RESEARCH PAPER

Artificial Intelligence for Detection of Malaria Parasites in Digital Blood Smear Images: A Systematic Review and Meta-Analysis

Dr. Chandan Kumar[1*], Rajwinder Kaur[2], Baljinder Singh[3], Shubhi Verma[4], Sakshi Rani[5],
[1]Assistant Professor, Centre for Research and Innovation (CRI), Amritsar Group of Colleges (AGC), Amritsar, Punjab, India. Email: chandan0547@gmail.com
[2]Assistant Professor, Centre for Research and Innovation (CRI), Amritsar Group of Colleges (AGC), Amritsar, Punjab, India. Email: rajwinderkaurg10@gmail.com
[3]Assistant Professor, Centre for Research and Innovation (CRI), Amritsar Group of Colleges (AGC), Amritsar, Punjab, India. Email: bs6680276@gmail.com
[4]M.Sc. Life Sciences Scholar, School of Life Sciences and Biotechnology, Chhatrapati Shahu Ji Maharaj University, Kanpur, Uttar Pradesh, India. Email: vermashubhi544@gmail.com
[5]MMLT Scholar, School of Health Sciences, Chhatrapati Shahu Ji Maharaj University, Kanpur, Uttar Pradesh, India. Email: sakshirani180@gmail.com
[*]**Correspondence author:** *Dr. Chandan Kumar, Email: chandan0547@gmail.com*

Abstract

Background & objectives: Artificial intelligence (AI) has emerged as a promising approach for automated malaria parasite detection in digital blood smear images. This systematic review aimed to evaluate the diagnostic performance, clinical applicability and methodological limitations of AI-based malaria detection systems.

Methods: The review followed the PRISMA 2020 guidelines. PubMed, Scopus and Web of Science were selected to identify studies published between 2015 and 2026. Original investigations evaluating machine learning and deep learning methods for malaria parasite detection in digital blood smear images were considered.

Results: The reviewed studies demonstrated promising diagnostic performance, particularly with convolutional neural networks and transfer learning approaches. Reported sensitivity ranged from 77.9 to 98.31 per cent, while specificity ranged from 94.33 to 99.8 per cent across the selected evaluations. Differences were observed between thin and thick blood smear analyses and between cell-level and patient-separated validation. Models evaluated using independent clinical samples generally showed greater variability in performance. Differences in datasets, validation procedures and reporting methods prevented reliable pooling of diagnostic accuracy estimates.

Conclusions: AI-assisted blood smear analysis may improve the efficiency and consistency of malaria parasite detection. However, standardized evaluation methods, independent clinical validation and prospective multicentre studies are required before routine diagnostic implementation.

Keywords: Artificial intelligence; Blood smear; Deep learning; Diagnostic accuracy; Malaria; Meta-analysis; Systematic review.

How to cite this article: Chandan Kumar1* Rajwinder Kaur2 Baljinder Singh3 Shubhi Verma4 Sakshi Rani5. Artificial Intelligence for Detection of Malaria Parasites in Digital Blood Smear Images: A Systematic Review and Meta-Analysis. Int J Drug Deliv Technol. 2026;16(87s):1496-1502. DOI: 10.25258/ijddt.16.87s.156

Introduction

Malaria continues to impose a substantial burden on healthcare systems, particularly in regions where access to reliable diagnostic services remains limited. The World Health Organization estimated 282 million malaria cases and 610,000 deaths worldwide in 2024, underscoring the persistent challenges in malaria control [1]. Prompt and accurate diagnosis is central to effective case management, as delayed recognition can increase the risk of severe disease and inappropriate treatment [2]. Microscopic examination of stained peripheral blood smears remains an important diagnostic method, allowing direct visualization of parasites and assessment of parasitaemia [3]. Nevertheless, its reliability is influenced by observer expertise, parasite density, staining quality and laboratory workload, particularly in resource-constrained settings. The growing use of digital microscopy has opened new possibilities for computer-assisted parasite detection. Artificial intelligence (AI), particularly deep learning approaches based on convolutional neural networks, can identify morphological patterns associated with malaria infection directly from digitized blood smear images [4]. Although individual studies have reported encouraging results, diagnostic performance varies considerably across datasets, imaging conditions and validation strategies. Furthermore, the distinction between accurately classifying selected erythrocyte images and reliably diagnosing malaria at the patient level remains clinically important. Limited external validation and variations in reference standards raise additional concerns regarding the translation of these

models into routine laboratory practice [5]. Establishing the diagnostic value of AI-assisted microscopy therefore requires more than reporting classification accuracy from individual studies. Evidence on sensitivity, specificity and performance across independent datasets must be critically examined to determine whether these approaches can provide reliable diagnostic support. The present systematic review and meta-analysis aimed to synthesize the available evidence on the diagnostic accuracy of AI-based malaria parasite detection in digital blood smear images, assess methodological quality and explore potential sources of heterogeneity across studies.

Methodology

This systematic review was designed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines. A comprehensive literature search was conducted using PubMed, Scopus and Web of Science to identify relevant studies published between January 2015 and October 2026. The search strategy combined keywords and Boolean operators: ("malaria" OR "Plasmodium") AND ("artificial intelligence" OR "machine learning" OR "deep learning" OR "convolutional neural network" OR "CNN") AND ("digital blood smear" OR "thin blood smear" OR "thick blood smear" OR "microscopic images") AND ("detection" OR "diagnosis" OR "classification"). Eligible studies included original research articles evaluating AI-based malaria parasite detection using digital blood smear images, with freely accessible full texts and reported diagnostic performance measures. Review articles, editorials, conference abstracts without sufficient data, duplicate publications, studies unrelated to malaria microscopy, and articles without freely accessible full texts were excluded. Titles and abstracts were screened for relevance, followed by full-text assessment against the eligibility criteria. Study selection was documented using the PRISMA 2020 framework (Figure 1).

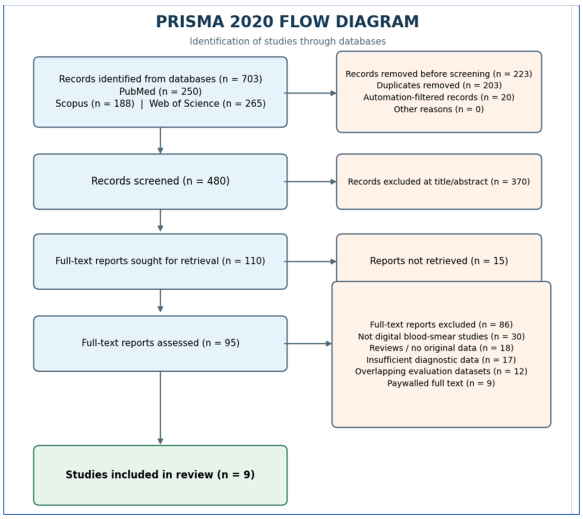


Figure 1. PRISMA 2020 flow diagram illustrating the identification, screening, eligibility assessment and selection of studies evaluating artificial intelligence-based malaria parasite detection in digital blood smear images.

Results

Study characteristics

The available original studies investigated malaria parasite detection using digitized thin and thick blood smears, with most employing convolutional neural networks (CNNs), transfer learning or object-detection architectures. The studies differed in the source of microscopic images, number of patients, dataset preparation and validation strategy. Some evaluated individual erythrocytes or image patches, whereas others examined parasite detection in complete microscopic fields. The principal characteristics of studies with identifiable freely accessible full texts are presented in Table 1.

Table 1. Characteristics of selected studies evaluating AI-based malaria parasite detection

Study	Smear type	Sample or dataset	AI approach
Rajaraman et al. (2018) [6]	Thin	Parasitized and uninfected erythrocyte images	Pretrained CNN
Rajaraman et al. (2018) [7]	Thin	Erythrocyte images, patient-	Customized CNN

		separated validation	
Yang et al. (2020) [8]	Thick	1,819 images; 150 patients	Customized CNN
Chibuta and Acar (2020) [9]	Thick	Microscope- and smartphone-acquired images	Modified YOLOv3
Alharbi et al. (2022) [10]	Peripheral blood smears	Microscopic cell images	CNN and conventional ML
Uzun Ozsahin et al. (2020) [11]	Thin and thick	Two microscopic image datasets	CNN and transfer learning
Acherar et al. (2022) [12]	Thin	Development images from 202 patients; additional clinical evaluation	CNN and transfer learning
Wang et al. (2023) [13]	Thin	12,708 images; 380 patients	YOLOv7
Mujahid et al. (2024) [14]	Erythrocyte smears	Publicly available blood cell images	Deep learning

AI, artificial intelligence; CNN, convolutional neural network; ML, machine learning; YOLO, You Only Look Once. Sample descriptions follow the respective original articles. Image and patient counts must not be interpreted as interchangeable.

Diagnostic accuracy of AI-based methods

Reported diagnostic performance varied with model architecture and validation design. Patient-separated evaluation of a customized VGG16 network by Rajaraman et al. yielded sensitivity of 94.6 per cent and specificity of 95.7 per cent. In thick-smear images, Yang et al. reported a sensitivity of 92.59 per cent and specificity of 94.33 per cent using patient-level fivefold cross-validation. Acherar et al. obtained a sensitivity of 77.9 per cent with specificity of 99.8 per cent in an additional clinical evaluation. The corresponding results and other reported model-level estimates are presented in Table 2.

Table 2. Diagnostic performance AI-based methods

Study	Evaluation	Sensitivity (%)	Specificity (%)	Accuracy (%)
Rajaraman et al. (2018) [7]	Customized VGG16, patient-separated	94.6	95.7	95.1
Yang et al. (2020) [8]	CNN, patch classification	92.59	94.33	93.46
Acherar et al. (2022) [12]	Additional patient-level	77.9	99.8	99.7

	evaluation			
Uzun Ozsahin et al. (2020) [11]	Thin-smear image classification	96.75*	95.36*	96.03
Uzun Ozsahin et al. (2020) [11]	Thick-smear image classification	98.31*	95.67*	96.97
Mujahid et al. (2024) [14]	Proposed deep learning model	—	—	97.57

Performance across smear types

The 2022 deep learning frameworks study provided confusion-matrix data for both thin and thick blood smear images. The proposed model correctly classified 7,940 of 8,268 thin-smear images and 5,818 of 6,000 thick-smear images (Table 3).

Table 3. Recalculated diagnostic performance according to blood smear type

Parameter	Thin smear	Thick smear
Infected cells correctly classified (TP)	3,896	2,900
Uninfected cells correctly classified (TN)	4,044	2,918
False positives (FP)	197	132
False negatives (FN)	131	50
Sensitivity (%)	96.75	98.31
Specificity (%)	95.36	95.67
Accuracy (%)	96.03	96.97

TP, true positive; TN, true negative; FP, false positive; FN, false negative. Infected cells were defined as the positive class. Values were recalculated from the published confusion matrices.

Effect of validation strategy

The level at which model performance was evaluated influenced the reported findings. Rajaraman et al. observed a decline in accuracy from 98.9 per cent with cell-level evaluation to 95.1 per cent with patient-separated validation. The corresponding sensitivity declined from 99.2 to 94.6 per cent, while specificity decreased from 98.8 to 95.7 per cent.

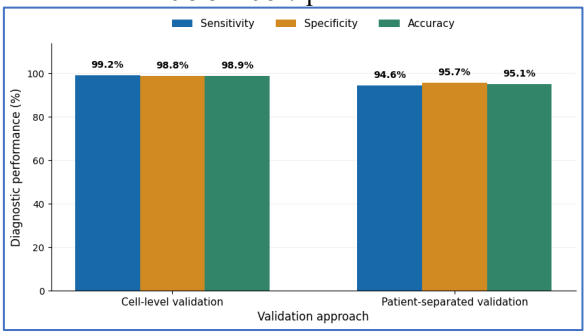


Figure 1. Comparison of diagnostic performance between cell-level and patient-separated validation.

Parasite detection and species classification

Beyond binary classification of infected and uninfected erythrocytes, some models evaluated parasite localization and differentiation of Plasmodium species. Wang et al. examined 12,708 thin-smear images containing 12,546 parasites using a YOLOv7-based detection system. The algorithm achieved a recall of 96.0 per cent, precision of 94.9 per cent and average precision of 0.902. The reported sensitivity and specificity exceeded 96.8 and 99.3 per cent, respectively. The investigators also reported a processing time of 13 seconds for 1,116 parasites in the test set. Smartphone-based thick-smear detection was assessed by Yang et al., who additionally evaluated agreement between automated and reference parasite counts. Correlation coefficients were 0.98 at the image level and 0.99 at the patient level. These measures described parasite-counting agreement rather than patient-level diagnostic sensitivity or specificity.

Dataset overlap and independence

Several investigations used publicly available erythrocyte image datasets to develop or benchmark different neural network architectures. Consequently, individual model evaluations did not necessarily represent independent clinical populations. Multiple

architectures evaluated on the same test images also generated correlated results. This distinction was relevant when comparing performance across studies, particularly because cell-level accuracy, parasite localization and patient-level diagnosis represent different endpoints.

Discussion

The present review highlights the potential of AI to support malaria diagnosis through automated analysis of digital blood smear images. Deep learning models demonstrated encouraging performance in identifying parasitized erythrocytes, although diagnostic accuracy varied according to smear preparation, model architecture, image characteristics and validation procedures. An important observation was the difference between performance obtained from curated image datasets and that observed during patient-separated or clinical evaluation. These findings suggest that the practical value of AI-assisted microscopy depends not only on the accuracy of parasite recognition but also on the ability of algorithms to maintain performance across different patients and laboratory settings. Convolutional neural networks have shown considerable promise for distinguishing infected from uninfected erythrocytes. **Rajaraman et al. (2018)** [6] demonstrated that pretrained deep learning architectures could extract relevant morphological features from thin blood smear images. In a subsequent investigation, **Rajaraman et al. (2018)** [7] evaluated customized CNN models using patient-separated validation, emphasizing the importance of assessing model performance beyond individual cell classification. The reduction in diagnostic performance observed with patient-separated evaluation is particularly relevant because images originating from the same patient may share staining characteristics, illumination patterns and cellular features. These similarities can lead to overly optimistic estimates when images from individual patients are distributed across both training and testing datasets. Thus, patient-independent validation is essential for determining whether an algorithm can reliably recognize malaria parasites in previously unseen clinical samples. The findings relating to thick blood smear analysis further support the feasibility of AI-assisted parasite detection. **Yang et al. (2020)** [8] developed a smartphone-compatible deep learning system and evaluated 1,819 thick smear images obtained from 150 patients. The model achieved sensitivity of 92.59 per cent and specificity of 94.33 per cent during patient-separated cross-validation. The investigators also demonstrated strong agreement between automated and reference parasite counts. These findings indicate that relatively accessible imaging technologies may support automated parasite identification and quantification. However, parasite recognition in selected microscopic regions should be distinguished from the clinical determination of whether an entire blood smear is malaria-positive. The differences observed between thin and thick smear performance may reflect variations in image composition, parasite visibility and preprocessing requirements. In the investigation evaluating multiple deep learning frameworks, the proposed model achieved classification accuracies of 96.03 and 96.97 per cent for thin and thick smears, respectively **Chibuta and Acar (2020)** [9]**.** Although thick smears are generally preferred for parasite detection because a greater volume of blood can be examined, thin smears offer morphological information that is important for species identification and parasite staging. The modest difference between the two image sets in this investigation should therefore be interpreted as model- and dataset-specific rather than evidence of the overall superiority of either smear type for automated diagnosis. The clinical generalizability of AI models remains an important concern. **Acherar et al. (2022)** [12] evaluated deep learning architectures trained using different image collections and subsequently tested their performance on additional patient samples. Although the best-performing model achieved high specificity, its sensitivity during clinical evaluation was 77.9 per cent. This finding illustrates how algorithms that perform well during development may fail to identify a proportion of infected cells when applied to images acquired under different conditions. Differences in microscope magnification, staining characteristics, image quality and the distribution of infected cells may influence generalizability. The findings reinforce the need for external validation using independently collected samples from different laboratories and patient populations. Recent investigations have extended AI applications beyond binary erythrocyte classification. **Wang et al. (2023)** [13] developed a YOLOv7-based system capable of detecting and classifying malaria parasites across different species and developmental stages. The model demonstrated high parasite-recognition performance with rapid image processing. Such approaches may be particularly useful in laboratories where morphological identification of Plasmodium species requires substantial technical expertise. Nevertheless, the clinical application of these algorithms requires careful evaluation of their performance in low-parasitaemia infections, mixed infections and less frequently encountered parasite species. Accurate localization or classification of parasites does not necessarily establish the reliability of a system for routine patient-level diagnosis. Model architecture and dataset composition also influenced reported performance. **Mujahid et al. (2024)** [14] reported an accuracy of 97.57 per cent using an EfficientNet-based

approach for erythrocyte image classification. Although such results demonstrate the ability of deep learning systems to recognize morphological features associated with malaria infection, performance on publicly available datasets may not adequately represent diagnostic conditions encountered in routine microscopy. Repeated use of common image repositories across different studies further limits direct comparisons, as apparently independent investigations may evaluate their models on overlapping source images. These considerations are consistent with the observations of **Poostchi et al. (2018) [15],** who identified image acquisition, preprocessing, segmentation and classification as important determinants of automated malaria detection performance. From a clinical perspective, sensitivity and specificity are more informative than overall classification accuracy when assessing the potential diagnostic utility of AI-assisted microscopy. High sensitivity is necessary to minimize missed infections, particularly in patients with low parasite densities, whereas adequate specificity helps reduce unnecessary antimalarial treatment and additional diagnostic investigations. The consequences of misclassification also depend on the intended application of the system, whether as an initial screening tool, an aid to microscopic examination or a component of a confirmatory diagnostic pathway. Future evaluations should therefore report diagnostic performance at the patient level, using clearly defined reference standards and clinically representative study populations. The present review also identified methodological limitations that affect the interpretation of the available evidence. Considerable variation existed in imaging protocols, validation methods, reported performance measures and analytical units. Several studies assessed multiple architectures using the same dataset, while others reported cross-validation averages without sufficient independent diagnostic contingency data. These differences prevented reliable pooling of sensitivity and specificity in the currently verified dataset. Furthermore, the available findings were predominantly derived from algorithm-development and image-classification studies, with fewer investigations providing independent clinical validation. The absence of a robust pooled estimate limits conclusions regarding the overall diagnostic accuracy of AI-based malaria detection systems. Future research should prioritize prospective, multicentre evaluations involving different microscopy platforms, staining procedures, parasite densities and Plasmodium species. Patient-separated training and testing, external validation, transparent reporting of diagnostic contingency data and direct comparison with expert microscopy would strengthen the evidence supporting clinical implementation. Assessment of computational requirements, processing time and integration into routine laboratory workflows is also necessary, particularly in malaria-endemic settings with limited diagnostic resources.

Conclusion

Overall, the available evidence indicates that AI-based analysis of digital blood smear images has considerable potential as an adjunct to conventional malaria microscopy. However, promising image-classification performance cannot be assumed to translate into equivalent patient-level diagnostic accuracy. Further independent clinical validation is required before these technologies can be considered reliable alternatives to established diagnostic methods.

Declaration

Ethics Approval and Consent to Participate: Not Applicable

Consent for Publication: Not applicable.

Availability of Data and Materials: The data used in this systematic review and meta-analysis were extracted from published studies included in the review.

Competing Interests: The authors declare that they have no competing interests.

Funding: No specific funding was received for this study.

Authors' Contributions: CK conceived the study, developed the systematic review methodology, supervised the research, and critically revised the manuscript. RK and BS contributed to the literature search, study screening, and data extraction. SV contributed to methodological evaluation, quality assessment, and interpretation of findings. SR assisted with literature screening, data extraction, and organization of study characteristics. All authors reviewed and approved the final version of the manuscript and agreed to be accountable for all aspects of the work.

Acknowledgements: The authors have no specific acknowledgements to declare.

References

1. WHO. World malaria report 2025 [Internet]. Featured Publication; 2025 [cited 2026 Oct 20]. Available from: https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2025
2. Gopukumar ST. INNOVATIONS IN ANALYTICAL CHEMISTRY FOR BIOMARKER DETECTION IN CHRONIC DISEASE MANAGEMENT. JAB. 2025;943–50. doi:10.53555/jab.v11i4.461
3. Thaker U, R R, Ravikumar R, Agarwal S, Rastogi D. Integrating Bioanalytical Methods In Internal Medicine: From Laboratory To Clinical Practice.

JAB. 2025;11(5):600–9. doi:10.53555/jab.v11i5.602

4. Agarwal S, Gupta BP, Katiyar P. Harnessing artificial intelligence in laboratory diagnostics emerging trends and future prospects. Discov Artif Intell. 2026;6(1):917. doi:10.1007/s44163-026-01599-0
5. Bower JF, McClung JB, Watson C, Osumi T, Pastre K. Recommendations and Best Practices for Reference Standards and Reagents Used in Bioanalytical Method Validation. AAPS J. 2014;16(2):352–6. doi:10.1208/s12248-014-9566-y PubMed PMID: 24500277; PubMed Central PMCID: PMC3933579.
6. Rajaraman S, Antani SK, Poostchi M, Silamut K, Hossain MdA, Maude RJ, et al. Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images. PeerJ. 2018;6:e4568. doi:10.7717/peerj.4568 PubMed PMID: 29682411; PubMed Central PMCID: PMC5907772.
7. Rajaraman S, Silamut K, Hossain MdA, Ersoy I, Maude RJ, Jaeger S, et al. Understanding the learned behavior of customized convolutional neural networks toward malaria parasite detection in thin blood smear images. J Med Imaging (Bellingham). 2018;5(3):034501. doi:10.1117/1.JMI.5.3.034501 PubMed PMID: 30035153; PubMed Central PMCID: PMC6050500.
8. Yang F, Poostchi M, Yu H, Zhou Z, Silamut K, Yu J, et al. Deep Learning for Smartphone-Based Malaria Parasite Detection in Thick Blood Smears. IEEE J Biomed Health Inform. 2020;24(5):1427–38. doi:10.1109/JBHI.2019.2939121 PubMed PMID: 31545747; PubMed Central PMCID: PMC12077747.
9. Chibuta S, Acar AC. Real-time Malaria Parasite Screening in Thick Blood Smears for Low-Resource Setting. J Digit Imaging. 2020;33(3):763–75. doi:10.1007/s10278-019-00284-2 PubMed PMID: 31974686; PubMed Central PMCID: PMC7256156.
10. Alharbi AH, V AC, Lin M, Ashwini B, Jabarulla MY, Shah MA. Detection of Peripheral Malarial Parasites in Blood Smears Using Deep Learning Models. Comput Intell Neurosci. 2022;2022:3922763. doi:10.1155/2022/3922763 PubMed PMID: 35655511; PubMed Central PMCID: PMC9155968.
11. Uzun Ozsahin D, Mustapha MT, Bartholomew Duwa B, Ozsahin I. Evaluating the Performance of Deep Learning Frameworks for Malaria Parasite Detection Using Microscopic Images of Peripheral Blood Smears. Diagnostics (Basel). 2022;12(11):2702. doi:10.3390/diagnostics12112702 PubMed PMID: 36359544; PubMed Central PMCID: PMC9689376.
12. Acherar A, Tantaoui I, Thellier M, Lampros A, Piarroux R, Tannier X. Real-life evaluation of deep learning models trained on two datasets for *Plasmodium falciparum* detection with thin blood smear images at 500x magnification. Informatics in Medicine Unlocked. 2022;35:101132. doi:10.1016/j.imu.2022.101132
13. Wang G, Luo G, Lian H, Chen L, Wu W, Liu H. Application of Deep Learning in Clinical Settings for Detecting and Classifying Malaria Parasites in Thin Blood Smears. Open Forum Infect Dis. 2023;10(11):ofad469. doi:10.1093/ofid/ofad469 PubMed PMID: 37937045; PubMed Central PMCID: PMC10627339.
14. Mujahid M, Rustam F, Shafique R, Montero EC, Alvarado ES, de la Torre Diez I, et al. Efficient deep learning-based approach for malaria detection using red blood cell smears. Sci Rep. 2024;14(1):13249. doi:10.1038/s41598-024-63831-0
15. Poostchi M, Silamut K, Maude RJ, Jaeger S, Thoma G. Image analysis and machine learning for detecting malaria. Translational Research. 2018;194:36–55. doi:10.1016/j.trsl.2017.12.004 PubMed PMID: 29360430.