

Malaria Detection Using Support Vector Machine (SVM) Based on Microscopic Blood Cell Images

Azzah Aqila^{1*}, Naishatanha Agatha Balqis², Chantika Febrin Hardiani³, Muhammad Yunus⁴

¹ Health Information Management Study Program, Department of Health, Politeknik Negeri Jember, Indonesia; g41240037@student.polije.ac.id

² Health Information Management Study Program, Department of Health, Politeknik Negeri Jember, Indonesia; g41240059@student.polije.ac.id

³ Health Information Management Study Program, Department of Health, Politeknik Negeri Jember, Indonesia; g41240233@student.polije.ac.id

⁴ Health Information Management Study Program, Department of Health, Politeknik Negeri Jember, Indonesia; m.yunus@polije.ac.id

Correspondence: g41240037@student.polije.ac.id

Abstract: Malaria continues to be a serious infectious disease that significantly affects healthcare systems in tropical and resource-limited regions. Conventional malaria diagnosis is commonly performed through microscopic examination of blood cell images using a microscope; however, this process is time-consuming and highly dependent on the expertise of medical personnel. This study proposes an automated malaria image classification framework using the Support Vector Machine (SVM) algorithm combined with image feature extraction techniques. The experimental dataset was collected from Kaggle and contained 4,000 microscopic blood cell images categorized into infected and uninfected classes. Several preprocessing procedures were applied, including image resizing, grayscale transformation, and pixel normalization to improve feature consistency. Furthermore, feature extraction was performed using the Histogram of Oriented Gradients (HOG) method. The dataset was separated into training and testing subsets using an 80:20 distribution scheme. An SVM classifier with a Radial Basis Function (RBF) kernel was implemented and evaluated using confusion matrix analysis and standard classification metrics. Performance evaluation demonstrated that the proposed model achieved 78.87% accuracy, accompanied by balanced precision, recall, and F1-score values. The obtained findings indicate that the proposed HOG-SVM framework provides stable classification capability with efficient computational performance for automated malaria detection tasks.

Keywords: Malaria Detection; Support Vector Machine; Machine Learning; Image Classification; Microscopic Images

Citation: A. Aqila, N. A. Balqis, C. F. Hardiani, and M. Yunus, "Malaria Detection Using Support Vector Machine (SVM) Based on Microscopic Blood Cell Images", *infynex*, vol. 1, no. 1, pp. 22–32, Jun. 2026.

Received: 27-05-2026

Accepted: 20-06-2026

Published: 30-06-2026



Copyright: © 2026 by the authors. This article is distributed under the terms and conditions of the Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

1. Introduction

Malaria is a parasitic disease spread through bites of infected *Anopheles* mosquitoes and continues to affect many tropical regions worldwide. [1]. Based on reports from the World Health Organization (WHO), malaria continues to be a significant public health challenge, particularly in tropical and developing countries where millions of cases are reported each year [1]. Therefore, a fast and accurate diagnostic method is required to support early malaria detection and treatment [2].

Malaria diagnosis is commonly performed through microscopic examination of blood smear preparations using a microscope [2]. This method remains the standard diagnostic approach because it can directly identify the presence of parasites in red blood cells [2]. Conventional microscopic diagnosis is often time-consuming and strongly influenced by the experience of laboratory personnel, which may affect diagnostic consistency. [3]. In addition, variations in microscopic image quality may lead to diagnostic errors [3].

Recent advancements in artificial intelligence and digital image analysis have contributed to the development of automated disease detection systems [2], [6]. Support Vector Machine (SVM) is one of the commonly applied machine learning algorithms for image classification tasks. [4]. Support Vector Machine (SVM) has been widely applied in classification problems because of its capability to separate complex data patterns efficiently. [4], [5]. Furthermore, SVM has relatively lower computational complexity compared to deep learning methods [6].

Previous research has demonstrated that deep learning approaches, including Convolutional Neural Networks (CNN), can produce high accuracy in medical image classification [7]. However, CNN models generally require larger computational resources and longer training times [6], [10]. Therefore, SVM remains relevant as a simpler yet effective alternative for image classification tasks [8], [11].

This research focuses on applying the Support Vector Machine (SVM) algorithm to classify malaria using microscopic blood cell images. The dataset used was obtained from Kaggle and consisted of two classes, namely Parasitized and Uninfected. The research procedure involved multiple stages such as image preprocessing, feature extraction, SVM training, and performance assessment of the proposed model. [4], [7].

2. Materials and Methods

This research applied a machine learning-based framework for automated malaria classification using microscopic blood cell images [3], [9]. The research stages included dataset collection, image preprocessing, feature extraction, Support Vector Machine (SVM) model training, and model performance evaluation [3], [4].

2.1 Research Flowchart

At this stage, a flowchart was used to illustrate the overall process of malaria detection using the Support Vector Machine (SVM) method, starting from dataset collection until model evaluation.

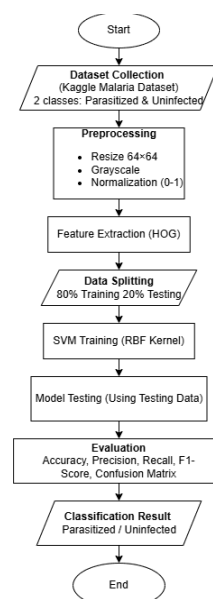


Figure 1. Research Flowchart of Malaria Detection Using the SVM Method

Figure 1 presents the complete workflow of the proposed malaria classification system based on the Support Vector Machine (SVM) algorithm. The workflow started with dataset acquisition from Kaggle, consisting of microscopic blood cell images categorized as Parasitized and Uninfected. The images then undergo preprocessing stages, including resizing, grayscale conversion, and normalization. Feature extraction was then conducted using the Histogram of Oriented Gradients (HOG) descriptor to obtain representative image features. The processed dataset was divided into training and testing subsets using an 80:20 ratio. An SVM classifier utilizing the Radial Basis Function (RBF) kernel was subsequently trained and evaluated using the prepared dataset. Model performance was finally assessed using accuracy, precision, recall, F1-score, and confusion matrix analysis to measure classification effectiveness.

2.2 Dataset

This study used the Malaria Cell Images Dataset obtained from Kaggle and processed using Google Colab. The dataset consisted of two microscopic blood cell image classes, namely Parasitized and Uninfected, which were utilized for malaria image classification experiments. A total of 4,000 microscopic blood cell images were included in this study, consisting of 2,000 infected cell images and 2,000 uninfected cell images.

The dataset was separated into training and testing subsets using the train-test split technique with an 80:20 ratio. Approximately 80% of the dataset was allocated for model training, whereas the remaining 20% was reserved for classification performance evaluation. Prior to the classification stage, all images underwent preprocessing and feature extraction procedures using the Histogram of Oriented Gradients (HOG) descriptor. The balanced distribution between infected and uninfected image classes was intended to minimize classification bias and improve the model's capability to distinguish malaria-positive and normal blood cell images.

2.3 Preprocessing

Image preprocessing was performed to standardize the dataset before the classification stage. In this stage, all images were resized to 64×64 pixels to standardize image dimensions and reduce computational complexity during the model training process. The RGB images were transformed into grayscale format to reduce image complexity and improve computational efficiency.

Next, normalization was performed by dividing each pixel value by 255 so that the pixel values ranged from 0 to 1. After normalization, the images were used for the Histogram of Oriented Gradients (HOG) was utilized to obtain important visual features from the blood cell images. This preprocessing stage aimed to generate more consistent data, accelerate computational processing, and improve the model's ability to recognize patterns in infected and uninfected blood cell images.

2.4 Feature Extraction (Ekstraksi Fitur)

Following the preprocessing stage, feature extraction was conducted using the Histogram of Oriented Gradients (HOG) descriptor. The HOG descriptor was utilized to capture texture patterns, structural characteristics, and gradient orientation information from microscopic blood cell images. The extracted feature vectors were subsequently utilized as input data for the Support Vector Machine

(SVM) classification model. The use of HOG aimed to improve the model's ability to distinguish between Parasitized and Uninfected blood cell images based on visual characteristics contained in the microscopic images.

Following the preprocessing stage, feature extraction was conducted using the Histogram of Oriented Gradients (HOG) descriptor. The extracted HOG features were utilized to represent structural patterns, edge distribution, and visual characteristics contained in microscopic blood cell images.

2.5 Support Vector Machine (SVM)

This study applied the Support Vector Machine (SVM) algorithm as the primary classification method. SVM is a machine learning technique that classifies data by determining an optimal separating boundary between different classes. The algorithm is recognized for its capability to process high-dimensional data and produce reliable classification results even with limited datasets.

The Radial Basis Function (RBF) kernel was selected in this research due to its effectiveness in managing non-linear data distributions. The classification model was implemented in Python using the Scikit-learn framework. The extracted image features were utilized as input data during both training and testing stages.

Model development was performed using training data, while testing data were used to evaluate classification performance. The implementation of SVM in this study was intended to achieve efficient computational performance while maintaining stable classification accuracy.

3. Results and Discussions

The experimental evaluation showed that the SVM model was capable of classifying malaria blood cell images with stable performance. Model testing was performed using data that had previously been divided into training and testing sets with an 80:20 proportion. Several evaluation metrics, including accuracy, precision, recall, and F1-score, were applied to measure the effectiveness of the proposed model.

The proposed model obtained 78.87% accuracy, with precision, recall, and F1-score values of 79%, 80%, and 79%, respectively. These findings demonstrate that the model was able to differentiate infected and non-infected blood cell images effectively. The precision score reflects the model's ability to minimize false positive classifications during prediction, while the recall value reflects the model's ability to identify malaria images in the testing dataset.

According to the confusion matrix analysis, the model successfully identified 318 Parasitized images and 313 Uninfected images correctly. Despite the satisfactory performance, several prediction errors were still observed during classification, where 82 Parasitized images were predicted as Uninfected and 87 Uninfected images were predicted as Parasitized. Overall, the SVM approach demonstrated consistent performance in malaria image classification using microscopic blood cell data.

3.1 Evaluation Metrics

Model performance was assessed using several statistical evaluation metrics, including accuracy, precision, recall, and F1-score. These evaluation indicators were applied to analyze the classification capability of the Support Vector Machine (SVM) model in identifying malaria blood cell images. Accuracy represents the proportion of

correctly classified samples among all testing data, precision indicates the reliability of positive prediction results, recall reflects the model's capability to correctly recognize positive cases, whereas the F1-score describes the balance between precision and recall performance. The evaluation formulas used in this study are presented as follows:

- Accuracy = $(TP + TN) / (TP + TN + FP + FN)$ (1)
- Precision = $TP / (TP + FP)$ (2)
- Recall = $TP / (TP + FN)$ (3)
- F1-Score = $2 \times (Precision \times Recall) / (Precision + Recall)$ (4)

True Positive (TP) refers to malaria-infected images that were correctly predicted as malaria cases, whereas True Negative (TN) represents normal blood cell images correctly classified as non-malaria. False Positive (FP) occurs when normal blood cell images are incorrectly predicted as malaria cases, while False Negative (FN) occurs when malaria-infected images are incorrectly classified as normal.

Accuracy was utilized to evaluate the general correctness of the classification model. Precision was applied to analyze the reliability of positive prediction outcomes. Recall was utilized to assess the capability of the model in recognizing positive samples, whereas the F1-score was used to evaluate the balance between precision and recall values. These evaluation metrics enabled a more comprehensive interpretation of the classification performance results.

3.2 Confusion Matrix Results

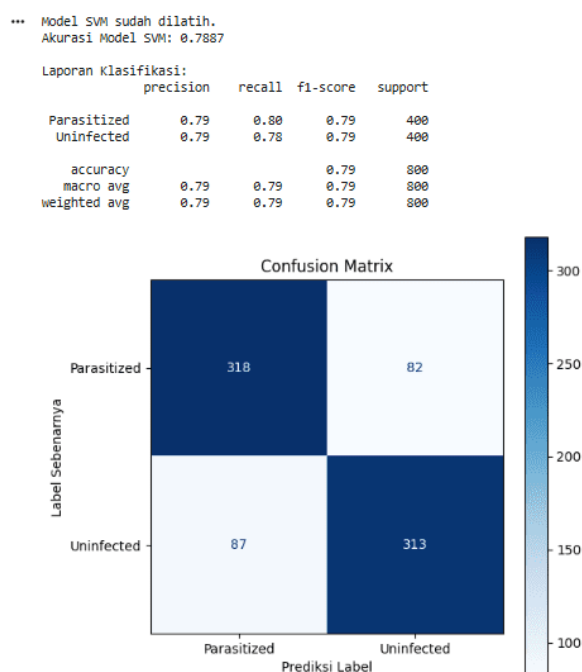


Figure 2. Confusion Matrix of SVM Model for Malaria Detection

Figure 2 illustrates the confusion matrix output generated by the Support Vector Machine (SVM) classification model for malaria blood cell image detection. The experimental evaluation showed that the proposed model achieved an accuracy value of 78.87%, demonstrating stable classification capability for both Parasitized and Uninfected classes. The classification model successfully identified 318 Parasitized images and 313 Uninfected images correctly.

Several classification errors were still observed, where 82 Parasitized images were classified as Uninfected and 87 Uninfected images were predicted as Parasitized. These classification inaccuracies were likely influenced by similarities in texture patterns, cell structure, and image intensity among several blood cell samples, which reduced the model's capability to differentiate both categories effectively.

Overall, the confusion matrix analysis demonstrates that the SVM method provided stable classification performance for automated malaria detection.

3.3 Evaluation Results

Table 1. Evaluation Results of SVM Model

Evaluation Metric	Value
Accuracy	78,87%
Precision	79%
Recall	80%
F1-Score	79%
Support	800

Table 1 summarizes the evaluation performance of the Support Vector Machine (SVM) classification model for malaria image detection. The experimental findings showed that the proposed model achieved 78.87% accuracy, 79% precision, 80% recall, and 79% F1-score during the testing process. These performance results demonstrate that the classification model was capable of distinguishing malaria blood cell images with relatively stable performance. The obtained precision score indicates that the model was capable of minimizing false positive classifications, whereas the recall result demonstrates the model's effectiveness in recognizing malaria-infected blood cell images. In addition, the F1-score value reflects a balanced relationship between precision and recall performance during classification. Overall, the SVM-based classification approach demonstrated effective performance in differentiating Parasitized and Uninfected blood cell images.

3.4 System Implementation

The malaria classification system was developed in Google Colaboratory using the Support Vector Machine (SVM) algorithm combined with Histogram of Oriented Gradients (HOG) feature extraction. Blood cell image datasets were processed through preprocessing, feature extraction, and model training stages to classify parasitized and uninfected blood cells. The trained model was then implemented into a Python-based application using Visual Studio Code and the Streamlit framework, allowing users to upload microscopic blood cell images and obtain automatic malaria classification results through the system interface.

- Parasitized

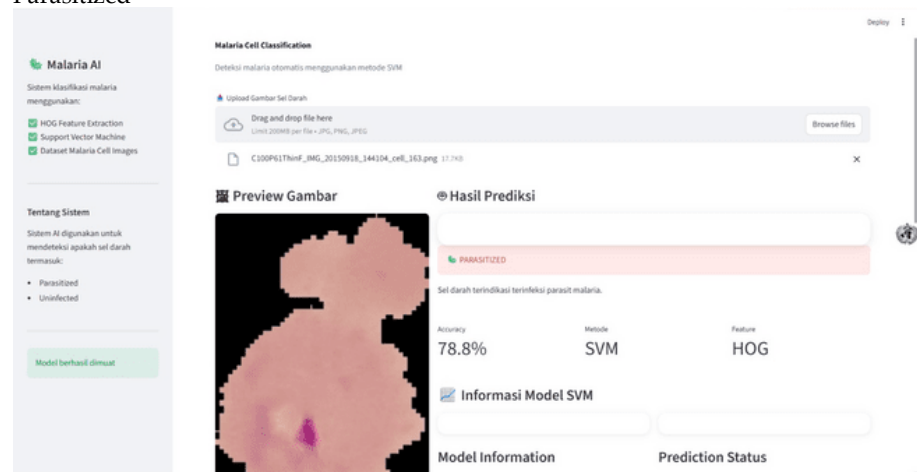


Figure 3. Interface of Malaria Cell Classification Using the SVM Method

Figure 3 illustrates the main interface of the malaria blood cell classification system using the Support Vector Machine (SVM) method. The application enables users to upload microscopic blood cell images and automatically classify them using HOG-based feature extraction. The prediction displayed in the figure belongs to the *Parasitized* category, meaning that the blood cell is identified as infected with malaria parasites. In addition, the system displays model information, including the model accuracy, classification approach, and features applied during detection.

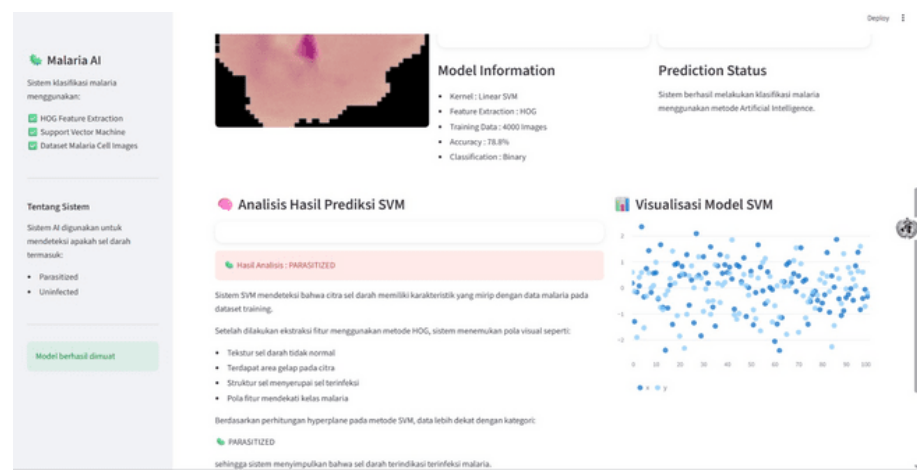


Figure 4. Analysis Result of Parasitized Blood Cell Classification

Figure 4 presents the analysis results of a blood cell image classified as *Parasitized*. The system identifies visual patterns that are similar to malaria-infected blood cell images contained in the training dataset. Using Histogram of Oriented Gradients (HOG) feature extraction, the system detects patterns such as abnormal cell texture, dark regions in the image, and cell structures resembling malaria-infected blood cells. In addition, the visualization of the SVM model is displayed in the form of a data distribution graph representing the classification results.

- Uninfected

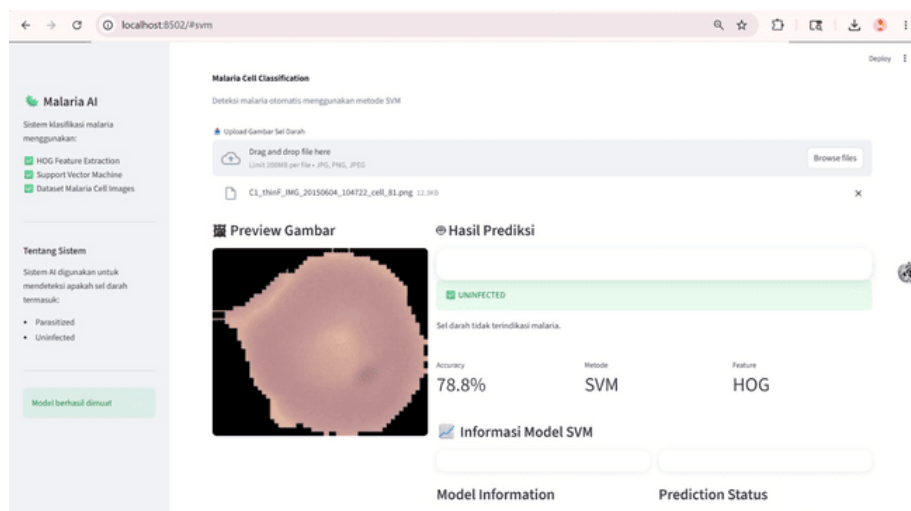


Figure 5. Interface of Uninfected Blood Cell Classification

Figure 5 shows the classification result of a normal blood cell image using the SVM method. The model classifies the uploaded image as *Uninfected*, indicating that the blood cell is not infected with malaria parasites. The interface displays the uploaded image preview, prediction result, model accuracy value, SVM method, and HOG features used during the classification process. These findings show that the system can differentiate normal blood cells from malaria-infected cells effectively.

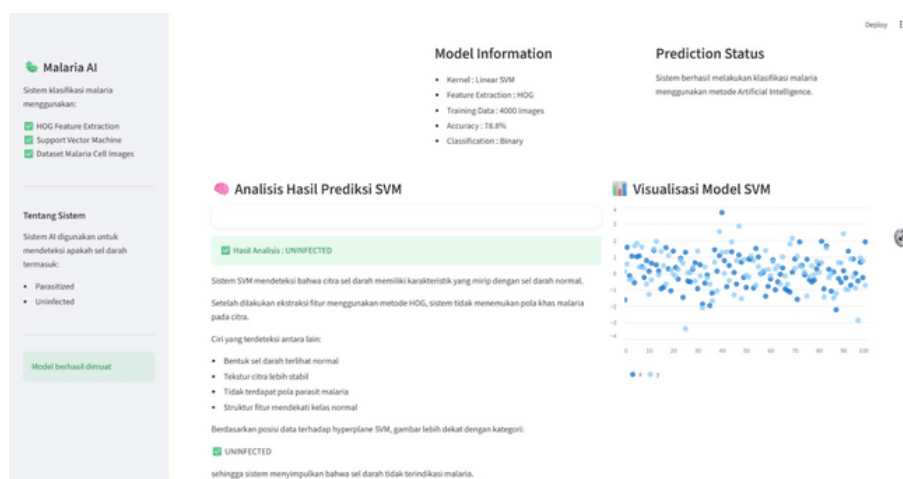


Figure 6. Analysis Result of Uninfected Blood Cell Classification

Figure 6 presents the analysis results of a blood cell image classified as *Uninfected*. The system identifies image characteristics that resemble normal blood cells, such as stable cell structure, normal image texture, and the absence of typical malaria parasite patterns. Based on the hyperplane calculation of the SVM method, the image data are classified closer to the normal category. In addition, the SVM model visualization is displayed to illustrate the distribution of the classification results visually.

The findings of this study demonstrate that the Support Vector Machine (SVM) algorithm achieved stable classification performance for malaria detection using microscopic blood cell images. The proposed classification model produced 78.87% accuracy, 79% precision, 80% recall, and 79% F1-score during the experimental evaluation process. The obtained results indicate that the SVM algorithm was able to

differentiate Parasitized and Uninfected blood cell images with stable classification capability.

The classification capability achieved in this study was affected by several preprocessing procedures performed before model training. Image resizing, grayscale transformation, and normalization procedures contributed to reducing data complexity and improving feature consistency during classification. Furthermore, the implementation of the Radial Basis Function (RBF) kernel allowed the model to process non-linear data distributions more efficiently, resulting in improved separation between classification categories [3], [4].

The experimental findings also indicate that the SVM algorithm still provides stable performance in medical image classification tasks, particularly for malaria detection based on microscopic blood cell images. Several previous studies also reported that the SVM algorithm achieved consistent results in medical data classification applications [4], [5]. Furthermore, the implementation of artificial intelligence approaches in healthcare systems is considered capable of supporting faster and more efficient disease diagnosis procedures [2], [6], [10].

Based on the confusion matrix analysis, the model correctly identified 318 Parasitized images and 313 Uninfected images. Nevertheless, several classification errors were still observed, with 82 Parasitized images predicted as Uninfected and 87 Uninfected images predicted as Parasitized. These misclassification cases were likely influenced by similarities in texture patterns, cell morphology, and image intensity among several blood cell samples, which reduced the model's capability to distinguish both image categories accurately. In addition, differences in microscopic image quality and illumination conditions may also have influenced the HOG-based feature extraction process.

Although deep learning approaches such as Convolutional Neural Networks (CNN) generally produce higher classification accuracy in medical image classification tasks, they require greater computational resources and longer training times [7], [10]. Therefore, the implementation of the Support Vector Machine (SVM) algorithm in this study remains an effective alternative approach due to its lower computational complexity and relatively efficient processing while still providing stable classification performance [8], [11].

This research still contains several limitations, "including the application of a relatively simple feature extraction technique and the absence of extensive parameter optimization procedures. Therefore, future studies may be improved by applying more advanced feature extraction techniques, performing parameter optimization using GridSearchCV, or applying deep learning approaches such as Convolutional Neural Networks (CNN) to enhance malaria image classification performance [7].

4. Conclusions

This research demonstrated the application of the Support Vector Machine (SVM) algorithm for malaria detection using microscopic blood cell images. The dataset used consisted of two classes, namely Parasitized and Uninfected, obtained from Kaggle. Prior to classification, the images were processed through resizing, grayscale transformation, normalization, and HOG-based feature extraction.

The experimental results showed that the SVM model achieved approximately 79% accuracy with balanced precision, recall, and F1-score values for both image classes. The confusion matrix results showed that the model successfully classified 318 Parasitized images and 313 Uninfected images correctly. The obtained results demonstrate that the

proposed model was able to differentiate infected and uninfected blood cell images effectively.

Overall, the proposed SVM-based approach showed potential as an alternative method for automated malaria image classification with relatively simple and efficient computational processes. Nevertheless, this research still has limitations related to dataset quantity and limited parameter optimization. Future studies are expected to utilize larger datasets and apply more advanced optimization techniques to improve classification performance.

5. Patents

Not applicable.

Author Contributions: Conceptualization, A.A., N.A.B., and C.F.H.; methodology, A.A. and N.A.B.; software, A.A. and N.A.B.; validation, C.F.H.; formal analysis, C.F.H.; investigation, A.A.; data curation, A.A. and N.A.B.; writing—original draft preparation, A.A.; writing—review and editing, N.A.B. and C.F.H.; visualization, A.A.; supervision, C.F.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The dataset used in this study is publicly available on Kaggle at [kaggle.com/datasets/iarunava/cell-images-for-detecting-malaria](https://www.kaggle.com/datasets/iarunava/cell-images-for-detecting-malaria).

Acknowledgments: The authors would like to express their sincere gratitude to all parties who supported and contributed to this research.

Conflicts of Interest: The authors declare no conflict of interest.

References

- [1] World Health Organization, *World Malaria Report 2023*. Geneva, Switzerland: WHO, 2023. [Online]. Available: <https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2023>. Accessed: May 17, 2026.
- [2] Y. Zhang, L. Chen, and J. Wang, "Artificial intelligence-based diagnostic systems in infectious disease detection," *BMC Infectious Diseases*, vol. 25, 2025. doi: 10.1186/s12879-025-12116-6.
- [3] S. Rahman, M. Iqbal, and T. Hidayat, "Machine learning algorithm comparison for disease prediction," *Building of Informatics, Technology and Science (BITS)*, vol. 4, no. 1, pp. 88–96, 2022.
- [4] R. F. Putri, S. Nugroho, and A. Kurniawan, "Application of support vector machine for health data classification," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 9, no. 3, pp. 455–463, 2022.
- [5] P. Wulandari and E. Saputra, "Optimization of support vector machine in medical dataset classification," *Syntax Literate: Jurnal Ilmiah Indonesia*, vol. 8, no. 4, pp. 202–211, 2024.
- [6] F. H. Ramadhan, A. Maulana, and R. Pratama, "Healthcare technology implementation using artificial intelligence approach," *International Journal of Health Information Technology*, vol. 3, no. 1, pp. 11–19, 2024.
- [7] D. P. Lestari, R. Handayani, and A. Setiawan, "Implementation of convolutional neural network for malaria image classification," *Indonesian Journal of Data Science*, vol. 4, no. 2, pp. 91–101, 2023.
- [8] A. K. Singh and R. Kumar, "Classification of medical image data using support vector machine for disease detection," *International Journal for Research Trends and Innovation*, vol. 3, no. 9, pp. 18–24, 2018.
- [9] N. A. Prasetyo, D. W. Sari, and M. Hidayat, "Implementation of machine learning for medical diagnostic systems," *Jurnal Sainstech*, vol. 29, no. 2, pp. 67–75, 2019.

-
- [10] M. A. Albahri, A. S. Hamid, and Z. T. Alwan, "Role of artificial intelligence techniques in COVID-19 diagnosis and classification," *Diagnostics*, vol. 10, no. 5, p. 329, 2020. doi: 10.3390/diagnostics10050329.
- [11] T. Siregar, B. Sinaga, and H. Nasution, "Comparative analysis of classification algorithms for healthcare datasets," *Majalah Ilmiah INTI (MIB)*, vol. 9, no. 2, pp. 112–120, 2021.
- [12] I. Arunava, "Cell Images for Detecting Malaria," Kaggle Dataset, 2018. [Online]. Available: <https://www.kaggle.com/datasets/iarunava/cell-images-for-detecting-malaria>. Accessed: May 17, 2026.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the author(s) and do not necessarily reflect the views of the publisher and editor(s). The publisher and editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.