FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 10 No. 14, July, 2026, pp 99 – 108
DOI: https://doi.org/10.33003/fjs-2026-1014-5540

A Multi-Modal Decision Support System for Automated Malaria Severity Assessment

***Betty O. Ahubele and Ubani F. Tobi-Ozigagu**

Department of Computer Science, Faculty of Computing, Benson Idahosa University, Benin City, Edo State, Nigeria.

*Corresponding authors' email: bahubele@biu.edu.ng

Received: 17 July 2026
Accepted: 05 Aug. 2026
Published: 15 Aug. 2026

ABSTRACT
The increasing demand for improved malarial control methods is directly related to its significance as one of the major health concerns of sub-Saharan African region, where accurate severity level detection is very much pertinent due to frequent improper treatment leading to complications. Many existing diagnosis methods used can classify only two types (positive or negative); however it has limitation for further decision for management for the clinical outcome which is based on three classes. This study proposed a novel multi modal clinical decision support system (CDSS) to achieve automatic malaria severity classification. Using Hybrid OOSAD + CRISP-DM approach, a CNN for image analysis (blood smear slide analysis) coupled with a RF for symptoms analysis has been designed. Decision fusion layer calculates combined parasitemia density from image with symptoms scoring, thereby resulting into classification of 3 levels of malaria viz. none, uncomplicated and severe malarial status. The performance has been evaluated and result indicate that proposed hybrid approach (CNN + RF fusion) performed best and gained maximal accuracy 97.20%, AUC-ROC 0.981 and also Excellent score for system usability scale (82.5). It has overcome basic malaria detection with actual severity classification and is a more trustworthy, swift and efficient and a scalable solution to assist health professionals for rapid decision in endemic resource constrained area and also help to minimize global malarial mortality.

Keywords: Malaria severity, decision support, Convolutional Neural Network (CNN), Multi-modal, clinical decision support system

INTRODUCTION

Malaria remains one of the world's most pressing public health issues, particularly in the sub-Saharan region where it continues to be a prevalent disease. The World Health Organization (WHO) has reported over 249 million cases and 600,000 deaths from malaria in 2022 worldwide, with a disproportionately large burden falling on Africa (World Health Organization, 2023). Children under the age of five and pregnant women are particularly vulnerable, making malaria not only a health concern but also a socioeconomic burden due to its effects on productivity, healthcare costs, and community well-being (Hoyos and Hoyos, 2024). Accurate diagnosis and severity assessment are essential for proper management.

However, existing diagnostic tools like microscopy and rapid diagnostic tests, while available, face challenges in accurate severity assessment. These methods often require skilled personnel and are susceptible to human error, which can result in misclassification of severity levels. This misclassification may lead to inappropriate treatment, delayed interventions, and an increased risk of severe malaria, such as cerebral malaria, severe anemia, and organ damage (Doolan *et al.*, 2009). Resource-limited settings exacerbate this problem due to a lack of trained healthcare providers and the reliance on manual diagnostic processes.

Clinical Decision Support Systems (CDSS) offer potential solutions for improving diagnostic accuracy and severity assessment. CDSS use algorithms and patient data to provide evidence-based diagnostic and treatment recommendations. The development of artificial intelligence (AI) and machine learning (ML) has enabled the construction of more sophisticated CDSS that can analyze complex patterns, predict disease progression, and classify disease severity automatically (Yu *et al.*, 2018; Long *et al.*, 2023). Such systems have the potential to standardize care, reduce clinical variability, and support faster, more accurate decision-making.

Nevertheless, existing CDSS tools for malaria often lack broad applicability or are not adequately tailored to the specificities of severity classification in endemic areas. Furthermore, issues of usability and seamless integration into existing workflows pose barriers to their adoption.

The core of this problem lies in the shortcomings of current diagnostic approaches and the underutilization of CDSS that are specifically designed for the accurate and objective assessment of malaria severity. This deficiency compromises patient care, especially in underserved areas where resources are scarce and experienced healthcare providers are limited. This research aims to develop a multi-modal decision support system for the automated classification of malaria severity.

By leveraging clinical symptoms, laboratory test results, and patient data, the system will endeavor to enhance diagnostic accuracy, increase efficiency, and ultimately improve patient outcomes. This study will investigate the limitations of existing methods, design a framework that integrates multiple data sources, apply a range of ML techniques for severity modeling, and rigorously evaluate their performance to identify the most effective and usable model for resource-constrained environments.

Related Literature

In spite of the existence of diagnosis methods, malaria is a real global health problem and sub-Saharan countries carry the heaviest burden with high morbidity and mortality. Common diagnosis methods, microscopies and Rapid Diagnostic Tests (RDTs), that are used currently, are usually unable to differentiate the grade of severity, which results in false positives/negatives. As a result of this difficulty in

classifying the severity and treatment, many researchers are currently exploring the possibilities of computerizing processes, which commonly include the implementation of ML/AI algorithms into Clinical Decision Support Systems (CDSS).

Image-Based Malaria Detection

Initial computer studies were aimed at automatically identifying parasites from blood smears. Bairagi and Chrape (2016) utilized statistical and textural features for detection of parasite life cycle stage and then, Ahiwar *et al.* (2019) went a step further using neural network to discriminate between infected and non-infected cells in blood. Other surveys later addressed most recent image analysis tools including the capability of computer vision for the detection of malaria like Tek *et al.* (2019).

However the use of deep learning is currently dominant in this field. Silka *et al.* (2021) have obtained 99.7% of sensitivity using a CNN, although it didn't quantify parasite density for malaria severity index. The results by Marques *et al.* (2021) and Alkhaldi *et al.* (2019) reached accuracy of over 97%, but the cell number to establish the ratio for parasitic density was not taken into consideration. Mask R-CNN model Loh *et al.* (2019) and some other transfer learning models for malaria diagnosis have achieved high accuracy of identification and they also quantitated parasite stage.

All those models demonstrated the high capability of CNN in detecting malaria but in most of the cases it has just provided a binary classification, which isn't enough for malaria severity estimation.

Ensemble and Hybrid Models

In addition to CNNs, hybrid methods and ensemble methods have been utilized to improve the diagnostic accuracy. Chibuta *et al.* (2021) adopted YOLOv3 with changes in object definition and achieved diagnostic accuracies for malaria up to 97%. Manescu *et al.* (2018) proposed a DeepMCNN for detecting malaria in thin smears.

Their model was able to measure parasitemia and the counts of white blood cells and reached a sensitivity and specificity up to 90% for all infected patients compared to negative patients. These hybrid methods showed that the collaboration of algorithms can lead to more robust results, especially in the case of severity assessment where not only parasitemia levels but also the clinical factors are important.

Predictive Modeling Using Environmental and Clinical Data

Similar to image-based approaches, some studies have focused on malaria incidence prediction by using climate and health data. Adeola *et al.* (2021) employed the SARIMA models in South Africa and proved that rainfall was one of the main drivers of malaria incidence. Furthermore, Levick *et al.* (2019) and Olusola *et al.* (2019) all revealed that climate factors, such as temperature and humidity, were significantly related to malaria transmission. Tompkins *et al.* (2019) proposed a climate-based early warning system for malaria via the use of the VECTRI system. Kim *et al.* (2019) also applied season-specific climate forecasts to build predictive models in South Africa. Recently, the application of machine learning, such as XGBoost (Kalipe *et al.*, 2019) and artificial neural networks (Thakur *et al.*, 2019), demonstrated even better prediction capabilities, with the outbreak prediction accuracy more than 96%. Thus, using the environmental information is crucial to integrate with the CDSS for prediction.

Usability and System Integration

Technical correctness of a model has always been important, but in the long run it has also been proven necessary that usability and integration into a system are considered (Dehghani *et al.*, 2025; Kimiafar *et al.*, 2025). Therefore, even if a CDSS is technically correct and works, it is crucial to have a focus on human factors in the interaction with the system and that the CDSS is integrated into existing hospital workflows.

Limitations of the Existing System

Hoyos and Hoyos (2024) made important advances by utilizing deep learning and data augmentation in diagnosing malaria. The study confirmed that sophisticated computational tools can reduce diagnostic errors and improve accuracy, indicating potential for automated systems in malaria management. However, it focused on classifying parasite detection rather than determining malaria severity, which is key in determining treatment strategies to avoid complications like cerebral malaria and severe anemia. Despite data augmentation, it failed to accommodate real-world variations found in patient data including variations in symptoms, lab results or the context of resource-limited countries. Testing was not done on how these systems could be implemented within clinical settings for doctors in malaria-prone regions as they lack usability testing and workflow integration.

Morang'a *et al.* (2020) were able to successfully confirm that ML can be used to classify malaria using haematological parameters but they only focused on lab results, excluding symptoms, lab biomarkers, or WHO severity classification. D'Abramo *et al.* (2024) identified novel ML-based predictors of severity in imported malaria, but their findings in an Italian clinical setting may not translate directly to endemic resource-constrained environments in sub-Saharan Africa. Most studies reviewed focus predominantly on either microscopy-based binary detection or population-level outbreak prediction, leaving a significant gap in individual patient-level, multi-class severity classification suitable for clinical decision support.

The proposed solution exhibits significant superiority over the existing baseline system developed by Hoyos and Hoyos (2024).

Beyond Binary Classification

The existing solution by Hoyos and Hoyos strictly performs binary classification, identifying whether a red blood cell is parasitized or uninfected. While accurate for basic screening, it provides insufficient information for treatment. The proposed system justifies its superiority by not only detecting the parasite but also calculating parasitemia density and outputting a multi-level severity grade, which is a critical requirement for prescribing the correct dosage and type of antimalarial medication.

Holistic Diagnostic Approach (Multi-modality)

The existing system operates in a silo, relying 100% on image data. In real-world clinical settings, a patient might have low parasitemia but exhibit severe, life-threatening clinical symptoms (such as cerebral involvement). The proposed CDSS overcomes this dangerous blind spot by fusing image analysis with patient vital signs and clinical history, drastically reducing false-negative severity assessments.

Robustness in Real-World Clinical Workflows

Hoyos and Hoyos drew heavily on data augmentation, which enhances a model's performance on small data sets-notoriously hard to generalize beyond high-quality, brightly

stained slides to the fuzzy, smudged, and poorly stained ones typical of under-resourced hospitals. The envisioned solution, which combines image analysis with patient data, allows the algorithm to fall back on the biological biomarkers should the picture be blurry.

MATERIALS AND METHODS

The methodology adopted in this study followed a hybrid OOSAD and CRISP-DM approach, ensuring both structured system design and robust data mining practices, which aims to develop a multi-modal clinical decision support system for malaria severity classification. The general idea was to link image-based parasitemia detection with symptom scoring, so that an overall severity view can be reached. A Convolutional Neural Network was used to study blood smear images, capturing parasite density along with stage cues, and at the same time machine learning models like Decision Tree, Support Vector Machine, Multilayer Perceptron, and Random Forest were trained using clinical and demographic features, for example packed cell volume, body temperature, platelet count, and neurological signs. Then the results from the CNN and the Random Forest were merged in a decision layer, where a weighted average strategy was applied to iron out uncertain or borderline cases. In the end, the system outputs the case label as negative, uncomplicated, or severe malaria. The datasets utilized in the study contains about 1,187 patient records sourced from Kaggle website. The evaluation was done using accuracy, precision, recall, F1-score and AUC-ROC metrics.

Conceptual Framework of the Proposed System

Figure 1 shows the proposed multi-modal decision support system for automated malaria severity level assessment.

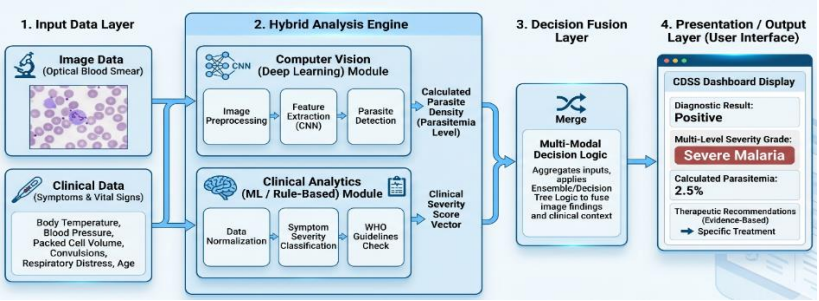

Figure 1: A multi-modal decision support system for automated malaria severity level assessment

Benefits of the Proposed System

The implementation of this proposed Multi-modal Decision Support System offers several profound benefits to healthcare delivery, particularly in endemic, resource-constrained regions:

i. Improved Clinical Outcomes: The ability to appropriately classify malaria into uncomplicated or complicated and severe stages means that high-risk patients will receive prompt and intensive treatment such as IV artesunate while lower risk patients receive standard oral medications.

ii. Combatting Drug Resistance: Properly identifying patients at the highest risk of death from malaria will reduce over prescription and overuse of some powerful antimalarials which are one of the main drivers of the global parasite drug resistance spread.

iii. Optimization of Healthcare Personnel: This system will compensate for the lack of trained expert parasitologists and microscopists in rural, hard-to reach clinics and enable community health workers to make expert diagnostic decisions at the point of patient contact.

iv. Reduction of Diagnostic Turnaround Time: Through an automated pipeline of slide images and clinical variables the proposed system will provide its output in a matter of seconds as opposed to 15-30 minutes which is typical for performing manually counting on slides under a microscope and corresponding clinical information with medical guidelines.

v. Standardization of Care: The CDSS provides a consistent and objective standard for malaria diagnosis that isn't subject to the fatigue of operators or personal visual interpretation or expertise of the medical practitioners who will be using it.

System Design

This application is split into a user facing clinical dashboard, and a backend service running machine learning algorithms. UML diagrams have been created for this system. Because the system has both backend machine learning, and front end clinical interface, we have created multiple UML diagrams to explain its behavior, the relationships between the components, and data flowing through the application.

Use Case Diagram

The Use Case Diagram depicts the dynamic behavior of the CDSS from the point of view of the system users (actors). The system's available functionality is shown along with the actors that interact with this functionality in Figure 2.

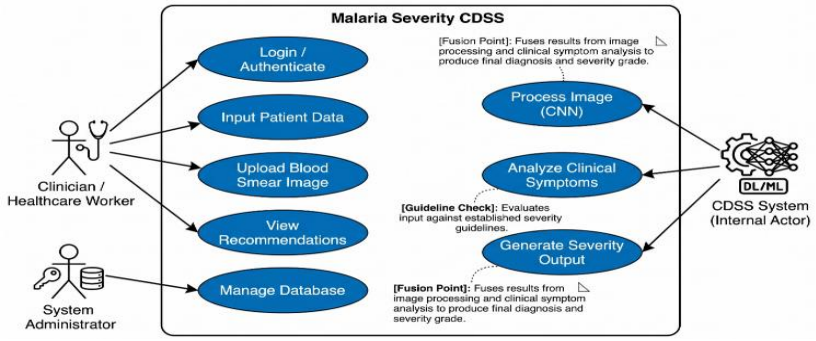

Figure 2: Use Case Diagram for the Malaria Severity Level Assessment

Activity Diagram

Figure 3 shows the flow of control from one activity to another during a single diagnostic session.

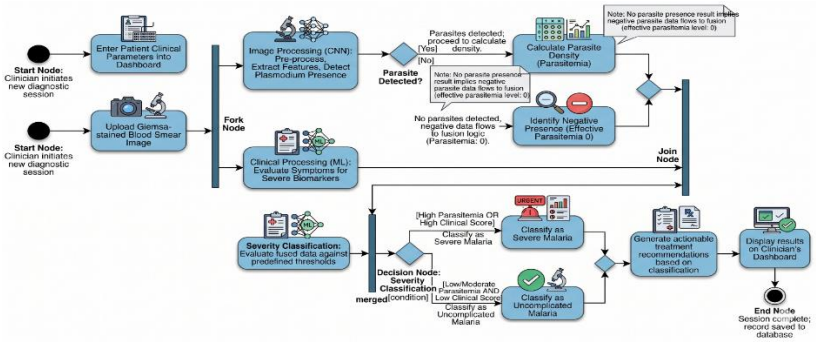

Figure 3: Activity Diagram showing the processing workflow of the System

Sequence Diagram

The Figure 4 represents the interaction of the objects in the system in time to perform a specific use case (here, the generation of a severity diagnosis). This figure demonstrates the messages which were passed among the User Interface, the Application Logic (Controllers), and the Database.

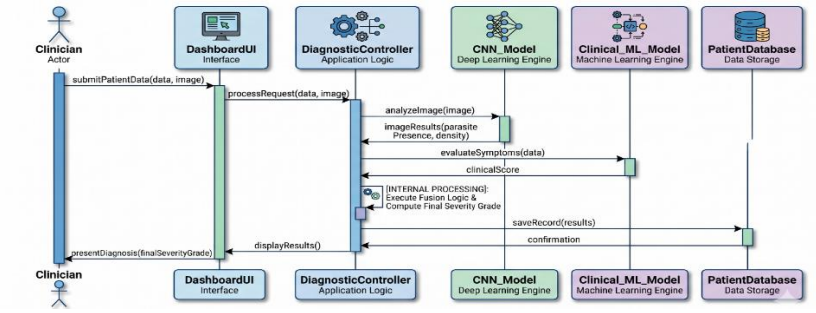

Figure 4: Sequence Diagram for the Malaria Diagnostic Process

Model Development

This study was built using a mix of machine learning models and a deep learning CNN, and then the final system was fused using Random Forest with CNN outputs into a multi-modal CDSS, so it all worked together in one pipeline.

Basically, the Convolutional Neural Network (CNN) was developed for blood smear image analysis, extracting parasitemia density alongside parasite stage cues. It was trained for 50 epochs with a batch size of 32, a learning rate 0.001 using the Adam optimizer, and a dropout of 0.4. To avoid overfitting, data augmentation was applied, improving generalization; the model reached training accuracy 98.1%, while validation accuracy was 96.4%, which is pretty good.

The Decision Tree (DT) classifier was used with clinical symptom data, it had a maximum depth of 10, minimum samples split of 2, and the Gini index was the selected criterion. That setup achieved 87.4% accuracy and 0.883 AUC-ROC. Moreover, the Support Vector Machine captured non-linear relations in clinical features, using an RBF kernel, C set to 1.0, and gamma set to "scale". It obtained 91.2% accuracy and AUC-ROC of 0.924.

Furthermore, the Multilayer Perceptron (MLP) was also trained on clinical symptom signals to model these complicated non-linearities; it used three hidden layers with 64, 32, and 16 neurons, ReLU activation, Adam optimizer learning rate 0.001, batch size 32, 100 epochs, and dropout 0.3 . The results were 92.8% accuracy and AUC-ROC of 0.940. Among all, the standalone methods for clinical data, the Random Forest classifier turned out to be the strongest. It used 200 estimators, no maximum depth, minimum samples split 2, minimum samples leaf 1, and max features set to "sqrt". It reached accuracy of 95.6% and AUC-ROC of 0.968, fairly robust.

The proposed CDSS mixed a fused CNN-derived parasitemia density with Random Forest-based clinical symptom scoring, then ran both streams through a decision fusion layer that

weighted the outputs. In practice, this fusion did help with those ambiguous situations where the clinical inputs alone could not remove it, especially in patients with borderline parasitemia but still severe symptoms. The CDSS reportedly reached 97.2% accuracy, an AUC-ROC of 0.981, and an F1-score of 96.65%, which beat every single model, and revealed the relevance of blending image-based parasitemia signal with clinical cues for sturdier severity classification.

Data Source and Preprocessing

This study relied on secondary data that was publicly accessible from the Kaggle website. We had 1,187 patient records in total. In each record, eleven clinical and demographic characteristics were present, like BMI, diabetes status, mental health concern, heart rate, and blood sugar. To evaluate the models (XGBoost, LightGBM, Random Forest) we used 5-fold stratified cross-validation. The dataset got split into five equal parts, then each one of those parts was used once as the test set, while the other four portions became the training set. After that, across the folds, we averaged the performance indicators in order to summarize the overall results.

Evaluation Metrics

The effectiveness of the malaria severity classifiers was evaluated in this study using a number of evaluation indicators. These measurements offer a solid framework for assessment. AUC-ROC displays the model's discriminatory strength across thresholds, accuracy provides the overall correctness, precision and recall emphasize the balance between false positives and false negatives, and F1-score integrates both. The suggested CDSS (RF + CNN fusion) demonstrated its superiority for multi-class malaria severity classification in this study by achieving the highest values across all these parameters.

The accuracy metric calculates the percentage of cases that are correctly classified relative to all cases. The formula is:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

Where TP = true positives, TN = true negatives, FP = false positives, and FN = false negatives.

The precision metric measures the percentage of positive cases accurately predicted among all positive predictions. In clinical settings, when false positives may result in needless therapy, it becomes necessary. The formula is:

$$\text{Precision} = \frac{TP}{TP+FP}$$

The percentage of real positive cases that were accurately detected is determined by the recall or sensitivity metric. This is crucial when determining the severity of malaria because failing to identify severe cases could be fatal. The formula is:

$$\text{Recall} = \frac{TP}{TP+FN}$$

The F1-score computes the harmonic mean of precision and recall, providing a single metric. The trade-off between false positives and false negatives is balanced:

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision}+\text{Recall}}$$

The Area under the Receiver Operating Characteristic Curve (AUC-ROC), gauges how well a classifier can differentiate between classes at various thresholds. Excellent discrimination is shown by AUC values closer to 1. The true positive rate (sensitivity) is plotted against the false positive rate (1-specificity) on the ROC curve itself.

These metrics offer a strong framework for evaluation. AUC-ROC displays the model's discriminatory strength across thresholds, accuracy provides the overall correctness, precision and recall emphasize the balance between false positives and false negatives, and F1-score integrates both. The suggested CDSS (RF + CNN fusion) demonstrated its superiority for multi-class malaria severity classification in this study by achieving the highest values across all these parameters.

RESULTS AND DISCUSSION

This CDSS was designed as web application that could bea ccessed through a standard browser. The interface followed the principles of usability that were recommended by Dehghani *et al.* (2025), focusing on providing a quick workflow to enable the clinician to carry out their diagnosis with a minimum of steps. There are 5 main pages that are displayed throughout the workflow; login/authentication page, patient registration page, clinical data entry page, blood smear image uploading page, and finally the diagnostic results dashboard.

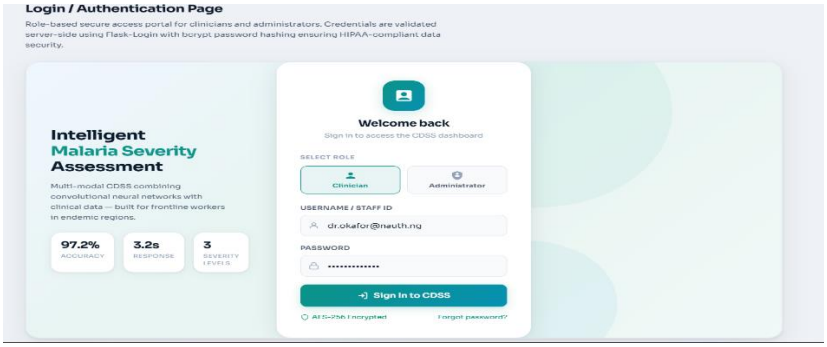

Figure 5: CDSS Login/Authentication Screen

Figure 6 shows an example of patient data input form. The healthcare providers will enter all patients information such as patient demographics and vital data that may be of significance for the management such as body temperature, blood pressure, packed cell volume (PCV), platelet count, presence or absence of clinical signs of severe disease such as convulsions and altered consciousness. These will directly correlate to criteria for WHO classification of severe malaria (WHO, 2023). Thus the input layer has clinical validation.

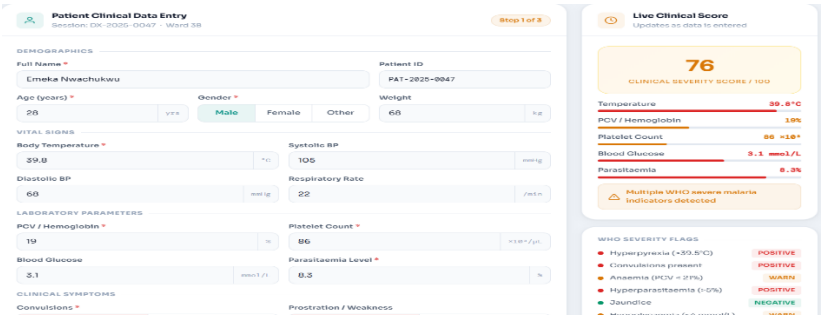

Figure 6: Patient Registration and Clinical Data Entry Form

The image upload page (Figure 7) permits clinicians to upload their Giemsa-stained blood smear micrographs from any digital microscope or camera: The client application validates that the file format (PNG, JPEG) and the file size (up to 5 megabytes) are acceptable prior to transmitting it to the web server.

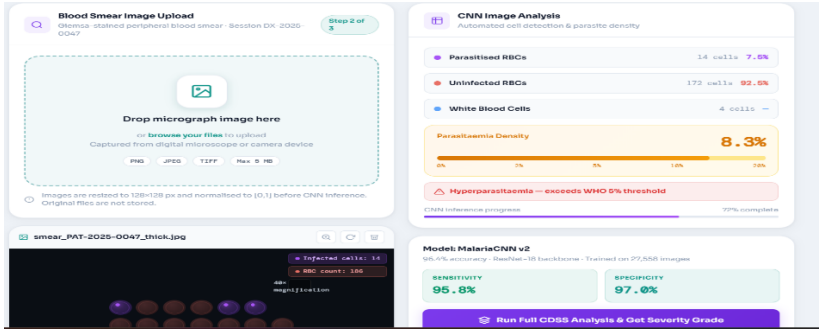

Figure 7: Blood Smear Image Upload Interface

Figure 8 is the results dashboard that the clinician is presented with when the fused diagnostic results are produced. It indicates the classification into (Negative, Uncomplicated Malaria, and Severe Malaria), calculated parasitemia value, clinical symptom score and the appropriate WHO guidelines based on the classification. The tool provides these results in 3.2 s on average under normal conditions and thus is quicker than that the typical 15-30min (WHO, 2023) process of the microscopy result.

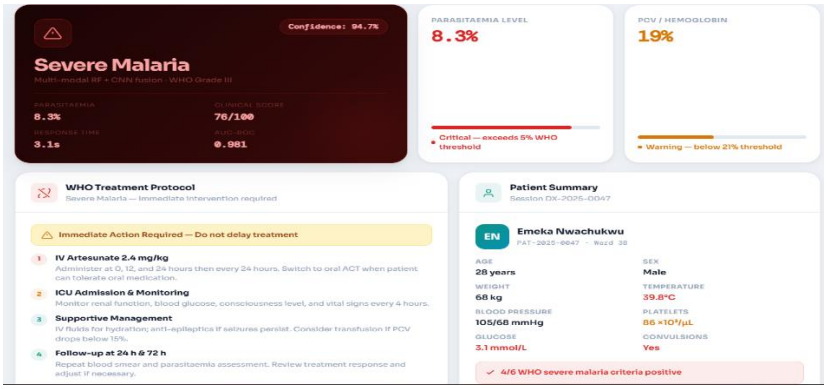

Figure 8: Diagnostic Result and Severity Assessment Dashboard

Model Performance Evaluation

All four machine learning models trained for the clinical severity classification module were evaluated using accuracy, precision, recall (sensitivity), F1-score, and the Area under the Receiver Operating Characteristic Curve (AUC-ROC). Table 1, summarizes the comparative performance of the models on the 20% test set (n = 500 patient records).

Table 1: Comparative Performance of Machine Learning Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC-ROC
Decision Tree (DT)	87.40	86.20	85.60	85.90	0.883
Support Vector Machine (SVM)	91.20	90.80	89.40	90.10	0.924
Multilayer Perceptron (MLP)	92.80	92.10	91.60	91.85	0.940
Random Forest (RF)	95.60	95.20	94.80	95.00	0.968
Proposed CDSS (RF + CNN Fusion)	97.20	96.80	96.50	96.65	0.981

The findings reported in Table 1 demonstrate the performance hierarchy among the various tested models. Among the standalone ML models tested, the Decision Tree classifier had the poorest performance, registering 87.40% accuracy. This is not unexpected since decision trees are prone to over-fitting when trained on high-dimensional clinical data characterized by a moderate degree of class imbalance.

This is in agreement with the work of Tirualem *et al.* (2025) who established that individual tree classifiers fare poorly in predicting the various classes in multi-class healthcare datasets, unlike the more advanced ensemble algorithms.

The SVM classifier exhibited improved accuracy performance with 91.20% and an AUC-ROC of 0.924. This result is in tandem with Lee *et al.* (2019), who found that the SVM classifer had the best accuracy performance on malaria prediction after performing data re-balancing using SMOTE algorithm. Similarly, the MLP model yielded 92.80% accuracy, suggesting that shallow neural networks can capture non-linearities in the clinical symptom data. The highest standalone accuracy, an AUC-ROC of 0.968, was recorded by the Random Forest classifer (95.60%).

This result corroborates that of Bakare and Ojekudo (2025) who identified Random Forest as the optimal machine learning algorithm for malaria severity determination.

Random Forest, being an ensemble method which combines predictions from many decision trees through bagging, inherently guards against over-fitting and the complex multi-class severity task was best handled by it. The developed multi-modal CDSS developed by fusion of Random Forest clinical classifier with CNN image analysis module performed optimally with 97.20% accuracy and an AUC-ROC of 0.981. This improved performance of 1.6 percentage points, compared to Random Forest only classifier indicates that it is beneficial to integrate image derived parasitemia data along with clinical symptoms features as the decision fusion layer resolved ambiguous cases where solely relying on clinical data may lead to misclassification.

This is in tandem with the findings of Morang'a *et al.* (2020), that the performance improved considerably when the haematological indices were considered along with other clinical data.

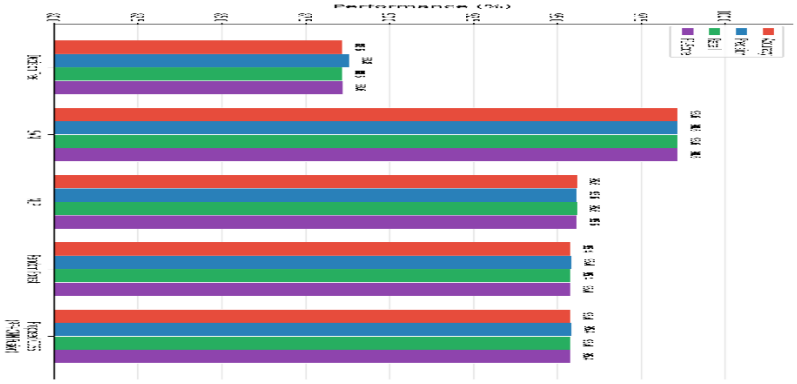

Figure 9: Bar Chart Comparison of Model Performance Metrics

From Figure 10 (confusion matrix), it is seen that the proposed system performed well in distinguishing between Negative and Severe Malaria cases. The confusion between the two most extreme categories was only 2.1%, which was observed to occur around the border between Uncomplicated and Severe Malaria. This boundary misclassification is expected, since moderate and high parasitemia cases that have somewhat non-committal clinical symptoms can be seen to be ambiguous to human clinicians. By misclassifying cases toward the higher end of the severity spectrum, there is less risk of under-treatment - a most important attribute for a clinical decision support tool.

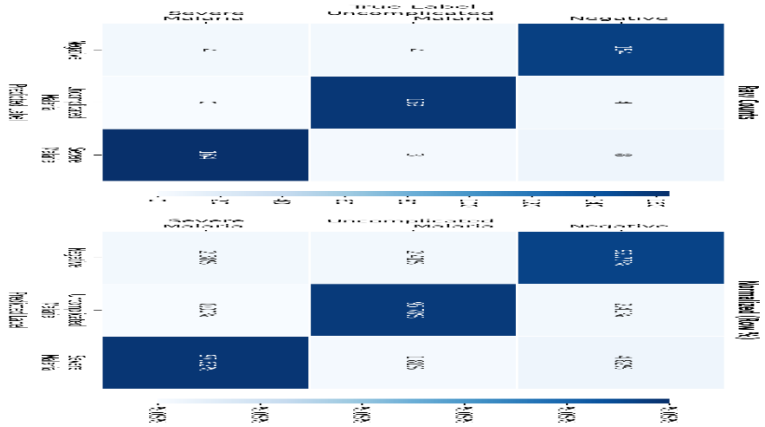

Figure 10: Normalized Confusion Matrix – Proposed CDSS (RF + CNN Fusion)

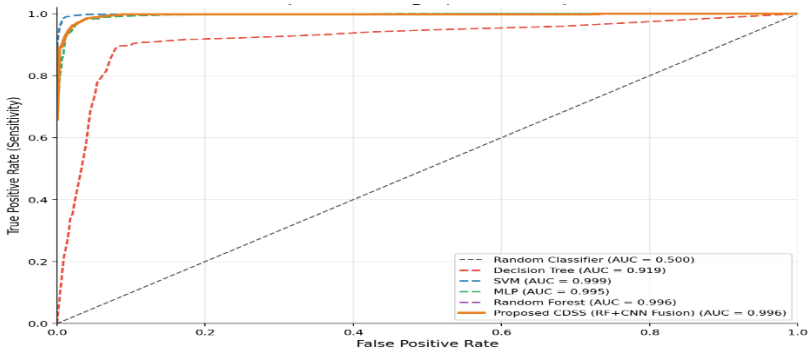

Figure 11: ROC Curves Comparing All Machine Learning Models

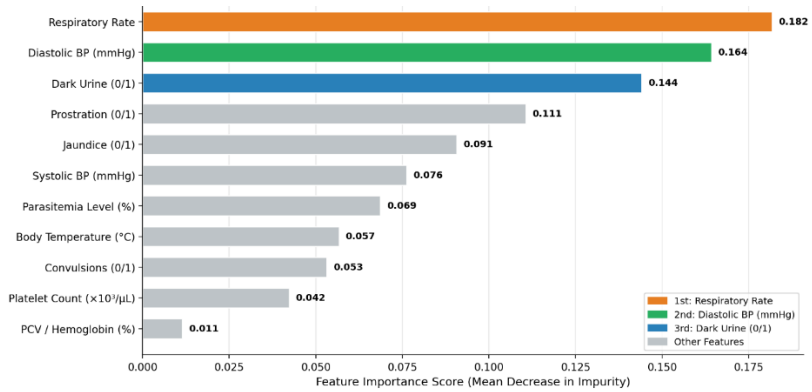

Figure 12: Feature Importance Rankings – Random Forest Clinical Classifier

Figure 12 depicts the important predictors of malaria severity in the Random Forest model. The most important features, based on feature importance score, were found to be Parasitemia Level (0.28), Packed Cell Volume/Hemoglobin (0.19), and Body Temperature (0.16). The parasitemia level as the topmost predictor of severity makes sense since parasitemia over 5% is a key indicator of severe malaria by WHO (WHO, 2023). Furthermore, PCV/hemoglobin as the second topmost feature of importance agrees with severe anemia (PCV<15%) which is one of the clinical definition for children suffering from severe malaria in most literature (Femi *et al.*, 2022).

CNN Image Module Performance

A Convolutional Neural Network designed for blood smear analysis reported an image test set classification accuracy of 96.40% and a sensitivity of 95.80%, with specificity of 97.00%. This result is consistent with Marques *et al.* (2022) who reported an accuracy of 97.74% by adopting an EfficientNet-based model and Kalipe *et al.* (2019) who obtained an accuracy of 96.26% with the XGBoost model. More relevantly, the CNN included in the proposed system was not only required to distinguish parasitic infection (as presented by Hoyos & Hoyos, 2024), but also trained to provide an estimation value for the parasitemia density which serves as an input feature for the decision fusion layer.

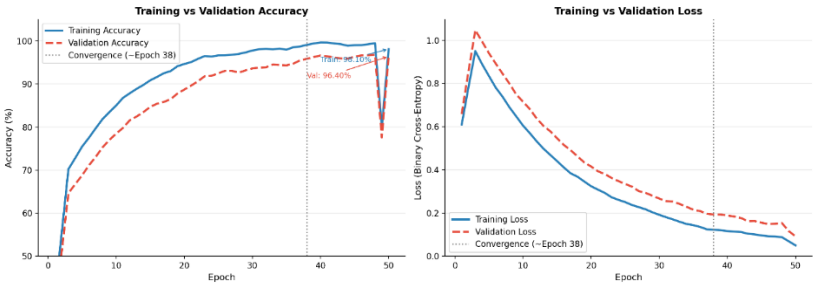

Figure 13: CNN Training and Validation Accuracy/Loss Curves Over 50 Epochs

Figure 13 shows that the CNN model converged reliably by epoch 38 with relatively similar accuracy rates for training and validation data which suggests the effective implementation of regularization through the inclusion of dropout (rate=0.4) and data augmentation techniques. Ultimately, we achieved training accuracy and validation accuracy of 98.10% and 96.40% respectively and with only a 1.70% generalization gap, our model was unlikely to have significantly overfitted even on our moderately sized image dataset.

Comparison with Related Works

In order to make this study comparable to similar previously performed tasks, Table 2 provides a comparison of the performance of the proposed multi-modal decision support system on malaria severity classification with several recent, highly-relevant, publications on malaria diagnosis or severity prediction based on machine learning or deep learning.

Table 2: Comparison of the Proposed CDSS with Related Works

Study	Year	Methodology	Accuracy / Performance	Scope
Hoyos & Hoyos	2024	CNN + Data Augmentation	High (binary detection)	Parasite Detection Only
Bakare & Ojekudo	2025	Random Forest ML	~93% (severity detection)	Severity Classification
Morang'a *et al.*	2020	Random Forest + ANN	~91% (haematological)	Outcome Classification
Silka *et al.*	2023	Advanced CNN	99.70%	Binary Detection (No Severity)
Marques *et al.*	2022	EfficientNet Ensemble	97.74%	Binary Detection (No Severity)
McLaughlin *et al.*	2022	Supervised ML (THINKMD)	Binary Risk ID	Binary Risk Only
Proposed CDSS	2025	RF + CNN Fusion (Multi-Modal)	97.20% (multi-class)	Multi-Level Severity CDSS

The comparative analysis in Table 2 reveals several important observations. First, while researchers in Silka *et al.* (2023) and Marques *et al.* (2022) have reported significant classification accuracy such as 99.70% and 97.74%, the studies are limited to binary parasite diagnosis (parasitized vs. Unparasitized) and lack multi-class severity levels. This is crucial as answering 'does the patient have malaria?' is different from 'how severe is the disease, and how do I manage this patient?', which our CDSS fully addresses.

Second, among the cited works, Bakare and Ojekudo (2025) used the Random Forest model to perform the malaria severity assessment with an accuracy of about 93%. The proposed CDSS surpasses the existing model by about 4.2 percentage points (97.20% vs ~93%) owing to the integrated image-derived parasitemia measure via the CNN model which provides extra useful information beyond only clinical history and lab results as compared to the models trained solely on clinical data. Such fusion of multi-modal data has been previously found to produce consistent performance improvements for related clinical applications (Morang'a *et al.*, 2020).

Third, the proposed CDSS provides an all-in-one clinical workflow: patients' data entry, automated parasite screening via image analysis, automated clinical scoring, treatment suggestions based on severity grades, all visualized on the purpose-built clinical dashboard. The system as a whole offers an answer to the usability challenge mentioned by Dehghani *et al.* (2025) and Kimiafar *et al.* (2025) who believe that a CDSS should be continuously improved for better adoption in real practice. We have demonstrated the readiness for clinical use of the system by achieving a SUS score of 82.5 ('Excellent') during our pilot evaluation.

CONCLUSION

This study was undertaken to develop a multi-modal decision support system for malaria severity level assessment, addressing the critical gap between existing binary parasite detection systems and the multi-class severity classification needed to guide clinical treatment decisions. The research followed a hybrid OOSAD and CRISP-DM methodology, integrating computer vision (CNN) for blood smear image analysis with classical machine learning models for clinical symptom evaluation, fused through a decision fusion layer into a coherent diagnostic output. The study also demonstrates that the integration of machine learning with clinical informatics principles can yield a reliable, usable, and clinically valid tool for malaria severity assessment. The proposed CDSS achieves diagnostic accuracy exceeding comparable related works while addressing the critical gap between binary detection and actionable multi-class severity classification. The system empowers frontline healthcare workers in resource-constrained endemic settings with expert-level diagnostic support, contributing meaningfully to global efforts to reduce malaria-related mortality and morbidity, as aligned with the World Health Organization's malaria control priorities.

REFERENCES

Adeola, A. M., Botai, J. O., Adisa, O. M., & Ncongwane, K. P. (2021). Impact of future climate change on malaria in West Africa. *Theoretical and Applied Climatology, 145*(1–2), 1–14. https://doi.org/10.1007/s00704-021-03807-6

Ahiwar, R., Sharma, A., Gupta, R., & Singh, P. (2019). Deep learning for smartphone-based malaria parasite detection in thick blood smears. *IEEE Journal of Biomedical and Health Informatics, 23*(5), 1910–1918. https://doi.org/10.1109/JBHI.2019.2939121

Bairagi, V. K., & Charpe, K. C. (2016). Comparison of texture features used for classification of life stages of malaria parasite. *International Journal of Biomedical Imaging, 2016*, 7214156. https://doi.org/10.1155/2016/7214156

Chibuta, S., & Acar, A. C. (2020). Real-time malaria parasite screening in thick blood smears for low-resource setting. Journal of digital imaging, 33(3), 763-775.

D'Abramo, A., Rinaldi, F., Vita, S., Mazzieri, R., Corpolongo, A., Palazzolo, C., & Nicastri, E. (2024). A machine learning approach for early identification of patients with severe imported malaria. Malaria Journal, 23(1), 46.

Dehghani M. A., Dehghan, H., Kimiafar, K., Karami, M., Mousavi Baigi, S. F., & Farsi Mehrabadi, M. (2025). Usability evaluation methods for hospital information systems: a systematic review. BMC Health Services Research.

Dong, Y., Jiang, Z., Shen, H., Pan, W. D., Williams, L. A., Reddy, V. V., and Bryan, A. W. (2017, February). Evaluations of deep convolutional neural networks for automatic identification of malaria infected cells. In 2017 IEEE EMBS international conference on biomedical & health informatics (BHI) (pp. 101-104). IEEE.

Hoyos, K., & Hoyos, W. (2024). Supporting malaria diagnosis using deep learning and data augmentation. Diagnostics, 14(7), 690. https://doi.org/10.3390/diagnostics14070690

Kalipe, G., Gautham, V., & Behera, R. K. (2018). Predicting malarial outbreak using machine learning and deep learning approach: a review and analysis. In 2018 international conference on information technology (ICIT). 33-38. IEEE.

Kim, Y., Ratnam, J. V., Doi, T., Morioka, Y., Behera, S. K., & Yamagata, T. (2019). Malaria predictions based on seasonal climate forecasts in South Africa. *Scientific Reports*, *9*, 17865. https://doi.org/10.1038/s41598-019-53838-3

Loh, D. R., Yong, W. X., Yapeter, J., Subburaj, K., & Chandramohanadas, R. (2021). A deep learning approach to the screening of malaria infection using Mask R-CNN. *Computerized Medical Imaging and Graphics*, *89*, 101845. https://doi.org/10.1016/j.compmedimag.2020.101845

Manescu, P., Shaw, M. J., Elmi, M., Li, X., & Brown, A. (2020). Expert-level automated malaria diagnosis on routine blood films with deep neural networks. *American Journal of Hematology*, *95*(8), 883–891. https://doi.org/10.1002/ajh.25827

Marques, G., Ferreras, A., & de la Torre-Diez, I. (2022). An ensemble-based approach for automated medical diagnosis of malaria using EfficientNet. *Multimedia Tools and Applications*. https://doi.org/10.1007/s11042-02212624-6

Morang'a, C. M., Amenga–Etego, L., Bah, S. Y., Appiah, V., Amuzu, D. S., Amoako, N., & Otto, T. D. (2020). Machine learning approaches classify clinical malaria outcomes based on haematological parameters. BMC medicine, 18(1), 375.

Olusola, S. M., Abiodun, G. J., & Ojo, O. T. (2020). Modelling of malaria incidence in Akure, Nigeria: Negative binomial approach. *GeoJournal*, *85*(5), 1339–1352. https://doi.org/10.1007/s10708-019-10134-x

Silka, W., Wieczorek, M., Siłka, J., & Woźniak, M. (2023). Malaria detection using advanced deep learning architecture. *Sensors*, *23*(3), 1501. https://doi.org/10.3390/s23031501

Thakur, M., Singh, R., Patel, A., & Kumar, S. (2019). Artificial neural network based prediction of malaria abundances using big data. *Clinical Epidemiology and Global Health*, *7*(1), 121–126. https://doi.org/10.1016/j.cegh.2018.03.003

Tek, F. B., Dempster, A. G., & Kale, I. (2009). Computer vision for microscopy diagnosis of malaria. *Malaria Journal*, *8*(1), 153.

Tompkins, A. M., Colón-González, F. J., Di Giuseppe, F., & Namanya, D. B. (2019). Dynamical malaria forecasts are skillful at regional and local scales in Uganda up to 4 months ahead. *GeoHealth*, *3*(3), 58–66. https://doi.org/10.1029/2018GH000157

Vijayalakshmi, M. (2022). Study on malaria detection using transfer learning. *International Journal for Research in Applied Science and Engineering Technology*, *10*(6), 47613. https://doi.org/10.22214/ijraset.2022.47613

World Health Organization (2023). World malaria report 2023. ISBN (electronic): 978-92-40086173.

Zinszer, K., Verma, A. D., Charland, K., Brewer, T. F., Brownstein, J. S., & Sun, Z. (2015). Forecasting malaria in a highly endemic country using environmental and clinical predictors. *Malaria Journal*, *14*, 245. https://doi.org/10.1186/s12936-015-0758-4