neighbours (KNN), Support Vector Machine (SVM), & Gradient Boosting, are used to classify the extracted features. Each classifier's performance is calculated based on accuracy, F1 score, precision, and recall metrics. The results of our experiments showed that the hybrid model has high accuracy in classification, where the Logistic Regression classifier could achieve above 93% accuracy. This hybrid method is a powerful diagnostic for malaria disease, accomplishing a more satisfactory compromise between the efficacy of the deep learning architectures such as CNNs, and the computational capabilities of more conventional classifiers. It holds promise for deployment into resource-limited settings where fast, automated threading diagnostic systems are much needed.

KEYWORDS: Malaria classification, Convolutional Neural Networks, Feature extraction, Random Forest, Logistic Regression.

1. INTRODUCTION

Malaria is a highly prevalent and life-threatening parasitic disease that affects humans worldwide, especially in tropical and subtropical areas. Caused by protozoan parasites of the genus Plasmodium, the disease is transmitted to humans via the bites of infected female Anopheles mosquitoes. According to the World Health Organization (WHO), malaria is a major public health threat and is most common in low- and middle-income countries. In 2019 alone, there were an estimated 229 million cases of malaria around the world, causing more than 400,000 deaths, mostly in sub-Saharan Africa. Malaria is a leading cause of death among children under five years of age. Despite extensive control and elimination programs, malaria remains a global health burden (Al-Awadhi *et al*., 2021).

Early and accurate diagnosis is one of the most important components of a comprehensive approach to combat malaria. Traditional diagnostic tools include blood smear microscopy, rapid diagnostic tests (RDTs), and molecular diagnostic tools such as polymerase chain reaction (PCR). The most common approach is microscopy of stained blood smears to detect malaria parasites by professionals who have been properly trained to perform this task. Despite the availability of microscopy and rapid diagnostic tests (RDTs), malaria diagnosis faces significant challenges, including dependence on skilled personnel, potential for human error, variability in test quality, high costs of molecular methods, and limited access to diagnostic facilities in rural or low-resource regions (Awosolu *et al*., 2022).

One way to overcome some such challenges is to automate the diagnosis process. Using image analysis techniques and machine learning algorithms, computers can be trained to recognise malaria parasites in blood smear images quickly and accurately, helping to reduce diagnostic error and improve access to reliable malaria diagnosis, especially in rural areas where the availability of expertise may be limited. Deep learning has shown impressive results in image classification problems in various fields, particularly in medical imaging. These technologies have the potential for solving and enhancing malaria detectability (Grignaffini *et al*., 2024), yielding faster and accurate results.

Deep Learning and CNNs techniques have dominated tasks of image classification in the last years. CNNs are a class of artificial neural networks whose structure is inspired by the biological structure of the animal visual cortex, which takes in and processes structured grid data like images. Due to their nature of learning feature representations from pixel data automatically and hierarchically, they are very suitable for image-related tasks. A CNN can learn different patterns like edges, textures and shapes from an image at its lower layers are patterns to help the network construct complex representations and objects or in the case of medical imaging, pathological features (Gao *et al*., 2022). CNNs have been successfully applied to medical imaging, including malaria diagnosis, skin cancer detection, and breast cancer classification in few studies. CNNs have been used in the context of diagnosing malaria, where images of blood cells are classified as parasitized or uninfected with significant success. However, CNNs are very powerful, but training deep neural networks require a lot of labeled data, computational resources, and time. This requirement poses an additional challenge with regard to practical implementation, particularly in low-resource

---

[*] Corresponding author

settings where computational capability and large-scale labelled data sets may not be available (Balasubramaniam *et al*., 2023).

However, deep learning architectures such as CNNs do have their shortcomings. A major problem lies in the computational expense of training and deploying such models. While designing a CNN from scratch to solve unique data problems, is often time-consuming and hardware-intensive (GPUs and TPUs), making it a challenge in resource-limited settings. More importantly, the implementation of these complex models in real-time diagnosis systems can be a challenging task due to speed and resource cost. Also, deep learning models typically demand a lot of annotated training data, which can be quite challenging to obtain in the healthcare domain, given privacy concerns, expenses, and the necessity for expert labels (Salehi *et al*., 2023).

DL methods may involve complex architectures, which can limit their performance on small datasets. Hybrid models can work for leveraging the benefits of CNNs in feature extraction and overcoming computational complexity due to deep learning (by utilizing simpler and cheaper classifiers in the final stages)(Wen *et al*., 2023)Here, CNN is used to extract more higher-level representative features from image inputs and pass them through classical machine learning classifiers such as Random Forest, SVM, KNN, Logistic Regression, etc. Our hybrid method achieves high diagnostic accuracy while employing less processing power than other CNN-based systems, making it well-suited to real-time clinical applications.

Most recent work has adopted using CNNs for feature extraction combined with a traditional machine learning classifier. Because of this, CNNs can learn representations of edges, textures, and higher-order relationships that are almost impossible to hand-engineer on images, from the pixel values of raw images themselves. After extracting those features, typical classifiers like Random Forest or SVM, which are less time-consuming and computationally expensive compared to deep neural networks, can be applied to classify the images based on their features (Sothe *et al*., 2020).

A key advantage of this hybrid method is the ability to train and deploy the model with much lower computational costs. The CNN is only employed as a feature extractor rather than going through an entire CNN process with deep classification layers, which reduces the amount of time spent training and applying the model to learn new parameters. The subsequent conventional classifiers can also be trained faster, using the CNN-extracted features, and require less memory space. Also, since CNN in this case acts as a feature extractor, the model can still leverage transfer learning with pre-trained CNNs such as VGG16, ResNet or Inception, pre-trained on large datasets like ImageNet. Even if the dataset is small for the specific domain (for instance, the domain of malaria diagnosis), the CNN still draws on features that were learned from general-purpose image classification tasks (Das *et al*., 2021).

(Rosnelly *et al*., 2023) Compared CNN and SVM classifiers for the classification of malaria parasites. An algorithm was used to extract features from images; that is, PEMA and KEHE algorithms were used to extract characteristics such as eccentricity and area. (Sonawane & Telang, 2020) made use of Bins Approach Algorithm and other feature extraction methods like color moments, texture features (GLCM). In this study, after preprocessing the blood smear images, SVM, RF and KNN based machine learning algorithms were used for classification. For malaria detection from blood smear images, a hybrid approach of CNN and SVM was used by (Vijayalakshmi, 2020). The satisfactory classification accuracy was obtained when VGG19 architecture was combined with SVM as even it was found to meet the threshold criteria thus it holds importance.

Deep learning techniques for automatic malaria detection have been further explored in recent work. Research conducted by (Ramos-Briceño *et al*., 2025) utilized a CNN derived model for identification of parasitized and uninfected cells with a remarkable accuracy, using images of thick blood smear to generalise the model in both P. falciparum and P. vivax species. Their model focused on the superior performance of deep architectures for medical imaging problems. (Ahamed *et al*., 2025) study introduced a group of tailored CNN structures (PCNN, SPCNN and SFPCNN), which were found to enhance malaria parasite classification. The SPCNN model, especially, improved robustness and interpretability by incorporating the spatial feature extracting network into the convolution process. These recent works validate the applicability of CNN-based methods to malaria diagnosis and also demonstrate the capabilities of architecture adaptation in improving the accuracy of diagnosis which closely relates with the motivations of our proposed hybrid model.

In this paper, we propose a hybrid scheme using the VGG16 CNN for feature extraction, and traditional multi-class classifiers for classification. In this paper, we specifically try five classifiers, namely: Random Forest, Logistic Regression, K-Nearest Neighbors, Support Vector Machine, and Gradient Boosting. The accuracy, precision, recall and F1-score metrics are the most common metrics used to quantify classifier performance in a classification problem, and therefore are reported for each one. We seek to explore whether this hybrid functional approach can provide excellent diagnostic performance while being computationally tractable for real-world applications in low-resource environments.

The remainder of this paper is organized as follows: In Section 2 we review related work in the area of automated malaria detection and hybrid machine learning models. In Section 3, we explain the methodology, covering the dataset, preprocessing steps, CNN-based feature extraction, and classifiers applied. In Section 4, we show the experiment results and discuss the performance of the hybrid model using various classifiers. Section 5 wraps up the paper by proposing future work.

2. METHODOLOGY

The Methodology includes several phases beginning from Dataset and Image Acquisition, Data Preprocessing where the images are resized and normalized. In Hybrid Model Design, CNN (VGG16) extracts features, which classified using traditional machine learning algorithms like Random Forest, Logistic Regression, SVM. Finally, Performance Evaluation calculates models using accuracy, precision, recall and f1 score for an efficient and accurate approach towards malaria diagnosis. As shown in figure 1.

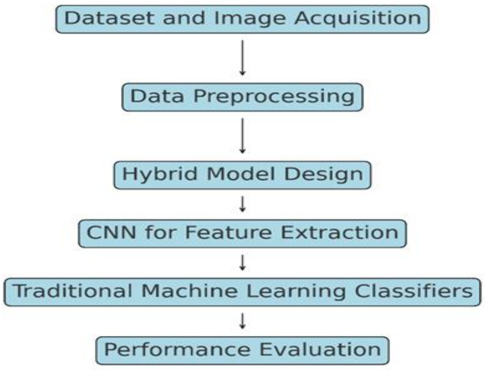

Figure 1: Data flow of the proposed model

Dataset and Image Acquisition:

The dataset used in this study is the NIH Malaria Dataset, which is publicly available through the TensorFlow Datasets (TFDS) repository. This dataset is widely considered a standard benchmark for malaria image classification tasks because it contains a large number of labeled, high-quality blood smear images, is balanced between infected and uninfected samples, and is captured under realistic microscopy settings, reflecting practical diagnostic challenges. While the dataset is balanced between classes, it primarily focuses on a specific imaging protocol and may not cover the full diversity of real-world variations such as differences in microscope brands, slide preparation techniques, or patient demographics across regions. This dataset containing 27,558 blood smear slide images. The dataset consists of two classes: parasitized (infected) and uninfected of the cells. The data is appropriately labeled, and the dataset is moderately balanced with almost the same amount of parasitized and uninfected images. Images are obtained under a regular light microscope and they have three color channels (RGB) and pixels dimensions vary from 100x100 pixels to 200x200 pixels.

It is one of the image-based malaria diagnosis datasets, and hence a good choice for our research. And because it is available via TFDS, it can be used in the TensorFlow framework for simple data processing, training and evaluation.

Data Preprocessing:

Processing image data is an important step of ensuring a successful machine learning pipeline, and that process takes preprocessing of the data. Since the images in malaria dataset do not have the same dimensions, we standardized the input size to be 224×224 pixels by resizing all images to this size. This dimension was selected, as it is a common input size for CNNs, e.g., VGG16 used for feature extraction. It helps in normalizing the input size, and also decreases computational cost which helps the model to concentrate on the essential parts of the image.

Image Normalization:

After resizing, we normalized the pixel values of all the images. Normalization is transforming the pixel value originally in range [0, 255] into a new range [0, 1]. This is an important step since, in general, CNNs benefit from having input data scaled to a smaller normalized range. So Normalized images helps in keeping the process of learning stable and enables the model to converge faster during training Without normalization, it can lead to worse performance or divergence due to high variance in the pixel values while making gradient updates.

Hybrid Model Design:

We use two components from different empirical paradigms of machine learning: Deep Learning (in the form of CNNs for feature extraction and traditional machine learning classifiers for classification. This utilizes the best of both worlds: CNN helps to extract the features from the images while a traditional classifier does the classification of the extracted features.

CNN for Feature Extraction:

In our pipeline, the CNN will serve the primary purpose of extracting features. Given the effectiveness of CNNs for automatically learning hierarchical feature representations directly from raw pixel data, they have emerged as the gold standard for image classification problems (Litjens *et al*., 2017; Shen *et al*., 2017). To extract features, we used pretrained VGG16 networks, and this model was one of the well-known CNN architectures, which has been pretrained on the ImageNet dataset that is a comprehensive dataset of around millions of labeled images across thousands of categories.

VGG16 Architecture: VGG16 has 16 layers, 13 convolutional and 3 fully connected. It uses a simple, yet effective, design pattern where a number of small 3×3 filters from a number of convolutional layers are stacked on top of each other followed by a max-pooling layer to reduce the size of the spatial dimension of the image. This architecture enables the network to learn progressively intricate features as information propagates through the layers, ranging from basic edge types in the lower layers to advanced object features in the higher layers.

Specifically, we will be using the VGG16 model's top classification layers (a.k.a the "head" of the model) in our hybrid approach. With this change, the network is only being used as a feature extractor and is not directly making classification. Concretely, we only removed the last fully connected layers and added a Global Average Pooling (GAP) layer. The GAP layer finds the average of each feature map, reducing the dimensionality of the output feature maps and providing a fixed length feature vector for each input image.

Moreover, VGG16 was chosen due to its elegant structure that has proven to be effective and widely accepted in the domain of computer vision. VGG16 is pretrained on ImageNet, so it has already learned a lot of features that are able to recognize objects. One observation is that while malaria is a specialized domain, their features transfer to our malaria classification task. Therefore, by employing transfer learning technique, we used the Pre-trained weights of VGG16 to Extract higher order features from malaria dataset instead of training the network from scratch. This allows for a large reduction in the computational cost and time for training.

Traditional Machine Learning Classifiers:

After the CNN extracts features from images, classical machine learning classifiers are used to perform the actual classification task. We tried out five non-neural network classifiers: Random Forest, Logistic Regression, KNN, SVM, and Gradient Boosting. Table 1 summarizes the key hyperparameters used for each classifier in our experiments to enhance reproducibility.

Random Forest: Random Forest is combined multiple decision trees to obtain a final prediction Individual trees are trained separately on random subsets of the data, and the overall prediction is determined by the majority vote among all trees. Random Forests are well-known for their robustness and handling high-dimensional feature spaces, which made it a good candidate for our feature-rich data.

Logistic Regression: Logistic Regression is a simple yet powerful linear classifier commonly used for binary classification tasks. Logistic Regression is a classification algorithm that estimates the probability of a binary outcome based on one or more predictor variables. Logistic Regression is very simple, but works good on separable data which could be the case with features extracted from CNN.

KNN: The KNN algorithm is a non-parametric method that classifies the image based on the majority class of the k-nearest neighbors in the feature space. Despite its simplicity and interpretability, KNN's performance might be sensitive to the value of k and the distance metric employed. Since prediction includes distance computations from all points in the training set, KNN can be slow on larger datasets.

SVM: SVM is a great classifier, works by choosing the hyperplane that best classifies the data by two. It finds the hyperplane that best separates the data points on either side and maximizes the distance between the support vectors (the data points that are closest to the hyperplane), resulting in a very strong decision boundary. SVM works well in high-dimensional space; therefore, the CNN extracted features would be best fitted to be classified with SVM.

Gradient Boosting Classifier: Gradient Boosting is yet another ensemble learning method that builds a series of models, each of which tries to rectify the mistakes made by the previous models. This results in a very accurate model (while possibly more computationally expensive than other classifiers) through iterating and updating. Gradient Boosting works very well for complex decision boundaries and also reduces overfitting.

Table 1 : Summarizes the key hyperparameters used for each classifier in our experiments.

Classifier	**Main Hyperparameters**
Random Forest	Number of trees = 100, Max depth = None
Logistic Regression	Regularization = L2, Solver = 'liblinear'
K-Nearest Neighbors	k = 5, Distance metric = Euclidean
Support Vector Machine	Kernel = RBF, C = 1.0
Gradient Boosting	Number of estimators = 100, Learning rate = 0.1

Traditional classifiers were selected because they offer efficient training and inference when working with fixed feature representations extracted from CNNs, allowing for high accuracy with reduced computational requirements, which is critical for deployment in resource-limited environments.

Performance Evaluation:

We then split the dataset and evaluate train-test performance of the classifiers based on scientific literature guidance, using 80% of the data for training and 20% for testing. We used the following evaluation metrics to see how each of the classifiers performed:

Accuracy: The number of correctly predicted instances divided by the total number of instances.

F1 Score: This is harmonic mean of Precision and Recall, hence a better statistic when the classes are imbalanced.

Precision: True positives of the predicted positives.

Recall: True positives / Actual positives

The CNN features extracted were used to train each classifier and to evaluate on the test set. They represent good measures of the classifiers performance, by describing both the accuracy along with the ratio between false positive and false negative cases.

Experimental Results:

Using five of the most popular classifiers — Random Forest, Logistic Regression, KNN, SVM and Gradient Boosting — we evaluated the hybrid method of extracting CNN-based features and passing them to a traditional machine learning classifier for malaria classification. This section shows the results and explains it on the basis of four performance metrics, namely, accuracy, F1 score, precision, and recall. These metrics provide an overall assessment of the success of the models at classifying malaria-infected and uninfected cells. All experiments were conducted using Python, TensorFlow, and Scikit-learn, and Keras on a standard workstation equipped with an Intel Core i7 CPU, 16 GB RAM, and no dedicated GPU. As shown in figure 2.

Figure 2: Performance Metrics of Classifiers

Random Forest Classifier:

The Random Forest classifier yielded an accuracy of 91.09%, indicative of a good classification performance but still with potential for enhancement. We also had an F1 score of 91.09%, which is an indication that we have a good balance between precision and recall. For precision, it was 91.17% and recall 91.09%. This means Random Forest had a slight advantage in making positive predictions correct, such that there is a small gap between precision and recall. The Random Forest has the advantage of being an ensemble of decision trees and taking the average of their predictions, so it is able to generalize well to unseen data, yet the model can be tuned further / trained on a larger dataset.

Its relatively high precision suggests that the model is good at avoiding false positives (e.g. not classifying uninfected cells as infected). If the recall indicates this is likely to be the case, it means that the model may sometimes miss some infected cells, which has obvious implications when considering how this might be used in the wild; if someone has malaria this kind of false negative could be life threatening. While it gets a good accuracy, we could get even higher accuracy with a little tweaking such as hyperparameter tuning (increasing no of trees, adjusting trees' depth, etc), or feeding the models with more data.

Logistic Regression Classifier:

Logistic Regression classifier is considered good for preserving interpretability in the model process, and despite being a simple technique, it outperformed Random Forest with an accuracy of 93.27% having proven robust for this type of binary classification problem, such as malaria detection. The F1 score stood at 93.26%, with 93.39% precision and 93.27% recall. Logistic Regression yielded good performance with reasonably well-balanced precision and recall, indicating success both in correctly identifying malaria-infected cells and ensuring that the cells predicted to be infected are largely so.

As Logistic Regression is a generalized linear model and assumed to perform a little worse than a more complex model such Random Forest and/or Gradient Boosting. Yet, its simplicity, interpretability and reasonably strong performance when used in some diagnostic systems can make it a competitive choice, especially in cases where faster prediction times and lower computational expenses are necessary. The high precision score for Logistic Regression indicates that it is quite accurate at avoiding false positives, which is crucial when utilizing it for healthcare purposes. But Logistic Regression is a linear model that may not perform well on more complex and non-linear decision boundaries.

K-Nearest Neighbors Classifier:

The K-Nearest Neighbors (KNN) classifier achieved the lowest score of the classifiers with 83.85% accuracy. The F1 score is 83.69% with a precision of 85.22% and recall of 83.85%, which reflects this relatively low accuracy. KNN outperformed the other algorithms because of the simplicity and sensitivity of the structure of the data. While more complex models learn complex decision boundaries or template ensemble methods, KNN only needs to learn the distance to nearest neighbors. The complexity of the malaria dataset and overlapping regions between parasitized and uninfected cells, which KNN may struggle with, in this case, could cause KNN to be less helpful than other approaches.

KNN had a higher precision than recall, which suggests it was more careful than accurate in its positive predictions, missing positive cases. This means KNN may classify malaria-positive cells as uninfected more often than we would like, resulting in a relatively high number of false negatives. In the case of malaria detection, this becomes particularly problematic as the goal is to detect all positive cases. Because of the low performance, KNN would not be a suitable choice for this classification problem in a real-world diagnostic context.

Support Vector Machine Classifier:

The SVM classifier has a validation accuracy of 90.71%, comparable to Random Forest and logistic regression. The F1 score was 90.69%, precision was 91.04% and recall was 90.71%. SVM's performance reflects on its ability to handle high-dimensional feature space effectively as it finds the optimal hyperplane which separates the parasitized cells from the uninfected cells. Both precision and recall were reasonably good, indicating this model was very effective in minimizing the number of false positives as well as false negatives leading to a reliable model for malaria detection.

This method works well when the classes in the training set are not linearly separable and it can map the data to higher dimensions using kernel trick. It performs competitively in this experiment and thus can serve as a strong candidate for image classification tasks. But SVM has also high computational cost which makes it use in less suitable that is not real-time application in resource-constrained environments.

Gradient Boosting Classifier:

Lastly, The Gradient Boosting classifier recorded an accuracy of 92.22%, putting it, performance-wise, between Logistic Regression and SVM. The F1 score was 92.21%, precision — 92.32%, recall — 92.22%. This ability of Gradient Boosting to emphasize the mistakes of the preceding models to incrementally enhance its predictions is very powerful and suits complex tasks extremely well. It is close to the best-performing models, and its balanced precision and recall suggests it is good at minimizing both false positives and false negatives.

Gradient Boosting — Gradient Boosting builds models in a sequential manner, making it very good at generating very accurate predictions. However, Gradient Boosting is also computationally intensive, similar to SVM, mainly with a large number of trees as well as boosting rounds. Although it is more accurate than the previous methods, its complexity may restrict its use in many real-time applications unless there are a lot of computational resources available.

3. DISCUSSION

To summarize, Logistic Regression and Gradient Boosting were the winners in terms of accuracy, with Logistic Regression slightly better than Gradient Boosting. Nonetheless, all classifiers except the KNN classifier produced strong metrics, boasting sensitivity levels of over 90%, indicating their potential for use in malaria classification. The low accuracies of KNN indicate that it is not suitable for complex datasets like the malaria dataset. The other classifiers displayed a more balanced approach between precision and recall; thus, they can be alternative options for the implementation of malaria classification in diagnostic systems. Improvements could be made from tuning hyper parameters and adding more data augumentation techniques to further avoid overfitting and improve generalization. As we employed a pretrained VGG16 model solely for feature extraction and combined it with traditional machine learning classifiers, the computational demands were significantly lower than training deep CNNs end-to-end. A more detailed evaluation of inference speed and memory consumption will be considered in future work to further validate the approach for deployment in resource-limited environments.

In many low-resource settings, access to high-performance computing infrastructure, such as GPUs or large memory servers, is limited. Training and deploying deep CNN models from scratch often requires significant computational resources and annotated datasets, which are not readily available. By using a pretrained VGG16 model solely for feature extraction and combining it with lightweight traditional classifiers, our hybrid approach reduces the computational burden substantially. This makes it more feasible for deployment on standard CPUs or mobile devices, aligning with the practical constraints typically encountered in rural and resource-limited environments.

CONCLUSION

This work proposes a hybrid approach for understanding the use of CNN based feature extraction in conjunction with traditional machine learning classifiers for the purpose of malaria diagnosis. The combination of VGG16 and various classification algorithms (like Logistic Regression and Gradient Boosting) allows for a balance between accuracy and computational efficiency in the proposed model. Out of all the classifiers that were evaluated, Logistic Regression was able to produce the best overall performance, boasting an accuracy of over 93.27% for this binary classification problem, and hence it is a strong and efficient choice for detecting malaria.

In this hybrid approach, we aim at a practical solution of its application in more resource-constrained settings by minimizing the computational intensity of deep learning architectures. Moreover, this approach strikes an adequate balance between the feature representation ability of CNNs and the interpretability and simplicity of classical classifiers.

Potential future works include tuning of hyperparameters, experiment with different CNN architectures for feature extraction, or advanced data augmentation techniques to increase the generalizability of the model. This work lays the groundwork for further development of this approach for use in the field for rapid and near-elimination of malaria or other medical diagnoses.

Acknowledgments:

The authors thank the staff of Computer Information System Department, Technical College of Zakho, Duhok Polytechnic University.

Ethical Approval:

The research protocol was evaluated and approved by the Research Ethics Committee at the College of Medicine, University of Zakho (UOZ), adhering to established ethical guidelines (Approval Code: UOZE521; Year: 2024).

Author Contributions

W.A.A., B.N.M., A.K.I, and O.M.A., contributed equally to the conception, implementation, analysis, and writing of this manuscript.

Funding:

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

Abd Alnabi, D. L., Ahmed, S. S., & Abd Alnabi, N. L. (2025). knee osteoarthritis stage classification based on hybrid fusion deep learning framework. science *journal of university of zakho,* 13(2), 262-278. https://doi.org/10.25271/sjuoz.2025.13.2.1450

Ahamed, M. F., Nahiduzzaman, M., Mahmud, G., Shafi, F. B., Ayari, M. A., Khandakar, A., Abdullah-Al-Wadud, M., & Islam, S. M. R. (2025). Improving Malaria diagnosis through interpretable customized CNNs architectures. *Scientific Reports*, *15*(1), 6484. https://doi.org/10.1038/s41598-025-90851-1

Al-Awadhi, M., Ahmad, S., & Iqbal, J. (2021). Current status and the epidemiology of malaria in the Middle East Region and beyond. *Microorganisms*, *9*(2), 338. https://doi.org/10.3390/microorganisms9020338

Awosolu, O. B., Yahaya, Z. S., Farah Haziqah, M. T., & Olusi, T. A. (2022). Performance evaluation of nested polymerase chain reaction (Nested PCR), light microscopy, and plasmodium falciparum histidine-rich protein 2 rapid diagnostic test (PfHRP2 RDT) in the detection of falciparum malaria in a high-transmission setting in Southwestern Nigeria. *Pathogens*, *11*(11), 1312. https://doi.org/10.3390/pathogens11111312

Balasubramaniam, S., Velmurugan, Y., Jaganathan, D., & Dhanasekaran, S. (2023). A modified LeNet CNN for breast cancer diagnosis in ultrasound images. *Diagnostics*, *13*(17), 2746. https://doi.org/10.3390/diagnostics13172746

Das, R., Kumari, K., De, S., Manjhi, P. K., & Thepade, S. (2021). Hybrid descriptor definition for content based image classification using fusion of handcrafted features to convolutional neural network features. *International Journal of Information Technology*, *13*(4), 1365–1374. https://doi.org/10.1007/s41870-021-00722-x

Gao, Z., Lu, Z., Wang, J., Ying, S., & Shi, J. (2022). A convolutional neural network and graph convolutional network-based framework for classification of breast histopathological images. *IEEE Journal of Biomedical and Health Informatics*, *26*(7), 3163–3173. https://doi.org/10.1109/JBHI.2022.3153671

Grignaffini, F., Simeoni, P., Alisi, A., & Frezza, F. (2024). Computer-Aided Diagnosis Systems for Automatic Malaria Parasite Detection and Classification: A Systematic Review. *Electronics*, *13*(16), 3174. https://doi.org/10.3390/electronics13163174

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., Van Der Laak, J. A., Van Ginneken, B., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, *42*, 60–88. https://doi.org/10.1016/j.media.2017.07.005

Ramos-Briceño, D. A., Flammia-D'Aleo, A., Fernández-López, G., Carrión-Nessi, F. S., & Forero-Peña, D. A. (2025). Deep learning-based malaria parasite detection: convolutional neural networks model for accurate species identification of Plasmodium falciparum and Plasmodium vivax. *Scientific Reports*, *15*(1), 3746. https://doi.org/10.1038/s41598-025-87979-5

Rosnelly, R., Riza, B. S., & Suparni, S. (2023). *Comparative Analysis of Support Vector Machine and Convolutional Neural Network for Malaria Parasite Classification and Feature Extraction*. https://doi.org/10.58346/JOWUA.2023.I3.015

Salehi, A. W., Khan, S., Gupta, G., Alabduallah, B. I., Almjally, A., Alsolai, H., Siddiqui, T., & Mellit, A. (2023). A study of CNN and transfer learning in medical imaging: Advantages, challenges, future scope. *Sustainability*, *15*(7), 5930. https://doi.org/10.3390/su15075930

Shen, D., Wu, G., & Suk, H.-I. (2017). Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, *19*(1), 221–248. https://doi.org/10.1146/annurev-bioeng-071516-044442