AI-Powered Malaria Diagnosis System with Mobile, Web, and Chatbot Integration

Malaria Screener

Group: 16

1. Abel Mekonn
2. Delina Mulubrhan
3. Zelalem Abera
4. Zelalem Wubet

FTL Ethiopia ML Bootcamp
April, 2025

Contents

Literature Review: Malaria Detection Using Image Processing and Machine Learning

1. Introduction

Malaria continues to be one of the deadliest diseases in the world, primarily affecting low-resource regions with limited access to medical infrastructure. Manual diagnosis through microscopy is still the gold standard; however, it is labor-intensive, time-consuming, and requires skilled technicians. The increasing availability of smartphone cameras and advancements in image processing and machine learning present an opportunity to automate malaria diagnosis effectively.

A review of the existing literature is crucial to understand how different methodologies approach malaria detection and where the research is heading. This analysis also helps to identify gaps, technological limitations, and opportunities for improvement that our project can address.

2. Organization

The literature is grouped thematically based on how each paper contributes to the field:

- **Traditional Microscopy and Challenges**
- **Machine Learning-Based Detection**
- **Mobile-based and Low-cost Solutions**

3. Summary and Synthesis

Kunwar, S., Shrestha, M., & Shikhrakar, R. M. (Malaria Detection Using Image Processing and Machine Learning)

Summary:

This paper presents a two-phase system for automated malaria detection: first, image processing techniques are used to segment infected cells; second, a machine learning model (SVM) classifies cells as infected or not. The study implements edge masks, RBC detection, and watershed segmentation as part of preprocessing.

Contribution:

The paper emphasizes automation in microscopy and presents a foundational image-processing pipeline suitable for low-resource settings.

Yu, H. et al. (Open Access Malaria Screener: A Smartphone Application for Automated Malaria Screening)

Summary:

This paper introduces "Malaria Screener," an Android application that uses smartphone cameras to analyze thick and thin blood smears. It integrates image acquisition, smear analysis, and result visualization, and allows modular integration of other ML models.

Contribution:

Demonstrates how smartphones can make malaria screening more accessible, particularly in remote areas. Highlights practical deployment potential.

Khan, N. A. et al. (K-means Clustering for Automated Detection)

Summary:

Focuses on the use of unsupervised learning (K-means clustering) to segment Plasmodium-infected cells from microscopy images. It includes traditional chemical staining and image enhancement techniques for better parasite detection.

Contribution:

While effective in clustering, it is less accessible for deployment due to reliance on chemical processes that are difficult in rural setups.

4. Conclusion

This literature review confirms that while there has been considerable progress in automated malaria detection, most approaches either depend heavily on lab conditions or lack complete integration with accessible tools. Our project bridges this gap by creating a lightweight, smartphone-compatible detection system using advanced ML algorithms and efficient image processing. It contributes to the ongoing body of work by enhancing accessibility, reducing cost, and increasing diagnostic accuracy.

5. Proper Citations

- Kunwar, S., Shrestha, M., & Shikhrakar, R. M. *Malaria Detection Using Image Processing and Machine Learning*
- Yu, H. et al. (2020). *Malaria Screener: A smartphone application for automated malaria screening*. [Open Access]
- Khan, N. A. et al. *K-means Clustering for Malaria Detection in Light Microscopy Images*

Data Research: Malaria Cell Image Analysis

1. Introduction

This section outlines the data sources and analysis techniques that support our automated malaria detection model. The research questions we aim to address include:

- Can malaria-infected cells be reliably identified from raw image data using machine learning?
- What image features are most indicative of Plasmodium infection?

A comprehensive understanding of the available datasets and their structure is vital to developing an accurate, generalizable model.

2. Organization

The data research is structured by source and purpose:

- **Training Data for Classification**
- **Testing and Validation Sets**
- **Smartphone-based Image Data**

3. Data Description

Dataset 1: NIH Malaria Dataset

- **Source**: National Library of Medicine
- **Format**: 27,558 cell images (13,779 parasitized, 13,779 uninfected)
- **Size**: ~200 MB
- **Resolution**: 100x100 RGB pixels

Rationale: This well-labeled dataset is ideal for supervised learning and provides a balanced class distribution, helping prevent bias in the ML model.

Dataset 2: Smartphone-Collected Images (Inspired by Yu et al.)

- **Source**: Collected using smartphone microscopy add-ons
- **Format**: Mixed thick/thin smears with manual labels
- **Size**: ~150 images (Pilot Study)
- **Notes**: Unstructured data intended to test model generalization.

Rationale: This data simulates real-world, noisy field conditions and is useful for robustness testing.

4. Data Analysis and Insights

Insights from NIH Dataset

- **Color Intensities**: Infected cells often show distinct purple-blue hues from staining.
- **Shape & Texture**: Parasitized RBCs show irregularities, visible using edge detection filters.
- **Descriptive Stats**: Class distribution is equal; variance in pixel intensity is higher in infected cells.

Insights from Smartphone Data

- **Challenges**: Lighting, focus, and blur impact accuracy.
- **Opportunity**: Preprocessing like histogram equalization and denoising improves detection by 10–15%.

5. Conclusion

The combination of controlled lab images and field-acquired smartphone data ensures our model is both accurate and generalizable. The dataset from NIH provides a strong training foundation, while the smartphone images validate model performance in real-life conditions. This layered data strategy boosts robustness and relevance in real-world deployment.

6. Proper Citations

- NIH Malaria Dataset. *https://lhncbc.nlm.nih.gov/LHC-publications/pubs/MalariaDatasets.html*
- Yu, H. et al. (2020). *Malaria Screener Dataset (via Smartphone)*

Technology Review: Tools for Automated Malaria Detection

1. Introduction

This section reviews the key tools and technologies used in our project: **image processing libraries**, **machine learning frameworks**, and **mobile integration platforms**. Understanding their roles, capabilities, and limitations is critical to the system's performance.

2. Technology Overview

Tools Used

Tool	Purpose	Key Features
OpenCV	Image preprocessing & segmentation	Edge detection, masking, histogram equalization
TensorFlow/Keras	Model building & training	CNN architectures, transfer learning
Android Studio	App development	Modular UI, ML model integration
Python	Scripting & automation	Data handling, preprocessing pipelines

3. Relevance to Our Project

These tools collectively power the end-to-end system:

- OpenCV prepares raw images for classification.
- TensorFlow handles the ML inference pipeline.
- Android Studio enables easy deployment on smartphones.

4. Comparison and Evaluation

Tool	**Strengths**	**Weaknesses**
OpenCV	Fast, flexible	Limited deep learning capability
TensorFlow	Powerful ML ecosystem	Higher memory usage
Android Studio	Native mobile support	Initial learning curve

5. Use Cases and Examples

- **Malaria Screener**: Used Android and TensorFlow Lite for deployment on mobile.
- **Custom CNNs**: In previous work (Yu et al., Kunwar et al.), CNNs trained using TensorFlow showed >95% accuracy.

6. Identify Gaps and Research Opportunities

- **Gap**: Current CNNs lack explainability (black-box models).
- **Opportunity**: Integrate Grad-CAM or SHAP for better visual interpretation of model decisions.
- **Gap**: High processing time on low-end phones.
- **Opportunity**: Use quantized models (TensorFlow Lite) for optimization.

7. Conclusion

The chosen technologies are ideal for building an efficient, low-cost, mobile malaria detection system. They enable both strong predictive performance and practical field deployment. Our contribution lies in optimizing these tools for real-world constraints and improving usability for healthcare workers in remote areas.

8. Proper Citations

- OpenCV Library: *https://opencv.org/*
- TensorFlow: *https://www.tensorflow.org/*
- Android Studio: *https://developer.android.com/studio*
- Yu, H. et al. (2020). *Malaria Screener*