

GEOMETRIC FEATURES: A CRITICAL COMPONENT FOR ACCURATE MALARIA PARASITE STAGE CLASSIFICATION IN THIN SMEAR MICROSCOPY

Gunjan Aggarwal ¹✉, Mayank Kumar Goyal ²✉

¹ School of Engineering & Technology, Sharda University, India

² School of Engineering & Technology, Sharda University, India



Corresponding Author

Gunjan Aggarwal,
2019000123.gunjan@dr.sharda.ac.in

DOI

[10.29121/shodhkosh.v5.i5.2024.4332](https://doi.org/10.29121/shodhkosh.v5.i5.2024.4332)

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

Copyright: © 2024 The Author(s). This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

With the license CC-BY, authors retain the copyright, allowing anyone to download, reuse, re-print, modify, distribute, and/or copy their contribution. The work must be properly attributed to its author.



ABSTRACT

Malaria-a global health problem always demands an accurate and timely diagnosis of a disease for its proper treatment. Traditional methods like microscopic examination are time-consuming and require specialized expertise. It thus poses challenges in resource-limited areas. Automated classification of malaria parasite stages helps in improving the diagnostic efficiency. In this paper, the importance of geometric features in malaria parasite stage classification using machine learning techniques has been realized. Geometric features, including area, perimeter, and shape descriptors, offer valuable information regarding the morphological differences between the various stages of the parasite. We compare the performance of the following machine learning models using geometric features: Random Forest, GaussianNB, XGBoost, and MLPClassifier. The results show that the inclusion of geometric features improves the accuracy and robustness of the machine learning models for classification. Among the different models tested in this study, MLP Classifier had 95.90% accuracy thus shows tremendous potential for a geometric feature in a malaria diagnosis program. This current study, therefore, gives way to advancement in automated diagnosis of malaria among others and further pursuit of geometric-based applications in their fields.

Keywords: Geometric Features, Malaria Diagnosis, Random Forest Classifier, XGBoost, MLP Classifier, Multiclass Classification

1. INTRODUCTION

Malaria is a life-threatening disease that is caused by Plasmodium parasites and transmitted through the bites of infected Anopheles mosquitoes (World Health Organization, n.d.). The species of Plasmodium comprises four species, namely Plasmodium falciparum, Plasmodium vivax, Plasmodium ovale and Plasmodium malariae. While in human's host each species separately progresses in four stages namely ring, trophozoite, schizont and gametocyte. Among these species, the most common causes of severe complications are Plasmodium falciparum and Plasmodium vivax (Heide et al., 2019). Stage-specific identification of malaria parasites is very crucial to establish the correct treatment regimen. Conventional approaches involve the visual examination of blood smears using a microscope, which is cumbersome and error-prone (Talapko et al., 2019). Machine learning-based automated diagnostic devices are becoming an exciting new approach, offering more accuracy and speed.

In the recent years, several features from microscopic images, such as geometric, color and texture attributes, have been used to detect malaria and classify its stage. Table 1 gives an overview of different types of features used in the detection of the malaria parasite's stage, with their descriptions and references. Geometric features, including area and perimeter, are often used since they can be used to represent morphological differences between the stages. Texture features comprise GLCM and LBP that indicate surface pattern features, while color features measure the intensity distributions to classify into the respective stage. Morphological features, which are the aspect ratio and convexity, help well in the structural analysis. The pixel intensity variations are measured as statistical features. The spatial frequency patterns can be determined using features based on Fourier and wavelet transforms in the frequency domain. Nuclear-cytoplasmic ratio and similar other features related to the biological context have great significance for parasites. Hybrid features combine multiple feature types to enhance classification accuracy, thus showing the potential of combining diverse feature sets in automated malaria diagnosis.

Geometric features become very important characteristics in distinguishing between different species and life stages of malaria parasites (A. S. Nugroho et al., 2020). Geometric features may include area, perimeter, eccentricity, and circularity, which capture the morphological features across lifecycle stages. These features capture information that supports pixel-based features in image-based models. This paper elaborates on the significance of geometric features in the malaria parasite stage classification problem and highlights their use in comparing various machine learning techniques.

Table 1 Comparing different feature types used in malaria parasite stage detection

Feature Type	Description	Advantages	Challenges
Geometric Features(H. A. Nugroho et al., 2015)	Metrics such as area, perimeter, circularity, aspect ratio, elongation, etc.	Simple, interpretable, and strongly correlated with parasite morphology.	May not capture complex textures or color variations.
Color Features(Poostchi et al., 2018)	Mean intensity, color histograms, and pixel intensity distributions.	Effective for distinguishing parasites with unique staining patterns.	Sensitive to variations in staining quality and lighting conditions.
Texture Features(Tek et al., 2010)	GLCM, LBP, and wavelet transform features to capture surface patterns.	Captures fine details and differences in parasite stages effectively.	Computationally intensive and sensitive to noise in the image.
Morphological Features(Loddo et al., 2018)	Shape descriptors like convexity, solidity, and skeletonization.	Useful for describing the structural changes in parasites across stages.	Limited in capturing internal details like texture or color.
Statistical Features(Bashar, 2019)	Histogram of oriented gradients (HOG), intensity moments including mean, variance and skewness	Provides robust global statistical measures for classification.	May lose spatial information due to global averaging.
Frequency Domain Features(Poostchi et al., 2018)	Fourier transform and DCT to analyze spatial frequency patterns.	Captures patterns not visible in the spatial domain.	Requires pre-processing and is computationally demanding.
Biological Features(Tek et al., 2010)	Nuclear-cytoplasmic ratio, inclusion bodies, and other biological markers.	Directly relevant to malaria biology, aiding in species-specific detection.	Requires domain-specific knowledge and manual annotations.
Hybrid Features(Bashar, 2019)	Combines geometric, texture, and color features for enhanced performance.	Leverages strengths of multiple feature types for improved accuracy.	Increased complexity and potential redundancy in feature space.

Geometric features, especially area and perimeter, are particularly effective in malaria detection and classification due to their ability to capture essential morphological differences among parasite species and their developmental stages. For example, the ring stage of *Plasmodium falciparum* usually occupies less area compared to the trophozoite

stage. With this feature analysis, automatic systems are capable of accurate detection and classification of parasite life cycle stages.

The primary aim of this study is to propose an automated and efficient approach to classify malaria parasite stages using the integration of advanced image preprocessing techniques with robust machine learning. This study aims to:

- This paper introduces a novel approach to malaria parasite stage classification by emphasizing the critical role of geometric features, such as area, perimeter, and shape descriptors, in capturing morphological variations across different stages.
- The study evaluates the performance of multiple machine learning models, including Random Forest, GaussianNB, XGBoost, and MLPClassifier, demonstrating that the integration of geometric features significantly enhances classification accuracy and robustness.

The MLPClassifier achieved the highest accuracy of 95.90%, underscoring the effectiveness of these features in automated malaria diagnosis. This work provides a foundation for leveraging geometric features to advance diagnostic efficiency, particularly in resource-limited settings.

This paper is organised into sections as follow: section 1 will introduce malaria parasite stages and explain goal of this paper. In section 2, the existing literature for malaria detection is explored. Section 3 will detail the machine learning techniques used to identify and classify the different malaria parasite stages. Section 4 will detail the experiment and evaluation of various classification model. Section 5 will summarize the result and discuss future scope

2. LITERATURE SURVEY

The literature survey highlights diverse approaches to automated malaria diagnosis, including geometric, color, and texture features, advanced algorithms, and machine learning classifiers. These methods clearly reflect the substantial progress made in accuracy and efficiency in multi-stage malaria parasite classification.

The innovative morpho-geometrical approach by authors A. S. Nugroho et al. (A. S. Nugroho et al., 2020) for automated malaria diagnosis integrates Naive Bayes classification that differentiates malaria parasite species and life stages. The research accomplished commendable accuracy metrics with an F1 score of 80.60% using computational geometry combined with more complex algorithms such as recursive bottleneck detection and Otsu's thresholding.

Abbas et al. (Abbas & Dijkstra, 2020a) in 2022 utilized area and perimeter measurements, as well as other features, to classify *Plasmodium falciparum* stages using random forest classifiers with high accuracy in stage differentiation.

Thaqifah Ahmad Aris et al. (Aris et al., 2023) another research attempt that has built a strong image processing framework using the geometric features like area and perimeter to classify the intelligent multi-stage malaria parasite. Improved detection performance has been reported and thus the framework depicts high accuracy along with adaptability in different smear types and staining techniques. It shows the feature integration for the enhancement of malaria diagnosis.

Authors Jhonathan Sora-Cardenas et al. (Sora-Cardenas et al., 2025) focuses on the detection and classification of malaria parasites using color features in the RGB and HSV color spaces. The proposed method analyzes the staining patterns and pixel intensities in microscopic images. The results show the effectiveness of color features to differentiate between infected and uninfected cells. This approach increases the accuracy of diagnosis and minimizes the need for manual microscopy.

In study authors Abdurahman et al. (Abdurahman et al., 2021) comprehensively investigates using color features derived from images of thick smears for the detection of malaria parasites. It uses machine learning classifiers in an attempt to automate the diagnosis process. The study highlights the potential of color-based features in addressing the challenges of thick smear microscopy. It uses machine learning classifiers in an attempt to automate the diagnosis process

Study by Akcakir et al. (Akcakir et al., 2022) presents an approach combining geometric, morphologic, and texture features with deep learning techniques to detect malaria parasites. The system trains models for high diagnostic precision using labelled segmented cells. The hybrid feature set demonstrates potential in automating malaria diagnosis at various stages, opening up pathways toward scalable and efficient diagnostic solutions.

The literature survey focuses on the developments in malaria diagnosis, specifically highlighting geometric, color, and texture features for parasite detection and stage classification. Techniques such as Naive Bayes, Random Forest, and hybrid frameworks achieve commendable accuracy, underlining the importance of feature integration in automating diagnostics.

3. RESEARCH METHODOLOGY

The Research methodology outlines a systematic approach of this study, to classify malaria parasites. As illustrated in figure 1, the methodology can be split into three essential components: Preprocessing Techniques, Contour Detection, Feature Extraction and Machine Learning Techniques. Preprocessing techniques used in the image enhancement phase involved enhancing the microscopic images for detecting the malaria parasites accurately. Contour detection was then applied to the image for highlighting the outlines of parasites, thus enabling the derivation of meaningful geometric features like area and perimeter.

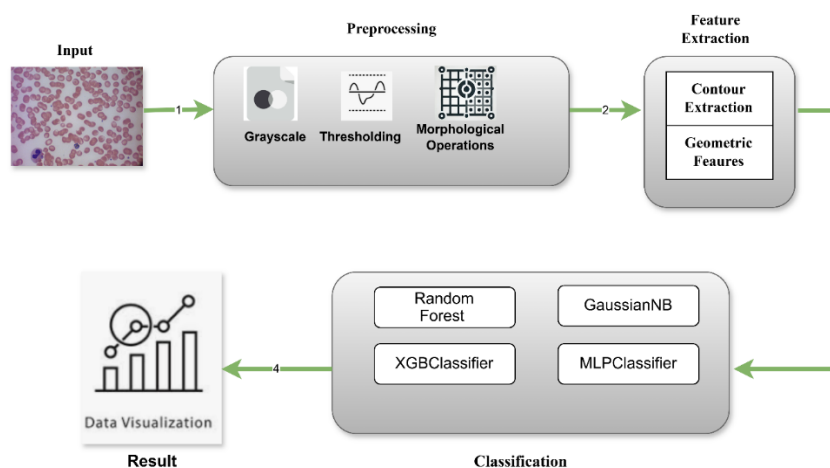


Figure 1 Classification of Malaria Parasite Stages

The extracted features were weighted and combined to identify the morphological characteristics of the parasites, and were subsequently used as input features for various machine learning algorithms, like Naive Bayes, XGBoost, Random Forest Classifier, and MLP, to determine the malaria parasite stages. It is worth noting that, the geometric features enhancement provides robust and accurate classification for all the algorithms. Therefore, the framework incorporates image processing with ML techniques, so that malaria stage diagnosis becomes more reliable and efficient.

3.1. PREPROCESSING TECHNIQUES AND FEATURE EXTRACTION

It has been observed that, preprocessing and feature extraction techniques are important for malaria parasite images for classification. In order to ensure the extraction of relevant features for classifying the malaria parasite stages correctly, the image quality is enhanced using the preprocessing pipeline.

Preprocessing Techniques

The steps used for preprocessing the image steps are discussed below:

Grayscale Conversion: The input images were converted from RGB to grayscale to reduce the amount of data and focus on intensity variations, which are important for the detection of morphological features.

Thresholding: To separate foreground (parasite-infected cells) from the background, Otsu's thresholding was used, which adaptive method automatically determined the optimal threshold value, thus providing robust segmentation against varying image conditions.

Morphological Transformations: Perform morphological operation using an opened or closed by 3×3 structural elements. These morphology transformations will enhance noise removal on the segmented images, which were obtained to effectively extract contours.

Contour Detection and Feature Extraction

Contours which identify the boundaries of connected components in the segmented image. The extracted contours formed a basis for calculating geometric features. It includes area, which is the count of pixels inside a contour, representing the size of the infected cell. The perimeter can be defined as the total length of the contour boundary, reflecting the cell's shape. In combined features, the area and perimeter have been weighted average computed to highlight prominent contours while at the same time reducing the influence of noise.

The combined area $A_{combined}$ and perimeter $P_{combined}$ were calculated using the following formulas:

$$A_{combined} = \frac{\sum(A_i \times A_i)}{\sum(A_i)} \quad (1)$$

$$P_{combined} = \frac{\sum(P_i \times A_i)}{\sum(A_i)} \quad (2)$$

where A_i and P_i are the area and perimeter of the i th contour respectively.

The preprocessing pipeline with feature extraction was applied to images from the MP-IDB dataset (Loddo et al., 2019), encompassing multiple parasite stages. The extracted features were stored in a structured format, and were fed as input to different machine learning classifiers. This approach effectively captured the subtle differences between stages, thereby improving classification accuracy.

3.2. MACHINE LEARNING TECHNIQUES

The weighted combination of area and perimeter features extracted from the images is used to train different machine learning algorithms like Naive Bayes, XGBoost Classifier, Random Forest classifier etc. for malaria parasite classification.

3.2.1. NAIVE BAYES

GaussianNB is a machine learning algorithm for classification tasks. GaussianNB (Pushpakumar et al., 2022) assumes features (combined area and perimeter) follow a normal distribution (bell-shaped curve) for each class (parasite stage). During training, it estimates the mean and variance of these distributions for each stage. When presented with a new image's features, it calculates the probability of that image belonging to each stage based on how well its features fit the corresponding stage's distribution. The stage with the highest probability becomes the predicted class. It's computationally efficient and requires minimal parameter tuning. The model provides insights into the distribution of features for each stage, aiding understanding of the classification process. It can perform well on problems with continuous features like the ones extracted here, especially when the data is well-separated between classes. However, it has some limitations to consider. It assumes feature independence, which might not always hold true (combined area and perimeter might be correlated). It can't directly handle categorical features (e.g., image filename). Techniques like label encoding are needed for such data. The training data consists of feature vectors (combined area, combined perimeter) and corresponding class labels (malaria parasite stages). For each class (e.g., Ring, Trophozoite, Schizont), the model estimates the mean and variance of each feature (combined area and perimeter). This is done by assuming that the features follow a Gaussian distribution. The model calculates the probability density function (PDF) (Virk et al., 2012) for each feature given the estimated mean and variance for each class. The features (combined area, combined perimeter) are extracted from a new, unseen image. The model calculates the probability of the new data point belonging to each class using the learned parameters and the PDF. The class with the highest probability is assigned to the new data point. Overall, GaussianNB (Soni et al., 2011) offers a good balance between simplicity, interpretability, and effectiveness for this specific task of classifying malaria parasite stages based on image-derived features.

3.2.2. XGBOOST CLASSIFIER

XGBoost, or Extreme Gradient Boosting (Hou et al., 2021), is a very efficient and scalable implementation of gradient boosting. It was designed to be faster and more efficient than any other available predictive modelling. It is widely used

for classification problems on big datasets with a high number of dimensions and is therefore also used for medical image analysis. It performs model building in a sequentially manner where each new model focusing on the residuals (errors) of the previous models and optimize the loss function on gradient descent. XGBoost applies L1 and L2 regularization to avoid overfitting, which is the major problem in most machine learning models. In contrast to other traditional boosting methods, XGBoost (Chen & Guestrin, 2016) allows parallel computation when creating trees, making training faster. XGBoost employs a depth-first strategy for tree pruning, which offers better performance as it doesn't have to build the entire tree before pruning, unlike some boosting methods. In this study, XGBoost classifies images using geometric features obtained from the contours of parasites. The features incorporate useful morphological information like area and perimeter combined. After applying several preprocessing techniques and feature extraction (contour-based geometric features), data can be fed into the XGBoost model for training. XGBoost has some hyperparameters (e.g., learning rate, max depth, number of estimators) that can be tuned to enhance the model's performance. Performance metrics can be used to determine the efficiency of the model like accuracy, precision, recall, F1-score.

3.2.3. RANDOM FOREST CLASSIFICATION

Random Forest (Tin Kam Ho, 1995) is a very powerful and efficient tool for malaria parasite stage classification. Its ensemble nature along with the capability to handle high-dimensional data makes it very suitable for image classification tasks. Random Forest (RF) is a machine learning technique that combines multiple decision trees to make predictions (Abbas & Dijkstra, 2020b). During training, it creates several decision trees and, for classification tasks, predicts the class with the majority vote. For regression tasks, it calculates the average of the predictions from all the trees. Random Forest is particularly more suitable for handling huge sets of data with increased dimensions like stages classification of malaria parasites using microscopic images. Random Forest assigns feature importance, where a researcher can understand which feature significantly contributes to the prediction from the model. Random Forest is friendly to use, with little parameter tuning hence, it is mostly often

the preferred choice for the practitioner. The primary metrics drawn out in the feature extraction process are combined area and combined perimeter of contours detected in the images. These are collected in a structured DataFrame such that each row corresponds to an image and carries out its respective stage label. Further, the data is divided into training and testing sets, using a 70-30 split ratio to ensure evaluation on unseen data. The model is trained on the training data, and predictions are made on the test set, followed by an assessment of the model's performance with accuracy, classification report and confusion matrix. The accuracy score is the measure of the model's overall performance, and the classification report gives details about precision, recall, and F1-score for each class. The confusion matrix visualizes the true positive and negative predictions and provides a deeper understanding of the model's classification capabilities.

3.2.4. MULTI-LAYER PERCEPTRON (MLP) CLASSIFIER

Multi-Layer Perceptron (MLP) classifier (Zare et al., 2013) is an artificial neural network, the most widely used and implemented machine learning algorithm for regression or classification. It comprises an input layer, one or more hidden layers of interconnected neurons, and an output layer. This makes the MLP classifier particularly useful for data with non-linearities, such as complicated applications like classifying malaria parasite stages. In the layers of an MLP, neurons process input through a linear transformation followed by a nonlinear activation function. An MLP is trained using backpropagation, which is a supervised learning technique. To enhance learning efficiency and accelerate training convergence, the Adam optimization technique will be used.

Table 2: Configuration of the Multi-Layer Perceptron (MLP) Classifier for Malaria Parasite Stage Classification

Parameter	Configuration
Input Features	Geometric features (e.g., area, perimeter)
Hidden Layers	2 layers (128 neurons, 64 neurons)
Activation Function	ReLU
Optimizer	Adam
Loss Function	Cross-Entropy Loss

Batch Size	Default (auto-tuned by MLPClassifier)
Learning Rate	Adaptive (Adam default)
Maximum Iterations	500
Random State	42
Evaluation Metrics	Accuracy, Precision, Recall, F1-Score

In this work, geometric parameters like area, perimeter, and shape descriptors are applied as MLP classifier input to distinguish the stages of malaria parasites. It will be possible for the MLP classifier to detect subtle distinctions between parasites on account of these features, effectively reflecting their morphological variations between stages. The `train_test_split` function is used to split the dataset into training (80%) and testing (20%) groups using a stratified split. As shown in Table 2, the MLP classifier is set with two hidden layers of sizes 128 and 64. Due to its ability to mimic nonlinear patterns and because it is computationally efficient, the ReLU activation function was used. The model was iterated to a maximum of 500 iterations. The training process involves learning the relationships between the geometric features and the corresponding parasite stages using the training dataset (X_{train}, y_{train}). After training, the classifier predicts the parasite stages for the test dataset (X_{test}). Model performance is evaluated using the `accuracy_score` function to compute overall accuracy and the `classification_report` provides detailed metrics, including precision, recall, and F1-score for each stage.

4. EXPERIMENTAL RESULTS

In this Experiment, the malaria parasite stage classification was analyzed based on the performance of various machine learning algorithms. Different performance metrics such as accuracy, weighted average precision, weighted average recall, and weighted average F1-score are used in the study. Models employed were Random Forest, GaussianNB, XGBoost Classifier, and MLP Classifier.

Table 3 Comparative Results of Machine Learning and Deep Learning Models for Malaria Parasite Stage Classification

Model	Accuracy (%)	Weighted Precision (%)	Average	Weighted Recall (%)	Average	Weighted	Average	F1-Score (%)
Random Forest	95.38		94		95			95
GaussianNB	93.85		94		94			94
XGBoost	94.62		94		95			94
MLPClassifier	95.90		92		96			94

As shown in Table 3, the results indicate that the MLP Classifier resulted in the maximum accuracy of 95.90% with the weighted recall value of 96%, which made it robust to classify malaria parasite stages.

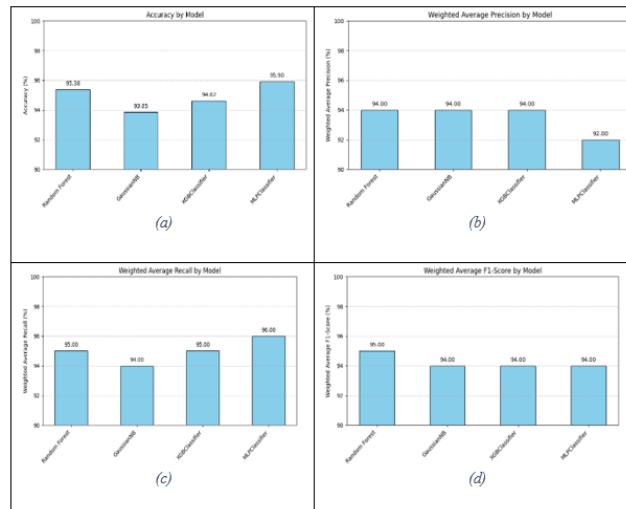


Figure 2 Graphical Representation of Model Performance Metrics(a) Accuracy(b) Precision(c) Recall(d) F1 score

It was very effective with low false positives and false negatives, and for this reason the classifier is a good candidate for the problem. The Random Forest classifier was the closest with accuracy of 95.38% and 95% of weighted F1. It ranked the second-best performer overall.

The XGBoost Classifier achieved competitive outcomes as it had an accuracy of 94.62% and a weighted recall of 95%, thus generalizing well across the dataset. On the other hand, the GaussianNB classifier has the lowest performance, with an accuracy of 93.85% and the lowest weighted metrics.

Traditional machine learning classifiers, such as Random Forest and XGBoost, showed competitive performance by using geometric features like area and perimeter, which captured morphological variations across different parasite stages. However, these models were not able to exploit the intricate patterns present in the dataset. GaussianNB, being simple and computationally efficient, failed to capture the complex relationships between features and performed relatively poor.

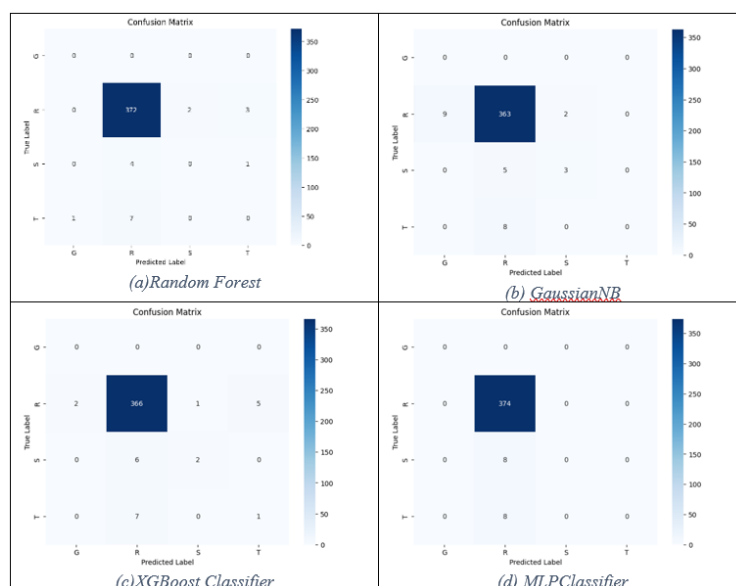


Figure 3 Confusion matrix showing the classification model's performance(a) Random Forest (b) GaussianNB (c) XGBoost (d)MLP Classifier

Figure 3 further explains the performance of MLPClassifier by using confusion matrix. Confusion matrix is very helpful in analysing the classification accuracy of the model at different stages of malaria parasites. Confusion matrix has the true positive, false positive, false negative, and true negative prediction at each stage that allows a very granular level of performance analysis of the classifier.

Overall, as seen in figure 2, the MLP Classifier was the best model for malaria parasite stage classification, followed by Random Forest and XGBoost Classifier. The results show that machine learning models, combined with carefully selected geometric features, have the potential to improve the accuracy and efficiency of malaria diagnosis. Future work could explore additional geometric or texture-based features to further enhance the performance of these classifiers.

5. CONCLUSION AND FUTURE SCOPE

This study demonstrates the critical role of geometric features in malaria parasite stage classification by machine learning techniques. Geometric features capture the morphological characteristics of parasites and enhance the performance of classification models with reliable and efficient automated malaria diagnosis. All the tested models were compared and MLP Classifier appears to be the most effective with an accuracy of 95.90%.

Although the outcome is promising, but the approach should improve its handling of imbalanced data between various stages of malaria. Techniques of data augmentation using rotation, flipping, and scaling are effective on classes underrepresented to boost model performance for all the classes. This approach could further be expanded with incorporation of both geometric and pixel-based features that could make it a superior technique to make better and

accurate classification results with more reliable diagnoses for multi-species malaria. The development of robust and scalable diagnostic tools will significantly have a high impact on the control and treatment of malaria globally.

CONFLICT OF INTERESTS

None.

ACKNOWLEDGMENTS

None.

REFERENCES

- Abbas, S. S., & Dijkstra, T. M. H. (2020a). Detection and stage classification of *Plasmodium falciparum* from images of Giemsa stained thin blood films using random forest classifiers. *Diagnostic Pathology*, 15(1), 130. <https://doi.org/10.1186/s13000-020-01040-9>
- Abbas, S. S., & Dijkstra, T. M. H. (2020b). Detection and stage classification of *Plasmodium falciparum* from images of Giemsa stained thin blood films using random forest classifiers. *Diagnostic Pathology*, 15(1), 130. <https://doi.org/10.1186/s13000-020-01040-9>
- Abdurahman, F., Fante, K. A., & Aliy, M. (2021). Malaria parasite detection in thick blood smear microscopic images using modified YOLOV3 and YOLOV4 models. *BMC Bioinformatics*, 22(1), 112. <https://doi.org/10.1186/s12859-021-04036-4>
- Akcakır, O., Celebi, L. K., Kamil, M., & Aly, A. S. I. (2022). Automated wide-field malaria parasite infection detection using Fourier ptychography on stain-free thin-smears. *Biomedical Optics Express*, 13(7), 3904. <https://doi.org/10.1364/BOE.448099>
- Aris, T., Nasir, A., Mustafa, W., Mashor, M., Haryanto, E., & Mohamed, Z. (2023). Robust Image Processing Framework for Intelligent Multi-Stage Malaria Parasite Recognition of Thick and Thin Smear Images. *Diagnostics*, 13(3), 511. <https://doi.org/10.3390/diagnostics13030511>
- Bashar, Md. K. (2019). Improved Classification of Malaria Parasite Stages with Support Vector Machine Using Combined Color and Texture Features. 2019 IEEE Healthcare Innovations and Point of Care Technologies, (HI-POCT), 135–138. <https://doi.org/10.1109/HI-POCT45284.2019.8962686>
- Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. <https://doi.org/10.1145/2939672.2939785>
- Heide, J., Vaughan, K. C., Sette, A., Jacobs, T., & Schulze Zur Wiesch, J. (2019). Comprehensive Review of Human *Plasmodium falciparum*-Specific CD8+ T Cell Epitopes. *Frontiers in Immunology*, 10, 397. <https://doi.org/10.3389/fimmu.2019.00397>
- Hou, Y., Li, Q., Zhang, C., Lu, G., Ye, Z., Chen, Y., Wang, L., & Cao, D. (2021). The State-of-the-Art Review on Applications of Intrusive Sensing, Image Processing Techniques, and Machine Learning Methods in Pavement Monitoring and Analysis. *Engineering*, 7(6), 845–856. <https://doi.org/10.1016/j.eng.2020.07.030>
- Loddo, A., Di Ruberto, C., & Kocher, M. (2018). Recent Advances of Malaria Parasites Detection Systems Based on Mathematical Morphology. *Sensors*, 18(2), 513. <https://doi.org/10.3390/s18020513>
- Loddo, A., Di Ruberto, C., Kocher, M., & Prod'Hom, G. (2019). MP-IDB: The Malaria Parasite Image Database for Image Processing and Analysis. In N. Lepore, J. Brieva, E. Romero, D. Racoceanu, & L. Joskowicz (Eds.), *Processing and Analysis of Biomedical Information* (Vol. 11379, pp. 57–65). Springer International Publishing. https://doi.org/10.1007/978-3-030-13835-6_7
- Nugroho, A. S., Winarta, T., Wibisono, Y., Galinium, M., Rozi, I. E., & Asih, P. B. S. (2020). Morpho-Geometrical Feature Extraction of Thin Blood Smear Microphotograph for Malaria Plasmodia Species and Life Stage Determination. 2020 International Conference on Advanced Computer Science and Information Systems (ICACSIS), 95–100. <https://doi.org/10.1109/ICACSIS51025.2020.9263220>
- Nugroho, H. A., Akbar, S. A., & Murhandarwati, E. E. H. (2015). Feature extraction and classification for detection malaria parasites in thin blood smear. 2015 2nd International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE), 197–201. <https://doi.org/10.1109/ICITACEE.2015.7437798>

- Poostchi, M., Silamut, K., Maude, R. J., Jaeger, S., & Thoma, G. (2018). Image analysis and machine learning for detecting malaria. *Translational Research*, 194, 36–55. <https://doi.org/10.1016/j.trsl.2017.12.004>
- Pushpakumar, R., Prabu, R., Priscilla, M., Renisha, P. S., Prabu, R. T., & Muthuraman, U. (2022). A Novel Approach to Identify Dynamic Deficiency in Cell using Gaussian NB Classifier. 2022 7th International Conference on Communication and Electronics Systems (ICCES), 31–37. <https://doi.org/10.1109/ICCES54183.2022.9835813>
- Soni, J., Ansari, U., Sharma, D., & Soni, S. (2011). Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction. *International Journal of Computer Applications*, 17(8), 43–48. <https://doi.org/10.5120/2237-2860>
- Sora-Cardenas, J., Fong-Amaris, W. M., Salazar-Centeno, C. A., Castañeda, A., Martínez-Bernal, O. D., Suárez, D. R., & Martínez, C. (2025). Image-Based Detection and Classification of Malaria Parasites and Leukocytes with Quality Assessment of Romanowsky-Stained Blood Smears. *Sensors*, 25(2), 390. <https://doi.org/10.3390/s25020390>
- Talapko, J., Škrlec, I., Alebić, T., Jukić, M., & Včev, A. (2019). Malaria: The Past and the Present. *Microorganisms*, 7(6), 179. <https://doi.org/10.3390/microorganisms7060179>
- Tek, F. B., Dempster, A. G., & Kale, İ. (2010). Parasite detection and identification for automated thin blood film malaria diagnosis. *Computer Vision and Image Understanding*, 114(1), 21–32. <https://doi.org/10.1016/j.cviu.2009.08.003>
- Tin Kam Ho. (1995). Random decision forests. *Proceedings of 3rd International Conference on Document Analysis and Recognition*, 1, 278–282. <https://doi.org/10.1109/ICDAR.1995.598994>
- Virk, A. S., Kaur, M., & Passrija, L. (2012). Performance Evaluation of Image Enhancement Techniques in Spatial and Wavelet Domains. *INTERNATIONAL JOURNAL OF COMPUTERS & TECHNOLOGY*, 3(1), 162–166. <https://doi.org/10.24297/ijct.v3i1c.2771>
- World Health Organization. (n.d.). Compendium of WHO malaria guidance: Prevention, diagnosis, treatment, surveillance and elimination Thumbnail V. World Health Organization. <https://iris.who.int/handle/10665/312082>
- Zare, M., Pourghasemi, H. R., Vafakhah, M., & Pradhan, B. (2013). Landslide susceptibility mapping at Vaz Watershed (Iran) using an artificial neural network model: A comparison between multilayer perceptron (MLP) and radial basic function (RBF) algorithms. *Arabian Journal of Geosciences*, 6(8), 2873–2888. <https://doi.org/10.1007/s12517-012-0610-x>