Bioinformatics Advances, YYYY, 0–0
doi: 10.1093/Bioinformatics Advances/xxxxx
Advance Access Publication Date: DD Month YYYY
Manuscript Category

Bioimage Informatics

An interactive mindmap of blood film images for automated malaria diagnosis: a comprehensive metadata repository of available datasets

Itunuoluwa Isewon[1,2,3,*], Faith Adegoke[1,2,3], Charles B. Delahunt[4], and Jelili Oyelade[1,2,3]

[1]Department of Computer and Information Sciences, Covenant University, P.M.B. 1023, Ota, Ogun State, Nigeria

[2]Covenant Applied Informatics and Communication African Centre of Excellence (CApIC-ACE), Covenant University, P.M.B. 1023, Ota, Ogun State, Nigeria

[3]Covenant University Bioinformatics Research (CUBRe), Covenant University, P.M.B. 1023, Ota, Ogun State, Nigeria

[4]Global Health Labs, Bellevue, WA, United States

*Corresponding author. Department of Computer and Information Sciences, Covenant University, P.M.B. 1023, Ota, Ogun State, Nigeria. E-mail:itunu.isewon@covenantuniversity.edu.ng

Associate Editor: XXXXXXX

Abstract

Summary: Malaria remains a significant public health challenge, necessitating innovative diagnostic solutions to combat high mortality rates. Here, we conducted a systematic review of malaria microscopy datasets published between 2014 and 2024 by querying Scopus and PubMed databases. We have a curated inventory of 71 malaria blood film image datasets. To further enhance usability, we developed an interactive mind map that visually represents the datasets based on five key features. The **Access** category differentiates between open and controlled access datasets. The **Smear Type** category classifies datasets by blood smear type. The **Stain Type** category organizes datasets by the staining methods used. The **Species Type** highlights the *Plasmodium* species present in the images, while the **Demography** category identifies the geographic origin of datasets. Each dataset has a dedicated page with a detailed description, a citation of the dataset's publication, and links to databases where it can be downloaded. This mind map has been carefully curated with up-to-date information and will be regularly updated as new datasets become available, ensuring its continued relevance and accuracy.

The mind map is a freely accessible resource, available at https://itunuisewon.github.io/Malaria_Blood_Film_Images/, which enables researchers to identify and utilize diverse datasets for training and validating machine learning algorithms.

Availability and implementation: Not applicable.

Keywords: Image analysis, Machine learning, Artificial intelligence, Medical informatics

1 Introduction

Computer-aided diagnosis (CAD) uses computers and computational tools to analyse and interpret medical data for disease diagnosis. Recent studies have highlighted the efficacy of CAD techniques in enhancing diagnostic accuracy. Such studies involve interdisciplinary collaboration among expert pathologists, computer scientists, and mathematicians (Molina *et al.,* 2021). Furthermore, these methods have also significantly

reduced the manual inspection workload for healthcare professionals by automating various aspects of data analysis.

Malaria remains a public health concern globally (Olasehinde *et al.,* 2019). It is caused by the *Plasmodium* parasite, which is transmitted to humans through the bite of female *Anopheles* mosquitoes. Among the species that infect humans, five variants are identified: *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium ovale*, *Plasmodium malariae*, and *Plasmodium knowlesi* (Oyewola *et al.,* 2022). There has been a concerning increase in malaria cases across 85 nations due to the COVID-19 pandemic, which disrupted healthcare systems (Benjamin *et al.,* 2023; G *et al.,* 2023). The 2024 World Health Organization (WHO) malaria report records 263 million cases and 597,000 deaths in 2023, with Africa bearing the majority of the impact. Children under the age of five are among the groups that are at high risk of malaria infection (WHO, 2024).

Early detection is vital for malaria diagnosis to ensure timely and appropriate treatment. However, diagnosing malaria can be particularly challenging in non-endemic regions, where clinicians may not initially consider it as a potential diagnosis due to limited familiarity with the disease, leading to misdiagnosis and inadequate treatment. Accurate detection of malarial parasites in blood samples requires expertise and experience, which laboratory technicians in these regions may lack, increasing the likelihood of diagnostic errors (Rehman *et al.,* 2020). In endemic regions, healthcare providers face a heavy workload, which can compromise their diagnostic efficiency and accuracy (Rees-Channer *et al.,* 2023). Malaria diagnosis typically includes methods such as microscopy, rapid diagnostic tests (RDTs), clinical evaluations, and molecular techniques like polymerase chain reaction (PCR) (Ojurongbe *et al.,* 2013). However, each of these methods has its drawbacks and limitations: PCR requires costly equipment and infrastructure, while RDTs have high detection thresholds and may fail to identify certain *Plasmodium* strains with HRP2 deletions (Alnussairi & İbrahim, 2022; Prosser *et al.,* 2021).

Globally, microscopy is a well-established technique for malaria diagnosis (Aqeel *et al.,* 2022). Despite its reliability, this method is time-consuming and labour-intensive, and the number of skilled professionals using it is insufficient to address the demands imposed by the disease, particularly in resource-constrained settings. Employing computer-aided diagnostic tools can accelerate the process and provide parasitologists with an opportunity for a second opinion. One key advantage of automated malaria microscopy is in quantifying parasitaemia (parasite density), enabling the calculation of clearance curves that help monitor drug-resistant strains of the *Plasmodium* parasite (Delahunt *et al.,* 2024). While numerous algorithms have been proposed for detecting malaria in Giemsa-stained blood films, only a small fraction have undergone field validation, and none to date have been integrated into clinical practice.

This paper provides a comprehensive catalogue of publicly available malaria image datasets to aid researchers in building reliable and effective automated malaria diagnosis models. The following sections first detail the microscopy method for malaria diagnosis and then present a compilation of malaria blood film datasets derived from a systematic review of existing literature.

2 Malaria Microscopy

The process of malaria microscopy involves examining stained thick and thin blood films under a microscope to detect and identify malaria parasites. The sensitivity of this technique is reduced in cases of low parasitemia, emphasising the necessity for skilled personnel to handle the staining process properly. Furthermore, microscopy demands considerable time and labour.

2.1 Blood sample collection

Blood samples are typically collected via a finger prick (or venipuncture in specific circumstances). A sterile lancet is used to puncture the third finger of the patient's hand, following alcohol swabbing to clean the site. A drop of blood is then transferred to a slide to prepare thick and thin films. In cases requiring venipuncture, a larger sample of venous blood is drawn and stored in tubes containing anticoagulants, subsequently used for preparing the blood films (WHO, 2016a).

2.2 Film preparation

Blood samples from the patients are used to prepare thick and thin blood films for malaria diagnosis using microscopy. In a thin film, red blood cells are arranged in a single layer, allowing for the observation of the parasite within the red blood cells (RBCs). This film is fixed with methanol and is useful for identifying the species of malaria parasite, determining life cycle stages, and quantifying high levels of parasitemia. However, its sensitivity decreases at low parasite densities. The distinction between RBCs and other blood components is clearer due to improved visibility (Rehman *et al.,* 2018). To prepare the thin film, a "spreader" slide is used to spread the blood drop quickly and smoothly, creating a feathery edge.

A thick blood film is more sensitive in detecting parasites than a thin blood film. It is prepared by swirling the blood drop, resulting in a denser film that allows for a larger volume of blood to be examined. In this film, RBCs are ruptured, leaving only the white blood cells (WBCs), platelets, and parasites visible. One challenge with this method is distinguishing the parasites from other blood components, i.e., a higher frequency of artefacts (Maturana *et al.,* 2022). Unlike the thin film, it is not fixed in methanol and can also be used to quantify parasitemia (Musumeci, 2014). Figure 1 shows a pictorial representation of the thick and thin Giemsa-stained blood films.

Figure 1: Giemsa-stained blood film (Top: thick film. Bottom: thin film.)

2.3 Staining of slides

In the microscopy method for malaria diagnosis, histological staining is a crucial step in preparing tissue samples. Enhancing their contrast, makes them more visible and easier to examine under the microscope. The process of histological staining consists of five key stages: fixation, processing, embedding, sectioning, and staining (Horobin, 2011). The Romanowsky stain, widely used since its introduction in the 1890s, is one of the primary stains used in this process (Musumeci, 2014). Methylene blue binds to acidic molecules, while eosin binds to neutral or basic cellular compounds. As a result, the nuclear material (DNA) stains purple, while the cytoplasm stains blue. As red blood cells have no nucleus, while examining a clean blood film, only white blood cells and *Plasmodium* parasites contain nuclei, thus stain purple. The cytoplasm of *Plasmodium* parasites typically appears distinct and stains blue. Variants of the Romanowsky stain include Giemsa, Leishman, Field, May Grunwald-Giemsa (MGG), Wright, and Jaswant Singh-Bhattacharji stains, each with its unique modifications. The Giemsa stain is widely favoured for its stability and consistent quality across varying temperatures, while Leishman's stain is also commonly used in malaria diagnosis. The Field stain is user-friendly, has a short staining time, and is cost-effective. However, if used by inexperienced personnel, there is a higher risk of false positives, as improper staining can result in an increased number of artefacts (Janiesch *et al.,* 2021). Even with correct application, the resulting blood film can vary in appearance due to factors such as the stain's age, concentration, and other variables. The

Jaswant Singh-Bhattacharji (JSB) stain, developed in India, offers a quick and inexpensive solution. However, its tendency to fade over time makes it unsuitable for archived slides (Kalinin *et al.,* 2024).

2.4 Imaging

Malaria parasites are minuscule (with morphological details around 0.1μm), so a high-resolution system is required for accurate diagnosis. Stained slides are typically examined using an optical (light) microscope with oil immersion at 100× magnification for high-resolution imaging. Alternatively, a microscope with a 40× objective (NA = 0.75) may be used for lower-magnification overviews or when oil immersion is not required (Horning *et al.,* 2021). The image is then adjusted for optimal clarity. Microscopy images of malaria-infected thick and thin blood films are shown in Figures 2 and 3, respectively. The thick blood film can be used to detect parasite density, while the thin blood film shows the parasite morphology in detail and can be used to identify different species of the parasite.

Figure 2: Thick Giemsa-stained blood film[1].

Figure 3: Thin blood film field of Views. (a) an ideal field of view. (b) Clumped red blood cells. The green arrow shows a white blood cell, and the red arrows show red blood cells infected by the *Plasmodium* parasites. Other purple-stained objects are artefacts.[2]

2.5 Malaria Diagnosis

As outlined in the WHO malaria microscopy standard operating procedure (WHO, 2016c, 2016b), a drop of immersion oil should be applied to the thick blood film, with the 100× objective lens making contact with the oil. The fine adjustment knob is then used to focus the slide, which should be examined systematically. For a slide to be considered non-parasitized, at least 100 fields of view must be examined without the detection of a parasite. If parasites are found, they should be counted along with WBCs until 200 WBCs have been counted. The formula in Eqn1 below is used to calculate parasitemia:

$$\frac{Plasmodium\ parasite\ counted \times 8000}{WBC\ counted} = parasites/\mu l\ blood \qquad \text{Eqn1}$$

The thin blood film should be examined at its feathery edge to identify parasite species after reviewing the thick blood film.

3 Machine learning for Malaria Diagnosis

Machine learning algorithms and computer vision techniques have been leveraged to aid malaria diagnosis. Machine learning enables a computer program to optimise tasks and improve performance metrics through experience and iterative learning (Rabbani *et al.,* 2013). It has been extensively used in medical imaging, employing traditional methods such as image acquisition, pre-processing, segmentation, feature extraction, feature selection, and classification (Molina *et al.,* 2021). In the past, researchers have relied on hand-crafted features like histograms of gradients, Haralick's texture features, RGB values, co-occurrence matrices, shape descriptors such as area, perimeter, compactness ratio, eccentricity, and bending energy, along with local binary patterns (Poostchi *et al.,* 2018).

Additional methods include Support Vector Machine (SVM) (Linder *et al.,* 2014), Decision Tree (Rabbani *et al.,* 2013), Naive Bayes (Das *et al.,* 2015), Linear Discriminant (Park *et al.,* 2016), and K-Nearest Neighbour (KNN) (Malhotra *et al.,* 2017). Shallow machine learning models have also been used in image classification for malaria diagnosis. With advancements in deep learning (DL), artificial intelligence has significantly improved in overall capabilities and performance. Deep learning has gained popularity and has shown results in domains such as Natural Language Processing and Computer Vision (Haralampieva *et al.,* 2020). This is due to its automated feature extraction within vast data frameworks, enabling DL models to learn autonomously, much like the human brain.

Among the various deep learning methods used for automatic diagnosis, the Convolutional Neural Network (CNN) is particularly prevalent. CNN excels in image classification due to its speed and efficiency. It outperforms other methods by adopting a hierarchical learning structure that can be robustly trained once the model structure is matched to the input features. Compared to traditional classifiers, CNN is a flexible learning framework that does not require prior feature extraction. The CNN architecture is constructed with the convolutional, pooling, and fully connected (FC) layers (Pan *et al.,* 2018). The input images are first processed through the convolutional layer, where convolution with a specific N × N filter extracts relevant image features. The pooling layer, usually positioned between two convolutional layers, reduces the image dimensions. Finally, the FC layer connects neurons based on the sum of the previous features to generate a specific output, often serving as a classifier (Hcini *et al.,* 2022). CNN architectures commonly used are VGGNet (Simonyan & Zisserman, 2015), GoogleNet (Szegedy *et al.,* 2014), ResNet (He *et al.,* 2015), and AlexNet (Krizhevsky *et al.,* 2012).

The ring stage of *Plasmodium falciparum* is a primary target in drug resistance research because it is during this stage that the parasite is most vulnerable to key antimalarial drugs. Understanding the mechanisms of drug resistance at this stage, such as mutations in the Pfkelch13 gene associated with artemisinin resistance, is critical for developing effective treatment strategies. (Delahunt *et al.,* 2022; Mehanian *et al.,* 2017) developed a framework using the ring stages of *Plasmodium* for parasite quantitation. They also categorised all late stages together for species identification. For optimal processing and precise object detection, a combination of techniques was used in a series, beginning with an initial detector. This distractor filter uses basic, manually designed features in conjunction with a CNN. The resulting output is a combination of patient-level quantitation and species identification predictions. (Manescu *et al.,* 2020) proposed a classifier called the Deep Malaria Convolutional Neural Network (DeepMCNN), which aggregates the number of *Plasmodium* parasites and white blood cells to estimate parasitaemia. With a Positive Predictive Value (PPV) of 92% and a Negative Predictive Value (NPV) of 90%, the system showed improved accuracy in detecting low parasitaemia levels, proving its effectiveness in adult malaria diagnosis.

In the research report by (Kassim *et al.,* 2021), Plasmodium VF-Net was proposed for malaria diagnosis on both patient and image levels using thick blood smears. It first determines the presence of infection and then identifies whether the parasite is P. falciparum or P. vivax. The study employed a Mask R-CNN (Regional-Convolutional Neural Network) to detect parasites and a ResNet50 classifier to minimise false positives. In the study conducted by (Yu *et al.,* 2023). The NLM Malaria Screener, an app developed for malaria diagnosis, was evaluated. The research was carried out in Sudan with 190 patients, and the Malaria Screener app was used to analyse Giemsa-stained blood smears, while

[1] Permission from Mehanian *et al.,* (Computer-automate malaria diagnosis and quantitation using convolutional neural networks, ICCV 2017) (Mehanian *et al.,* 2017).

[2] Permission from Delahunt *et al.,* (Full-automated patient-level malaria assessment on field-prepared thin blood film microscopy images, IEEE GHTC 2022) (Delahunt *et al.,* 2022).

expert microscopy and nested PCR were used as reference standards for comparison. The results indicated that the application achieved a 74.1% accuracy rate for detecting *Plasmodium falciparum*, meeting the WHO Level 3 criteria for parasite detection. Additionally, a novel algorithm named Plasmodium VF-Net achieved an 83.1% accuracy when compared to expert microscopy, and 81.0% accuracy when matched against PCR, meeting the WHO Level 2 criteria for detecting both *Plasmodium falciparum* and *Plasmodium vivax*. These findings suggest that both the Malaria Screener and Plasmodium VF-Net show promise as tools for malaria diagnosis, though further improvements are necessary to enhance their accuracy, identify malaria species, and count the number of parasites. For further insights into critical aspects of developing machine-learning models for malaria diagnosis, please see (Delahunt *et al.,* 2024).

Data plays a vital role in the development and performance of automated diagnostic systems for malaria. While machine learning models, especially deep learning frameworks, have shown promise in detecting and quantifying parasitemia, their accuracy and generalizability are heavily dependent on the quality and diversity of training data. A larger, well-annotated, and geographically diverse collection of microscopy image datasets is essential to improve model robustness.

4 Available Blood Film Images

Seeing the key factor data plays in developing this automated diagnostic system, a systematic review was conducted to look through publications in this domain using a defined search strategy. The review methodology adhered to the guidelines outlined in the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) protocol (Page *et al.,* 2021).

4.1 Eligibility Criteria

The eligibility criteria for the study, which determine the publications to be included, are shown in table 1. Each study from the structured search was assessed for image dataset inclusion.

TABLE 1

4.2 Search Strategy

To ensure comprehensive data collection, a search strategy was employed to craft precise search terms and pinpoint important databases for collecting relevant documents. Using a query string tailored to the study's objectives, articles were retrieved from Scopus and PubMed, as these databases host peer-reviewed content and other pertinent information. The search query included terms like "automated malaria diagnosis" and "malaria image dataset," with MeSH terms applied for PubMed searches. As illustrated in Figure 4, the search for datasets followed the PRISM guidelines, and the figure shows the search results and processing steps from 2014 to 2024. An overview of publications over the years of focus is shown in Figure 5.

Figure 4: PRISMA flowchart showing the screening process

Figure 5: Summary of all publications included in the study by year of publication

4.3 Summary of Identified Datasets

From the systematic review, 54,238 records were retrieved from searches conducted on PubMed and Scopus databases. After merging the records, 7,800 duplicates were identified and removed. Subsequently, 45,841 records were screened out based on their titles, narrowing down the list to 597 articles for further review. The screening process, based on the eligibility criteria provided in Table 1, resulted in 545 studies being included. An additional three records were added through a manual search. The study identified 71 blood film image datasets, summarized in Table 2.

TABLE 2

4.3.1 Stain Types

To identify malaria parasites, staining reagents were used to highlight cellular components in the selected studies. The staining reagents used ranged from Modified Romanowsky, Giemsa, May Grunwald-Giemsa, Leishman, and Field stains. The Giemsa stain accounts for 59% of all the datasets. Its stability and reliability make it the most widely utilised for identifying malarial parasites. Other reagents, such as Modified Romanowsky, May Grunwald-Giemsa, and Wright-Giemsa stains, were the least represented, as indicated in Figure 6d. The Giemsa stain is widely used due to its stability, and it is the recommended and most reliable stain for identifying malaria parasites (WHO, 2016c). Notably, 25% of the studies did not specify the staining reagent used.

4.3.2 Film Types

The analysis revealed that six datasets featured both thick and thin blood film images. Thin blood films were more commonly used, with forty-four (44) studies relying exclusively on this type, compared to 13 studies that used only thick blood films. This suggests a research preference for thin blood film images in developing the automated malaria diagnostic systems. The distribution of film types used in the studies is illustrated in Figure 6a.

4.3.3 Imaging Techniques

A range of equipment was used to capture the images for malaria diagnosis. Microscopes were the most common method, though some researchers attached smartphones to the microscope eyepiece to capture images. For example, (Rosado *et al.,* 2017) used a "mSmartScope" prototype paired with smartphones to capture their images. Other researchers attached cameras to microscopes for this task. For the image datasets featured in (Davidson *et al.,* 2021), six different microscopes were used by the research groups to obtain the images.

The statistics on the imaging techniques used are shown in Figure 6b, with further details on the methods provided in Table 3. There is a clear need for more images to be generated to advance research in this field, as indicated by the image datasets discovered. While deep learning models for automated malaria diagnosis require large datasets to train neural networks (Karasu Benyes *et al.,* 2022), it has been reported that the available datasets for malaria diagnosis are limited (Maqsood *et al.,* 2021). Furthermore, (Aqeel *et al.,* 2022) noted the need for more comprehensive repositories of annotated image data and the importance of collecting and standardising more data, given the various parasite species and the developmental stages.

This study also noted that some essential information about the image datasets was not provided by the researchers.

Figure 6: (a) Type of film used in selected studies, (b) Imaging techniques used in selected studies, (c) Distribution of datasets by continent, (d) Staining techniques used in selected studies.

TABLE 3

4.3.4 Data Availability

While some datasets discovered are publicly accessible, the availability of datasets from forty-two (42) studies remains unspecified. As previously mentioned, there is a clear need for publicly available and annotated image datasets to advance research in this domain. Table 4 outlines the data availability for each dataset, including the Digital Object Identifier (DOI) for each article. In cases where the DOI is unavailable, the article's title is listed in its place.

4.3.5 Study Location

The datasets identified originate from various regions, with twelve coming from Africa, while the remainder are from parts of Asia, South America, and Europe. Table 5 details the country, city, and study area as reported. As shown in Figure 6b, a large proportion of the datasets originated from Asia, while 12% of the dataset origins were not reported. 16% came from Africa. This highlights the need for more datasets from malaria-endemic regions, particularly Africa. This observation is substantiated by (Kassim *et al.,* 2021), who emphasised the importance of evaluating automated malaria diagnostic models using data from different global regions to assess and ensure their adaptability to diverse datasets. Variability in medical images, including the devices used for capturing, the lack of standardisation in medical practices, and patient genetic and phenotypic characteristics, along with parasite species and staining variations, influences the generalizability of machine learning models (Maleki *et al.,* 2023, Kuo *et al.,* 2020). The generalizability of these models is essential for real-world deployment and application in malaria diagnosis.

TABLE 4

TABLE 5

The systematic review carried out showed limitations in data availability, metadata consistency, and geographic representation of datasets. To address this, the next section introduces an open-source interactive webpage developed to be a resource for discovering diverse blood film image datasets for malaria diagnosis.

5 Data Accessibility and Interactive Mind Map

In automated malaria diagnosis, data accessibility remains a challenge. From the datasets discovered in the study, several datasets were either not available due to the loss of datasets or accessible only upon request. These data and metadata accessibility issues will be a challenge for reproducibility and the benchmarking of robust machine learning models. The National Library of Medicine have made most of their generated annotated image datasets accessible, and this includes the thick blood film *Plasmodium falciparum*-infected images from 150 patients, the thin blood film uninfected and *Plasmodium* falciparum-infected images from 193 patients, among others. Other image datasets can be found on Kaggle, Mendeley data, Zenodo and the research group's webpage. With all these efforts by the researchers, there is still a need for many image datasets to be made publicly available (Maturana et al., 2022).

In biomedical image analysis, data accessibility is very important. In the scientific community, the FAIR principles (Findability, Accessibility, Interoperability and Reusability) have been emphasised for data sharing and standardisation (Wilkinson, 2016). To follow these principles, some of the available datasets that have the license for redistribution have been hosted on Zenodo for easy access, and the link can be accessed at https://zenodo.org/records/16958076 This study also introduces an interactive mind map to facilitate datasets to be easily discovered by researchers in this domain.

The findings of this research are being presented interactively through a user-friendly mind map developed with the static Jekyll Cayman theme on GitHub pages. A dedicated repository "Malaria_Blood_Film_Images" was created to store markdown files detailing individual blood film datasets. The Jekyll Cayman theme was integrated into the repository, complete with its configuration, layout, and style files for webpage customisation. An interactive HTML landing page was also designed and deployed as a GitHub page. It categorises datasets by film type, *Plasmodium* species, staining method, imaging technology, geographical details and availability. It enables users to visually explore datasets and directly access their source URLs or DOIs. This approach provides a user-friendly alternative to traditional static tables, reducing search time and encouraging dataset reuse.

The structure of this repository allows for updates from the scientific community. Each dataset's metadata information has been documented in markdown files, allowing for easy revision and contributions by other users. The choice of GitHub for this is to ensure that the repository remains open source and can also benefit from version control (Russell et al., 2018). Just like was demonstrated by the MalariaScreener datasheet (National Library of Medicine, National Institutes of Health, Bethesda, 2020), Centralising these datasets' information will enhance their use for training A. I models.

As the field increasingly moves toward generalizable, cross-site diagnostic tools, access to a broad and well-documented pool of training data becomes essential. The mind map contributes to this goal by lowering the barrier for researchers, especially those in low-resource settings, to discover and leverage valuable image datasets for automated diagnosis development. The repository is publicly available and maintained at https://github.com/itunuisewon/Malaria_Blood_Film_Images, with the live interactive site accessible at https://itunuisewon.github.io/Malaria_Blood_Film_Images/.

A detailed guide for navigating and using the interactive mindmap webpage has been provided in Supplementary File 1. This includes a brief explanation of the nodes and how to interact with them.

6 Challenges and Future Direction

In the automated diagnosis of malaria, machine learning has seen significant progress, but there is a major need to translate all this progress into solutions that can be deployed in clinical settings.

From our study, one of the major limitations seen is the inconsistency in the formats of image datasets. Datasets vary in formats, magnification and even annotation types. This inconsistency is a major hindrance to the integration of images across studies, thereby obstructing reproducibility (Straiton, 2024). This lack of standardization is a result of differences in data collection, annotation and regional practices. Another challenge seen is the absence of some information, like imaging equipment, film type, and staining method, in some of the datasets. Having accurate metadata is important in understanding the data and training robust models. In addition, many of the datasets only provide binary labels, with just a few offering species differentiation or lifecycle-stage annotations.

The geographic bias is another major challenge. Many of the publicly available image datasets originate from a limited number of countries, mainly in Asia. Despite African countries bearing a high malaria burden, they remain underrepresented. This bias can reduce the effectiveness of

models when they are deployed in underrepresented populations, which will, in turn, cause diagnostic errors (Kabore & Guel, 2024; Mehrabi et al., 2021). The varying prevalence of species, including coinfection, is often not shown in training datasets, and this can lead to limited adaptability to clinical contexts.

To address these challenges in developing robust and generalizable models, we propose establishing a repository specific to malaria image datasets to host standardised datasets with accurate and complete metadata, similar to The Cancer Imaging Archive (TCIA). A standardized metadata schema can also be adopted to ensure that all imaging information is included in the datasets. Efforts can also be made to develop tools that can correct for the colour and contrast variations that can be caused using different stains and data heterogeneity.

7 Conclusion

In conclusion, this study provided an overview of the conventional microscopy method for malaria diagnosis, focusing on the curation of blood film image datasets, as data plays a vital role in developing automated diagnostic systems (Akinrinmade et al., 2023). The study revealed several limitations in the curated datasets. These include inconsistencies in image dataset formats, variations in magnification and annotation types, a notable absence of essential metadata (such as imaging equipment and staining methods) in many datasets, and a significant geographic bias. Furthermore, data accessibility remains a challenge, as many datasets are either unavailable or accessible only upon request, hindering reproducibility and the benchmarking of robust models. Larger, well-annotated image datasets are required to improve the classification accuracy of machine learning models for automated malaria diagnosis. To address these critical limitations and support the advancement of automated malaria diagnosis, this work introduces an open-source, interactive website (mind map). This centralized resource is designed to enhance data accessibility and discoverability by categorizing datasets by film type, Plasmodium species, staining method, imaging technology, geographical details, and availability, allowing researchers to easily explore and access source URLs or DOIs. It will act as a valuable resource for researchers and developers working on malaria detection and diagnosis. The interactive mind map is publicly accessible at https://itunuisewon.github.io/Malaria_Blood_Film_Images/, and the corresponding Zenodo repository containing the curated datasets is available at https://zenodo.org/records/16958076.

Acknowledgements

We acknowledge the contributors of publicly available malaria image datasets used in this review and those who maintain open-access repositories.

Supplementary data

Supplementary data are available at *Bioinformatics Advances* online.

Conflict of interest

None declared.

Funding

This work was supported by funding from the World Bank, granted to Covenant Applied Informatics, Communication Africa Centre of Excellence (CApIC-ACE) through the ACE Impact Project (2019–2024), and Covenant University Centre for Research, Innovation, and Discovery (CUCRID).

Data availability

The data underlying this article are available on GitHub at https://github.com/itunuisewon/Malaria_Blood_Film_Images.

References

Abdul-Nasir, A. S., Mashor, M. Y., Halim, N. H. A., & Mohamed, Z. (2015). The cascaded moving k -means and fuzzy c -means clustering algorithms for unsupervised segmentation of malaria images. *AIP Conference Proceedings*, *1660*. https://doi.org/10.1063/1.4915828

Acherar, A., Tantaoui, I., Thellier, M., Lampros, A., Piarroux, R., & Tannier, X. (2022). 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*, *35*(November), 101132. https://doi.org/10.1016/j.imu.2022.101132

Adeleke, O. T., Aworinde, H. O., Oboh, M., Oladosu, O., Ayenigba, A. B., Atobatele, B., Adeleke, O. V., Oladipo, T. S., & Adebayo, S. (2024). Dataset for a novel AI-powered diagnostic tool for Plasmodium parasite detection authors. *Data in Brief*, *57*, 110950. https://doi.org/10.1016/j.dib.2024.110950

Adi, K., Pujiyanto, S., Gernowo, R., Pamungkas, A., & Putranto, A. B. (2016). Identifying the developmental phase of plasmodium falciparum in malaria-infected red blood cells using adaptive color segmentation and back propagation neural network. *International Journal of Applied Engineering Research*, *11*(15), 8754–8759.

Akinrinmade, A. A., Adetiba, E., Badejo, J. A., & Oshin, O. (2023). NaijaFaceVoice: A Large-Scale Deep Learning Model and Database of Nigerian Faces and Voices. *IEEE Access*, *11*(June), 58228–58243. https://doi.org/10.1109/ACCESS.2023.3282618

Aladago, M. M., Torresani, L., & Rosca, E. V. (2019). Semantic Segmentation of the Growth Stages of Plasmodium Parasites using Convolutional Neural Networks. *IEEE AFRICON Conference*, *2019-Septe*. https://doi.org/10.1109/AFRICON46755.2019.9133937

Alnussairi, M. H. D., & İbrahim, A. A. (2022). Malaria parasite detection using deep learning algorithms based on (CNNs) technique. *Computers and Electrical Engineering*, *103*(September), 1–13. https://doi.org/10.1016/j.compeleceng.2022.108316

Aqeel, S., Haider, Z., & Wajihullah, K. (2022). Towards digital diagnosis of malaria: How far have we reached? *Journal of Microbiological Methods*, 37. https://doi.org/https://doi.org/10.1016/j.mimet.2022.106630

Aris, T. A., Nasir, A. S. A., Jaafar, H., Chin, L. C., & Mohamed, Z. (2021). Color constancy analysis approach for color standardization on malaria thick and thin blood smear images. *Lecture Notes in Electrical Engineering*, *666*(January), 785–804. https://doi.org/10.1007/978-981-15-5281-6_57

Arshad, Q. A., Ali, M., Hassan, S. ul, Chen, C., Imran, A., Rasul, G., & Sultani, W. (2022). A dataset and benchmark for malaria life-cycle classification in thin blood smear images. *Neural Computing and Applications*, *34*(6), 4473–4485. https://doi.org/10.1007/s00521-021-06602-6

Benjamin, G. Y., Mallam, F., Alhamdu, J. T., Yakubu, S. T., & Ladan, I. A. (2023). A review of the impact of COVID-19 pandemic on malaria control in Africa. *Microbes and Infectious Diseases*, *4*(4), 1081–1087. https://doi.org/10.21608/MID.2023.231844.1600

Chamidah, N., Lestari, B., Saifudin, T., Rulaningtyas, R., Wardhani, P., & Budiantar, I. N. (2024). ESTIMATING THE NUMBER OF MALARIA PARASITES ON BLOOD SMEARS MICROSCOPIC IMAGES USING PENALIZED SPLINE NONPARAMETRIC POISSON REGRESSION Malaria is one of the most common mosquito-borne diseases throughout the tropical and subtropical regions of the world wi. *Commun. Math. Biol. Neurosci.*, *2024*(Article ID 60), 1–19. https://doi.org/https://doi.org/10.28919/cmbn/8578

Cruz, C. J. G., Labreque, M., Cynthia Goh, M., & Rivera, P. T. (2016). APDS (automated parasite detection system) for field malaria diagnosis in the Philippines. *Southeast Asian Journal of Tropical Medicine and Public Health*, *47*(4), 600–606.

da Silva Ramos, J., Vieira, I. H. P., Rocha, W. S., Esquerdo, R. P., Watanabe, C. Y. V., & Zanchi, F. B. (2024). A transfer learning approach to identify Plasmodium in microscopic images. *PLoS Computational Biology*, *20*(8 August), 1–11. https://doi.org/10.1371/journal.pcbi.1012327

Dantas Oliveira, A., Carvalho, B. M., Prats, C., Espasa, M., Gomez i Prat, J., Lopez Codina, D., & Albuquerque, J. (2018). An automatic system for computing malaria parasite density in thin blood films. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *10657 LNCS*, 186–193. https://doi.org/10.1007/978-3-319-75193-1_23

Das, D. K., Maiti, A. K., & Chakraborty, C. (2015). Automated system for

characterization and classification of malaria-infected stages using light microscopic images of thin blood smears. *Journal of Microscopy*, *257*(3), 238–252. https://doi.org/10.1111/jmi.12206

Dave, I. R. (2018). Image analysis for malaria parasite detection from microscopic images of thick blood smear. *Proceedings of the 2017 International Conference on Wireless Communications, Signal Processing and Networking, WiSPNET 2017*, *2018-Janua*, 1303–1307. https://doi.org/10.1109/WiSPNET.2017.8299974

Davidson, M. S., Andradi-Brown, C., Yahiya, S., Chmielewski, J., O'Donnell, A. J., Gurung, P., Jeninga, M. D., Prommana, P., Andrew, D. W., Petter, M., Uthaipibull, C., Boyle, M. J., Ashdown, G. W., Dvorin, J. D., Reece, S. E., Wilson, D. W., Cunningham, K. A., Ando, D. M., Dimon, M., & Baum, J. (2021). Automated detection and staging of malaria parasites from cytological smears using convolutional neural networks. *Biological Imaging*, *1*, 1–13. https://doi.org/10.1017/s2633903x21000015

de Souza Oliveira, A., Guimarães Fernandes Costa, M., das Graças Vale Barbosa, M., & Ferreira Fernandes Costa Filho, C. (2022). A new approach for malaria diagnosis in thick blood smear images. *Biomedical Signal Processing and Control*, *78*(January). https://doi.org/10.1016/j.bspc.2022.103931

Delahunt, C. B., Gachuhi, N., & Horning, M. P. (2024). Metrics to guide development of machine learning algorithms for malaria diagnosis. *Frontiers in Malaria*, *2:1250220*(April), 1–14. https://doi.org/10.3389/fmala.2024.1250220

Delahunt, C. B., Jaiswal, M. S., Horning, M. P., Janko, S., Thompson, C. M., Kulhare, S., Hu, L., Ostbye, T., Yun, G., Gebrehiwot, R., Wilson, B. K., Long, E., Proux, S., Gamboa, Di., Chiodini, P., Carter, J., Dhorda, M., Isaboke, D., Ogutu, B., … Mehanian, C. (2019). Fully-automated patient-level malaria assessment on field-prepared thin blood film microscopy images. *2019 IEEE Global Humanitarian Technology Conference, GHTC 2019*, *i*. https://doi.org/10.1109/GHTC46095.2019.9033083

Delahunt, C. B., Jaiswal, M. S., Horning, M. P., Janko, S., Thompson, C. M., Kulhare, S., Hu, L., Ostbye, T., Yun, G., Gebrehiwot, R., Wilson, B. K., Long, E., Proux, S., Gamboa, Di., Chiodini, P., Carter, J., Dhorda, M., Isaboke, D., Ogutu, B., … Mehanian, C. (2022). Fully-automated patient-level malaria assessment on field-prepared thin blood film microscopy images. *2022 IEEE Global Humanitarian Technology Conference, GHTC 2022*, *iii*, 1–8. https://doi.org/10.1109/GHTC46095.2019.9033083

Delas Peñas, K., Rivera, P. T., & Naval, P. C. (2018). Analysis of Convolutional Neural Networks and Shape Features for Detection and Identification of Malaria Parasites on Thin Blood Smears. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *10752 LNAI*(March), 472–481. https://doi.org/10.1007/978-3-319-75420-8_45

Delgado-Ortet, M., Molina, A., Alférez, S., Rodellar, J., & Merino, A. (2020). A deep learning approach for segmentation of red blood cell images and malaria detection. *Entropy*, *22*(6), 1–16. https://doi.org/10.3390/e22060657

Devi, S. S., Roy, A., Singha, J., Sheikh, S. A., & Laskar, R. H. (2018). Malaria infected erythrocyte classification based on a hybrid classifier using microscopic images of thin blood smear. *Multimedia Tools and Applications*, *77*(1), 631–660. https://doi.org/10.1007/s11042-016-4264-7

Fong Amaris, W. M., Martinez, C., Cortés-Cortés, L. J., & Suárez, D. R. (2022). Image features for quality analysis of thick blood smears employed in malaria diagnosis. *Malaria Journal*, *21*(1), 1–12. https://doi.org/10.1186/s12936-022-04064-2

Fu, M., Wu, K., Li, Y., Luo, L., Huang, W., & Zhang, Q. (2023). An intelligent detection method for plasmodium based on self-supervised learning and attention mechanism. *Frontiers in Medicine*, *10*(June), 1–8. https://doi.org/10.3389/fmed.2023.1117192

G, A., S, E., I, I., J, E., T, D., S, R., J, O., & E., A. (2023). Mathematical modeling of malaria transmission dynamics in humans with mobility and control states. *Infectious Disease Modelling*, *8*(4), 1015–1031. https://doi.org/10.1016/j.idm.2023.08.005

Haralampieva, V., Rueckert, D., & Passerat-Palmbach, J. (2020). A Systematic Comparison of Encrypted Machine Learning Solutions for Image Classification. *PPMLP 2020 - Proceedings of the 2020 Workshop on Privacy-Preserving Machine Learning in Practice*, 55–59. https://doi.org/10.1145/3411501.3419432

Hartati, S., Harjoko, A., Rosnelly, R., Chandradewi, I., & Faizah. (2019). Performance of SVM and ANFIS for classification of malaria parasite and its life-cycle-stages in blood smear. In *Communications in Computer and Information Science* (Vol. 937). Springer Singapore. https://doi.org/10.1007/978-981-13-3441-2_9

Hcini, G., Jdey, I., & Ltifi, H. (2022). Improving Malaria Detection Using L1 Regularization Neural Network. *JUCS - Journal of Universal Computer Science*, *28*(10), 1087–1107. https://doi.org/10.3897/jucs.81681

He, K., Zhang, X., Ren, S., & Sun, J. (2015). *Deep Residual Learning for Image Recognition*.

Hendrawan, Y. F., Angkoso, C. V., & Wahyuningrum, R. T. (2017). Colour image segmentation for malaria parasites detection using cascading method. *2017 International Conference on Sustainable Information Engineering and Technology (SIET)*, 83–87. https://doi.org/10.1109/SIET.2017.8304114

Horning, M. P., Delahunt, C. B., Bachman, C. M., Luchavez, J., Luna, C., Hu, L., Jaiswal, M. S., Thompson, C. M., Kulhare, S., Janko, S., Wilson, B. K., Ostbye, T., Mehanian, M., Gebrehiwot, R., Yun, G., Bell, D., & Proux, S. (2021). Performance of a fully - automated system on a WHO malaria microscopy evaluation slide set. *Malaria Journal*, 1–11. https://doi.org/10.1186/s12936-021-03631-3

Horobin, R. W. (2011). How Romanowsky stains work and why they remain valuable - Including a proposed universal Romanowsky staining mechanism and a rational troubleshooting scheme. *Biotechnic and Histochemistry*, *86*(1), 36–51. https://doi.org/10.3109/10520295.2010.515491

Hung, Y. W., Wang, C. L., Wang, C. M., Chan, Y. K., Tseng, L. Y., Lee, C. W., & Tung, K. C. (2015). Parasite and infected-erythrocyte image segmentation in stained blood smears. *Journal of Medical and Biological Engineering*, *35*(6), 803–815. https://doi.org/10.1007/s40846-015-0101-0

Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine Learning and Deep Learning. *Electronic Markets*, *31*, 685–695. https://doi.org/10.4067/S0718-33052021000200180

Kabore, K. K., & Guel, D. (2024). *Addressing Challenges in Data Quality and Model Generalization for Malaria Detection*. https://doi.org/10.33140/JSNDC.04.03.09

Kalinin, V., Padnya, P., & Stoikov, I. (2024). Romanowsky staining: history, recent advances and future prospects from a chemistry perspective. *Biotechnic and Histochemistry*, *99*(1), 1–20. https://doi.org/10.1080/10520295.2023.2273860

Kanafiah, S. N. A. M., Mashor, M. Y., Mohamed, Z., Way, Y. C., Shukor, S. A. A., & Jusman, Y. (2022). An Intelligent Classification System for Trophozoite Stages in Malaria Species. *Intelligent Automation and Soft Computing*, *34*(1), 687–697. https://doi.org/10.32604/iasc.2022.024361

Kanakasabapathy, M. K., Thirumalaraju, P., Kandula, H., Doshi, F., Sivakumar, A. D., Kartik, D., Gupta, R., Pooniwala, R., Branda, J. A., Tsibris, A. M., Kuritzkes, D. R., Petrozza, J. C., Bormann, C. L., & Shafiee, H. (2021). Adaptive adversarial neural networks for the analysis of lossy and domain-shifted datasets of medical images. *Nature Biomedical Engineering*, *5*(6), 571–585. https://doi.org/10.1038/s41551-021-00733-w

Karasu Benyes, Y., Welch, E. C., Singhal, A., Ou, J., & Tripathi, A. (2022). A Comparative Analysis of Deep Learning Models for Automated Cross-Preparation Diagnosis of Multi-Cell Liquid Pap Smear Images. *Diagnostics*, *12*(8). https://doi.org/10.3390/diagnostics12081838

Karimi, A., Navidbakhsh, M., Haghi, A. M., & Faghihi, S. (2014). A morphology-based method for the diagnosis of red blood cells parasitized by Plasmodium malariae and Plasmodium ovale. *Scandinavian Journal of Infectious Diseases*, *46*(October 2013), 368–375. https://doi.org/10.3109/00365548.2014.880186

Kassim, Y. M., Yang, F., Yu, H., Maude, R. J., & Jaeger, S. (2021). Diagnosing malaria patients with plasmodium falciparum and vivax using deep learning for thick smear images. *Diagnostics*, *11*(11), 1–19. https://doi.org/10.3390/diagnostics11111994

Khan, N. A., Pervaz, H., Latif, A., & Musharaff, A. (2017). Unsupervised identification of malaria parasites using computer vision. *Pakistan Journal of Pharmaceutical Sciences*, *30*(1), 223–228.

Krappe, S., Benz, M., Gryanik, A., Tannich, E., Wegner, C., Stamminger, M., Wittenberg, T., & Münzenmayer, C. (2017). Automated plasmodia recognition in microscopic images for diagnosis of malaria using convolutional neural networks. *Medical Imaging 2017: Digital Pathology*, *10140*, 101400B. https://doi.org/10.1117/12.2249845

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). *ImageNet Classification with Deep Convolutional Neural Networks*. 1–9.

Kudisthalert, W., Pasupa, K., & Tongsima, S. (2020). Counting and classification

of malarial parasite from giemsa-stained thin film images. *IEEE Access*, *8*, 78663–78682. https://doi.org/10.1109/ACCESS.2020.2990497

Kuo, P. C., Cheng, H. Y., Chen, P. F., Liu, Y. L., Kang, M., Kuo, M. C., Hsu, S. F., Lu, H. J., Hong, S., Su, C. H., Liu, D. P., Tu, Y. C., & Chuang, J. H. (2020). Assessment of Expert-Level Automated Detection of Plasmodium falciparum in Digitized Thin Blood Smear Images. *JAMA Network Open*, *3*(2), e200206. https://doi.org/10.1001/jamanetworkopen.2020.0206

Li, S., & Yang, Z. (2020). *Microscopic Images of Parasites Species*. Mendeley Data. https://doi.org/10.17632/38jtn4nzs6.2

Linder, N., Turkki, R., Walliander, M., Ma¨rtensson, A., Diwan, V., Rahtu, E., Pietika¨inen, M., Lundin, M., & Lundin, J. (2014). A Malaria Diagnostic Tool Based on Computer Vision Screening and Visualization of Plasmodium falciparum Candidate Areas in Digitized Blood Smears. *PLoS ONE*, *9*(8). https://doi.org/10.1371/journal.pone.0104855

Liu, R., Liu, T., Dan, T., Yang, S., Li, Y., Luo, B., Zhuang, Y., Fan, X., Zhang, X., Cai, H., & Teng, Y. (2023). AIDMAN: An AI-based object detection system for malaria diagnosis from smartphone thin-blood-smear images. *Patterns*, *4*(9), 100806. https://doi.org/10.1016/j.patter.2023.100806

Loddo, A., Di Ruberto, C., Kocher, M., & Prod'hom, G. (2019). Mp-idb: The malaria parasite image database for image processing and analysis. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 11379). Springer International Publishing. https://doi.org/10.1007/978-3-030-13835-6_7

Loh, D. R., Yong, W. X., Yapeter, J., Subburaj, K., & Chandramohanadas, R. (2021). A deep learning approach to the screening of malaria infection: Automated and rapid cell counting, object detection and instance segmentation using Mask R-CNN. *Computerized Medical Imaging and Graphics*, *88*(July 2020), 101845. https://doi.org/10.1016/j.compmedimag.2020.101845

Maity, M., Jaiswal, A., Gantait, K., Chatterjee, J., & Mukherjee, A. (2020). Quantification of malaria parasitaemia using trainable semantic segmentation and capsnet. *Pattern Recognition Letters*, *138*, 88–94. https://doi.org/10.1016/j.patrec.2020.07.002

Maleki, F., Katie, O., Rajiv, G., Caroline, R., Spatz, A., & Reza, F. (2023). Generalizability of Machine Learning Models : Quantitative Evaluation of Three Methodological Pitfalls. *Radiology: Artificial Intelligence 2023;*, *5*(1). https://doi.org/https://doi.org/10.1148/ryai.220028

Malhotra, R., Joshi, D., & Shin, K. Y. (2017). *Approaching Bio Cellular Classification for Malaria Infected Cells Using Machine Learning and then Deep Learning to compare & analyze K-Nearest Neighbours and Deep CNNs*.

Manescu, P., Shaw, M. J., Elmi, M., Neary-Zajiczek, L., Claveau, R., Pawar, V., Kokkinos, I., Oyinloye, G., Bendkowski, C., Oladejo, O. A., Oladejo, B. F., Clark, T., Timm, D., Shawe-Taylor, J., Srinivasan, M. A., Lagunju, I., Sodeinde, O., Brown, B. J., & Fernandez-Reyes, D. (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

Maqsood, A., Farid, M., Khan, M., Sciences, M. G.-A., & 2021, undefined. (2021). Deep malaria parasite detection in thin blood smear microscopic images. *Mdpi.Com*. https://doi.org/10.3390/app11052284

Maturana, C. R., de Oliveira, A. D., Nadal, S., Bilalli, B., Serrat, F. Z., Soley, M. E., Igual, E. S., Bosch, M., Lluch, A. V., Abelló, A., López-Codina, D., Suñé, T. P., Clols, E. S., & Joseph-Munné, J. (2022). Advances and challenges in automated malaria diagnosis using digital microscopy imaging with artificial intelligence tools: A review. *Frontiers in Microbiology*, *13*(November), 1–17. https://doi.org/10.3389/fmicb.2022.1006659

Mehanian, C., Jaiswal, M., Delahunt, C., Thompson, C., Horning, M., Hu, L., Mcguire, S., Ostbye, T., Mehanian, M., Wilson, B., Champlin, C., Long, E., Proux, S., Gamboa, D., Chiodini, P., Carter, J., Dhorda, M., Isaboke, D., Ogutu, B., … Software, C. C. (2017). Computer-Automated Malaria Diagnosis and Quantitation Using Convolutional Neural Networks. *2017 IEEE International Conference on Computer Vision Workshops (ICCVW), Venice, Italy*, 116–125. https://doi.org/10.1109/ICCVW.2017.22.

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A Survey on Bias and Fairness in Machine Learning. *ACM Computing Surveys*, *54*(6). https://doi.org/10.1145/3457607

Molina, A., Alferez, S., Boldu, L., Acevedo, A., Rodellar, J., & Merino, A. (2020). Sequential classification system for recognition of malaria infection using peripheral blood cell images. *Journal of Clinical Pathology*, *73*(10), 665–670. https://doi.org/10.1136/jclinpath-2019-206419

Molina, A., Rodellar, J., Boldú, L., Acevedo, A., Alférez, S., & Merino, A. (2021). Automatic identification of malaria and other red blood cell inclusions using convolutional neural networks. *Computers in Biology and Medicine*, *136*(March). https://doi.org/10.1016/j.compbiomed.2021.104680

Molina, M. Á., Asencio-Cortés, G., Riquelme, J. C., & Martínez-Álvarez, F. (2021). A Preliminary Study on Deep Transfer Learning Applied to Image Classification for Small Datasets. *Advances in Intelligent Systems and Computing*, *1268 AISC*, 741–750. https://doi.org/10.1007/978-3-030-57802-2_71

Muhammad, F. A., Sudirman, R., Zakaria, N. A., & Noor Syakiylla Sayed Daud, S. (2025). Morphology classification of malaria infected red blood cells using deep learning techniques. *Biomedical Signal Processing and Control*, *99*(December 2023), 106869. https://doi.org/10.1016/j.bspc.2024.106869

Musumeci, G. (2014). Past, present and future: overview on histology and histopathology. *Journal of Histology and Histopathology*, *1*(1), 5. https://doi.org/10.7243/2055-091x-1-5

Nakasi, R., Mwebaze, E., & Zawedde, A. (2021). Mobile-aware deep learning algorithms for malaria parasites and white blood cells localization in thick blood smears. *Algorithms*, *14*(1), 1–16. https://doi.org/10.3390/a14010017

National Library of Medicine, National Institutes of Health, Bethesda, M. 20894. (2020). *NLM - Malaria Data*. https://lhncbc.nlm.nih.gov/LHC-research/LHC-projects/image-processing/malaria-datasheet.html

Nugroho, A. S., Gunawan, M., Pragesjvara, V., Riski, M. J., Desiani, Ashilah, I., Hariani, I., Rozi, I. E., Asih, P. B. S., Salamah, U., & Suryani, E. (2014). Two-stage feature extraction to identify Plasmodium ovale from thin blood smear microphotograph. *Proceedings of 2014 International Conference on Data and Software Engineering, ICODSE 2014*, 1–4. https://doi.org/10.1109/ICODSE.2014.7062676

Nugroho, H. A., Dendi Maysanjaya, I. M., Setiawan, N. A., Murhandarwati, E. E. H., & Oktoeberza, W. K. Z. (2019). Feature analysis for stage identification of Plasmodium vivax based on digital microscopic image. *Indonesian Journal of Electrical Engineering and Computer Science*, *13*(2), 721–728. https://doi.org/10.11591/ijeecs.v13.i2.pp721-728

Nugroho, H. A., Goratama, R. D., & Frannita, E. L. (2021). Saturation channel extraction of HSV color space for segmenting Plasmodium parasite. *IOP Conference Series: Materials Science and Engineering*, *1088*(1), 012073. https://doi.org/10.1088/1757-899x/1088/1/012073

Nugroho, H. A., Nurfauzi, R., Murhandarwati, E. E. H., & Purwono, P. (2022). PlasmoID : A dataset for Indonesian malaria parasite detection and segmentation in thin blood smear. *Arxiv.Org*, 1–16.

Ojurongbe, O., Adegbosin, O. O., Taiwo, S. S., Armstrong, O., Alli, T., Olowe, O. A., Ojurongbe, T. A., Bolaji, O. S., & Adeyeba, O. A. (2013). *Assessment of Clinical Diagnosis , Microscopy , Rapid Diagnostic Tests , and Polymerase Chain Reaction in the Diagnosis of Plasmodium falciparum in Nigeria*. *2013*, 5–10.

Olasehinde, G., Diji-Geske, R., Fadina, I., Arogundade, D., Darby, P., Adeleke, A., Dokunmu, T., Adebayo, A., Oyelade, & J. (2019). *Epidemiology of Plasmodium falciparum infection and drug resistance markers in Ota Area , Southwestern Nigeria*. 1941–1949.

Oyewola, D. O., Dada, E. G., Misra, S., & Damaševičius, R. (2022). A Novel Data Augmentation Convolutional Neural Network for Detecting Malaria Parasite in Blood Smear Images. *Applied Artificial Intelligence*, *36*(1). https://doi.org/10.1080/08839514.2022.2033473

Ozbilge, E., Guler, E., & Ozbilge, E. (2024). Ensembling Object Detection Models for Robust and Reliable Malaria Parasite Detection in Thin Blood Smear Microscopic Images. *IEEE Access*, *12*, 60747–60764. https://doi.org/10.1109/ACCESS.2024.3393410

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *The BMJ*, *372*. https://doi.org/10.1136/bmj.n71

Pan, W., Dong, Y., Advanced, D. W.-M. learning:, & 2018, U. (2018). Classification of malaria-infected cells using deep convolutional neural networks. In *books.google.com*. https://books.google.com/books?hl=en&lr=&id=EG-QDwAAQBAJ&oi=fnd&pg=PA159&dq=%22image+classification%22++in+malaria&ots=UXE59MWcWw&sig=-

tYKk8D__AKbbYXYUxduSq0gpXY

Park, H. S., Rinehart, M. T., Walzer, K. A., Chi, J. A., & Wax, A. (2016). Automated Detection of P . falciparum Using Machine Learning Algorithms with Quantitative Phase Images of Unstained Cells. *PLoS ONE*, *11*(9). https://doi.org/10.1371/journal.pone.0163045

Poostchi, M., Silamut, K., Maude, R. J., Jaeger, S., & Thoma, G. (2018). Image analysis and machine learning for detecting malaria. *Translational Research*, *194*, 36–55. https://doi.org/10.1016/j.trsl.2017.12.004

Preedanan, W., Phothisonothai, M., Senavongse, W., & Tantisatirapong, S. (2016). Automated detection of plasmodium falciparum from Giemsa-stained thin blood films. *2016 8th International Conference on Knowledge and Smart Technology, KST 2016*, 215–218. https://doi.org/10.1109/KST.2016.7440501

Prosser, C., Gresty, K., Ellis, J., Meyer, W., Anderson, K., Lee, R., & Cheng, Q. (2021). Plasmodium falciparum histidine-rich protein 2 and 3 gene deletions in strains from Nigeria, Sudan, and South Sudan. *Emerging Infectious Diseases*, *27*(2), 471–479. https://doi.org/10.3201/EID2702.191410

Quinn, J. A., Andama, A., Munabi, I., & Kiwanuka, F. N. (2014). Automated blood smear analysis for mobile malaria diagnosis. *Mobile Point-of-Care Monitors and Diagnostic Device Design*, *October*, 115–132. https://doi.org/10.1201/9781315215723-10

Rabbani, H., Zekri, M., & Talebi, A. (2013). Automatic diagnosis of malaria based on complete circle-ellipse fitting search algorithm. *Journal of Microscopy*, *September*. https://doi.org/10.1111/jmi.12081

Rahmanti, F. Z., Sutojo, Ningrum, N. K., Imania, N. K., & Purnomo, M. H. (2016). Plasmodium vivax classification from digitalization microscopic thick blood film using combination of second order statistical feature extraction and K-Nearest Neighbor (K-NN) classifier method. *Proceedings - 2015 4th International Conference on Instrumentation, Communications, Information Technology and Biomedical Engineering, ICICI-BME 2015*, 79–83. https://doi.org/10.1109/ICICI-BME.2015.7401339

Rajaraman, S., Antani, S. K., Poostchi, M., Silamut, K., Hossain, M. A., Maude, R. J., Jaeger, S., & Thoma, G. R. (2018). Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images. *PeerJ*, *2018*(4), 1–17. https://doi.org/10.7717/peerj.4568

Rees-Channer, R. R., Bachman, C. M., Grignard, L., Gatton, M. L., Burkot, S., Horning, M. P., Delahunt, C. B., Hu, L., Mehanian, C., Thompson, C. M., Woods, K., Lansdell, P., Shah, S., & Chiodini, P. L. (2023). Evaluation of an automated microscope using machine learning for the detection of malaria in travelers returned to the UK. *Frontiers in Malaria*, *1*(August), 1–10. https://doi.org/10.3389/fmala.2023.1148115

Rehman, A., Abbas, N., Saba, T., Mehmood, Z., Mahmood, T., & Ahmed, K. T. (2018). Microscopic malaria parasitemia diagnosis and grading on benchmark datasets. *Microscopy Research and Technique*, *81*(9), 1042–1058. https://doi.org/10.1002/jemt.23071

Rehman, M. Z. ur, Ahmad, A., Chaudhary, F., & Ashfaque, S. M. (2020). Malaria Cell Image Classification using Deep Learning. *International Journal of Recent Technology and Engineering (IJRTE)*, *7*(4), 6259–6263. https://doi.org/10.35940/ijrte.f9540.038620

Resource, P. E. I. (2014). *IPLab:Lab 11:Malaria --- Pathology Education Instructional Resource{,}*. https://peir.path.uab.edu/index.php?title=IPLab:Lab_11:Malaria&oldid=2227

Rosado, L., da Costa, J. M. C., Elias, D., & Cardoso, J. S. (2017). Mobile-based analysis of malaria-infected thin blood smears: Automated species and life cycle stage determination. *Sensors (Switzerland)*, *17*(10), 1–22. https://doi.org/10.3390/s17102167

Rosnelly, R., Hartati, S., Winarko, E., & Mulatsih, S. (2017). Identification of malaria disease and its stadium based on digital image processing. *Journal of Theoretical and Applied Information Technology*, *95*(3), 700–710.

Russell, P. H., Johnson, R. L., Ananthan, S., Harnke, B., & Carlson, N. E. (2018). A large-scale analysis of bioinformatics code on GitHub. *PLoS ONE*, *13*(10), 1–19. https://doi.org/10.1371/journal.pone.0205898

Salamah, U., Sarno, R., Arifin, A. Z., Nugroho, A. S., Rozi, I. E., & Asih, P. B. S. (2019). A robust segmentation for malaria parasite detection of thick blood smear microscopic images. *International Journal on Advanced Science, Engineering and Information Technology*, *9*(4), 1450–1459. https://doi.org/10.18517/ijaseit.9.4.4843

Sampathila, N., Shet, N., & Basu, A. (2018). Computational approach for diagnosis of malaria through classification of malaria parasite from microscopic image of blood smear. *Biomedical Research (India)*, *29*(18), 3464–3468. https://doi.org/10.4066/biomedicalresearch.29-18-970

Sengar, N., Burget, R., & Dutta, M. K. (2022). A vision transformer based approach for analysis of plasmodium vivax life cycle for malaria prediction using thin blood smear microscopic images. *Computer Methods and Programs in Biomedicine*, *224*, 106996. https://doi.org/10.1016/j.cmpb.2022.106996

Shaka, M. (2021). *EFFECTIVE CREATION OF GROUND TRUTH DATA- SET FOR MALARIA DIAGNOSIS - 1ST AFRICANLP WORKSHOP , INTERNATIONAL CONFERENCE ON LEARNING REPRESENTATIONS (ICLR 2020). Iclr 2020*.

Sharma, M., Devi, S. S., & Laskar, R. H. (2020). Automatic Detection of Malaria Infected Erythrocytes Based on the Concavity Point Identification and Pseudo-Valley Based Thresholding. *IETE Journal of Research*, *0*(0), 1–18. https://doi.org/10.1080/03772063.2020.1787238

Shi, L., Guan, Z., Liang, C., & You, H. (2020). Automatic Classification of Plasmodium for Malaria Diagnosis based on Ensemble Neural Network. *ACM International Conference Proceeding Series*, 80–85. https://doi.org/10.1145/3399637.3399641

Simonyan, K., & Zisserman, A. (2015). *VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION*. 1–14.

Straiton, J. (2024). Artificial intelligence: help or hindrance in solving the reproducibility crisis? *BioTechniques*, *76*(7), 291–294. https://doi.org/10.1080/07366205.2024.2355776

Sukumarran, D., Hasikin, K., Khairuddin, A. S. M., Ngui, R., Sulaiman, W. Y. W., Vythilingam, I., & Divis, P. C. S. (2024). An optimised YOLOv4 deep learning model for efficient malarial cell detection in thin blood smear images. *Parasites and Vectors*, *17*(1), 1–26. https://doi.org/10.1186/s13071-024-06215-7

Sumi, A. S., Nugroho, H. A., & Hartanto, R. (2021). Global Threshold Using Otsu and Active Contour for Detection of Malaria Parasites in Thick Blood Smear. *BIO Web of Conferences*, *41*. https://doi.org/10.1051/bioconf/20214104002

Sunarko, B., Williams, S., Prescott, W. R., Byker, S. M., & Bottema, M. J. (2017). Correlation between automatic detection of malaria on thin film and experts' parasitaemia scores. *AIP Conference Proceedings*, *1818*(March). https://doi.org/10.1063/1.4976918

Szegedy, C., Reed, S., Sermanet, P., Vanhoucke, V., & Rabinovich, A. (2014). *Going deeper with convolutions*. 1–12.

Vijayalakshmi A, & Rajesh Kanna B. (2020). Deep learning approach to detect malaria from microscopic images. *Multimedia Tools and Applications*, *79*(21–22), 15297–15317. https://doi.org/10.1007/s11042-019-7162-y

Wang, G., Luo, G., Lian, H., Chen, L., Wu, W., & Liu, H. (2023). Application of Deep Learning in Clinical Settings for Detecting and Classifying Malaria Parasites in Thin Blood Smears. *Open Forum Infectious Diseases*, *10*(11), 1–8. https://doi.org/10.1093/ofid/ofad469

WHO. (2016a). Collection of blood by venipuncture and preparation of blood films from venous blood collected in tubes containing anticoagulant. *Malaria Microscopy Standard Operating Procedure*, *Version 1*(MM-SOP-05B), 1–4.

WHO. (2016b). Malaria Parasite Counting. *Malaria Microscopy Standard Operating Procedure*, 1–5. https://apps.who.int/iris/handle/10665/274382

WHO. (2016c). Microscopy Examination of Thick and Thin Blood Films for Identification of Malaria Parasites: Malaria Microscopy Standard Operating Procedure-Mm-Sop-08. *World Health Organization*, 1–6.

WHO. (2024). World Malaria Report 2024. In *Word Malaria report Geneva: World Health Organization. (2024). Licence: CC*.

Wibisono, Y., Nugroho, A. S., & Galinium, M. (2020). Optimization on Malaria Computer Aided Diagnostic System. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/3429789.3429825

Wilkinson, M. D. (2016). FAIR Guiding Principles for Reproducibility. *ScientificData*, 1–9.

Xiong, W., Ong, S.-H., Lim, J.-H., Cheng, J., & Ying, G. (2015). *Blood Smear Analysis ,*.

Yang, F., Poostchi, M., Yu, H., Zhou, Z., Silamut, K., Yu, J., Maude, R. J., Jaeger, S., & Antani, S. (2020). Deep Learning for Smartphone-Based Malaria Parasite Detection in Thick Blood Smears. *IEEE Journal of Biomedical and Health Informatics*, *24*(5), 1427–1438. https://doi.org/10.1109/JBHI.2019.2939121

Yoon, J., Jang, W. S., Nam, J., Mihn, D.-C., & Lim, C. S. (2021). An Automated

Microscopic Malaria Parasite Detection System Using Digital Image Analysis. *Diagnostics*, *11*(527), 1–13.

Yu, H., Mohammed, F. O., Abdel Hamid, M., Yang, F., Kassim, Y. M., Mohamed, A. O., Maude, R. J., Ding, X. C., Owusu, E. D. A., Yerlikaya, S., Dittrich, S., & Jaeger, S. (2023). Patient-level performance evaluation of a smartphone-based malaria diagnostic application. *Malaria Journal*, *22*(1), 1–10. https://doi.org/10.1186/s12936-023-04446-0

Zhong, Y., Dan, Y., Cai, Y., Lin, J., Huang, X., Mahmoud, O., Hald, E. S., Kumar, A., Fang, Q., & Mahmoud, S. S. (2023). Efficient Malaria Parasite Detection From Diverse Images of Thick Blood Smears for Cross-Regional Model Accuracy. *IEEE Open Journal of Engineering in Medicine and Biology*, *4*, 226–233. https://doi.org/10.1109/OJEMB.2023.3328435

Zindi. (2024). *Lacuna Malaria Detection Challenge*. https://zindi.africa/competitions/lacuna-malaria-detection-challenge

Table 1: The Eligibility Criteria for the Study

Inclusion Criteria	Criteria
IC1	The study was published between 2014 and 2024.
IC2	The study involves the automated diagnosis of Malaria targeting microscopy.
IC3	The study is published in a conference, journal, or book chapter.
IC4	The full text of the study is available.
IC5	The study is in the English Language.

Exclusion Criteria	Criteria
EC1	The study did not give sufficient information about the dataset.
EC2	The study is not focused on automated microscopy.

[3]**Table 2: Summary of available malaria image datasets**

Identifier	**Country**	***Plasmodium* species**	**Film type**	**Optical Train**	**Stain type**	**No of Patients**	**No of images**	**Type of Image**	**Publicly Available?**
D1 (Yu *et al.,* 2023)	Sudan	Pv, Pf	Thick & Thin	Microscope + smartphone	Giemsa	95	Thick – 2944 Thin - 875	FoVs	Yes
D2 (Quinn *et al.,* 2014)	Uganda	Pf	Thick	Microscope + Camera	Field	133	2703	FoVs	Yes
D3 (Fong Amaris *et al.,* 2022)	Columbia	Pv	Thick	Microscope	Modified Romanowsky	42	420	FoVs	Yes
D4 (de Souza Oliveira *et al.,* 2022)	Brazil	Pv	Thick	Microscope	Giemsa	NR	676	FoVs	AOR
D5 (Arshad *et al.,* 2022)	Pakistan	Pv	Thin	Microscope	Giemsa	NR	345	FoVs	AOR
D6 (Nakasi *et al.,* 2021)	Uganda	Pf	Thick	Microscope + smartphone	Field	NR	903	FoVs	Yes
D7 (Manescu *et al.,* 2020)	Nigeria	Pf	Thick	Microscope	Giemsa	299	NR	FoVs	Yes
D8 (Delgado-Ortet *et al.,* 2020)	Spain	Pf	Thin	Microscope + Camera	May-Grunwald-Giemsa	NR	517	FoVs	Yes

[3] Pf – *Plasmodium falciparum*, Pv- *Plasmodium vivax*, Po – *Plasmodium ovale,* Pm – *Plasmodium malariae,* Pk – *Plasmodium knowlesi*, Pc – *Plasmodium cynomolgi,* NR – Not reported, FoVs – Field of Views, AOR – Available on Request

D9 (Loddo *et al.,* 2019)	Switzerland	Pf, Pm, Po, Pv	Thin	Microscope	Giemsa	NR	Pf – 122 Po - 29 Pm - 37 Pv - 46	FoVs	Yes
D10 (Rajaraman *et al.,* 2018), (Yang *et al.,* 2020)	Bangladesh	Pf	Thick & Thin	Microscope + smartphone	Giemsa	200	Thick – 1883 Thin - 965	Cells images & FoVs	Yes
D11 (Kassim *et al.,* 2021)	Bangladesh	Pv	Thick	Microscope + smartphone	Giemsa	200	NR	FoVs	Yes
D12 (Kanafiah *et al.,* 2022)	Malaysia	Pk, Pf, Pv	Thin	Microscope + Camera	NR	NR	300	FoVs	AOR
D13 (Kuo *et al.,* 2020)	Taiwan	Pf	Thin	Microscope	Giemsa	36	8145	FoVs	Yes
D14 (Aris *et al.,* 2021)	Malaysia	Pf, Pv	Thick & Thin	Microscope	NR	NR	Thick - 200, Thin - 400	FoVs	No
D15 (Maity *et al.,* 2020)	India	Pf, Pv	Thin	Microscope	NR	NR	210	FoVs	Yes
D16 (Aladago *et al.,* 2019)	Ghana	Pf	Thin	Microscope	Giemsa	NR	403	FoVs	No
D17 (Dave, 2018)	India	Pv	Thick	Microscope + Camera	Giemsa	5	87	FoVs	No
D18 (Delas Peñas *et*	Philippines	Pv, Pf	Thin	Microscope +	NR	NR	Pv = 142	FoVs	No

al., 2018)				Camera			Pf = 221		
D19 (Dantas Oliveira *et al.,* 2018)	Spain	Pf	Thin	Microscope + Camera	Giemsa	NR	50	FoVs	No
D20 (Devi *et al.,* 2018)	India	Pv, Pf	Thin	Microscope	Leishman	400	1302	FoVs	No
D21 (Rosado *et al.,* 2017)	Portugal	Pf, Po, Pm	Thin	mSmartScope prototype + Smartphone	Giemsa	7	566	FoVs	No
D22 (Rosnelly *et al.,* 2017)	Indonesia	Pf, Pm, Pv	Thin	Microscope	Giemsa	NR	600	FoVs	No
D23 (Khan *et al.,* 2017)	NR	Pv	Thick & Thin	Microscope	Leishman	118	330	FoVs	No
D24 (Preedanan *et al.,* 2016)	Thailand	Pf	Thin	Microscope + Camera	Giemsa	NR	15	FoVs	No
D25 (Cruz *et al.,* 2016)	Philippines	Pf, Pv	Thin	Microscope + Camera	Giemsa	21	267	FoVs	No
D26 (Abdul-Nasir *et al.,* 2015)	Malaysia	Pv	Thin	Microscope + Camera	Giemsa	NR	125	FoVs	No
D27 (Das *et al.,* 2015)	India	Pf, Pv	Thin	Microscope	Leishman	NR	750	FoVs	No
D28 (A. Molina *et al.,* 2020)	NR	NR	Thin	Microscope + Camera	Giemsa	87	NR	FoVs	No

D29 (Fu *et al.,* 2023)	China	Pf, Pv, Pm, Po	Thin	NR	NR	100	NR	Blood slice images	No
D30 (Linder *et al.,* 2014)	Finland	Pf	Thin	Microscope + Camera	Giemsa	47	42,787	FoVs	Yes
D31 (A. S. Nugroho *et al.,* 2014)	Indonesia	Po	Thin	NR	Giemsa	NR	NR	FoVs	No
D32 (Liu *et al.,* 2023)	Sierra Leone	NR	Thin	Microscope + smartphone	Giemsa	NR	177	FoVs	Yes
D33 (Loh *et al.,* 2021)	NR	Pf 3D7 strain	Thin	Microscope + Camera	Giemsa	NR	297	FoVs	No
D34 (Acherar *et al.,* 2022)	France	Pf	Thin	Microscope + smartphone	Giemsa	202	1250	FoVs	No
D35 (Salamah *et al.,* 2019)	Indonesia	NR	Thick	Microscope	NR	NR	30	fov	No
D36 (Sengar *et al.,* 2022)	Pakistan	NR	Thin	Microscope	Giemsa	NR	345	Cell images	No
D37 (Yoon *et al.,* 2021)	Korea	Pf, Pv	Thick & Thin	Microscope	Giemsa	71	NR	FoVs	No
D38 (Davidson *et al.,* 2021)	NR	Pf 3D7, NF54, DD2, and D10	Thin	Different Equipment	Giemsa	NR	108	FoVs	No

D39 (Sunarko *et al.,* 2017)	NR	NR	Thin	Microscope	Giemsa	7	NR	FoVs	No
D40 (Zhong *et al.,* 2023)	Sudan	Pf, Pv	Thick	Microscope + smartphone	Giemsa	NR	111	FoVs	No
D41 (Wang *et al.,* 2023)	China	Pf, Pv, Pm, Po, Pk, Pc	Thin	Microscope + camera	Wright-Giemsa	380	Pf – 4405 Pv - 3030 Pm - 225 Po - 117 Pk - 4630 Pc - 139	FoVs	No
D42 (Hartati *et al.,* 2019)	Indonesia	Pm, Pv, Pf	Thin	NR	Giemsa	NR	600	FoVs	No
D43 (H. A. Nugroho *et al.,* 2021)	Thailand	Pf	Thin	Microscope + camera	Giemsa	NR	280	FoVs	No
D44 (Sampathila *et al.,* 2018)	India	Pv	Thin	Microscope + camera	Leishman	NR	143	FoVs	No
D45 (H. A. Nugroho *et al.,* 2019)	Indonesia	Pv	Thin	NR	NR	NR	73	FoVs	No
D46 (Kudisthalert *et al.,* 2020)	Thailand	Pf	Thin	Microscope + camera	Giemsa	NR	23,248	Cell images	No
D47 (Shi *et al.,* 2020)	China		NR	NR	NR	NR	217	Cell	No

								images	
D48 (Vijayalakshmi A & Rajesh Kanna B, 2020)	India	Pf	Thin	Microscope + camera	Giemsa	50	NR	FoVs	No
D49 (Resource, 2014)	NR	Pf	Thin	NR	NR	NR	NR	Whole slide image	Yes
D50 (Sumi *et al.,* 2021)	Indonesia	NR	Thick	Microscope + camera	Giemsa	NR	36	FoVs	No
D51 (Karimi *et al.,* 2014)	Iran	Po, Pm	Thin	Microscope	Giemsa	27	27	FoVs	No
D52 (Wibisono *et al.,* 2020)	Indonesia	NR	Thin	NR	NR	NR	100	FoVs	No
D53 (Hung *et al.,* 2015)	NR	Pf, Pm, Po, Pv	NR	NR	NR	NR	96	FoVs	No
D54 (Adi *et al.,* 2016)	NR	Pf	NR	Microscope + camera	NR	NR	152	FoVs	No
D55 (Rahmanti *et al.,* 2016)	Indonesia	Pv	Thick	Microscope + camera	NR	NR	NR	Cell images	No
D56 (Xiong *et al.,* 2015)	Singapore	NR	NR	Microscope + camera	Giemsa	4	>15000	FoVs	No

D57 (Hendrawan *et al.,* 2017)	Indonesia	Pf, Pv, Pm, Po	NR	NR	NR	NR	574	FoVs	No
D58 (Krappe *et al.,* 2017)	NR	Pf, Pv, Pm	Thick	Microscope	NR	44	NR	Cell images	No
D59 (Sharma *et al.,* 2020)	India	Pf, Pv	Thin	NR	Leishman	200	2800	FoVs	No
D60 (H. A. Nugroho *et al.,* 2022)	Indonesia	Pf, Pv, Pm, Po	Thin	Microscope	NR	NR	559	FoVs	No
D61 (Delahunt *et al.,* 2019)	South America, Africa, Asia, and London	Pf, Pv, Pm, Po	Thin	Microscope	Giemsa	765	323000	FoVs	No
D62 (Kanakasabapathy *et al.,* 2021)		Pf	Thin	Microscope	Giemsa	8	12635, 11397, 14469	Cell images	Yes
D63 (Shaka, 2021)	Tanzania	NR	Thick	Microscope + smartphone	Giemsa	100	10000	FoVs	Yes
D64 (Zindi, 2024)	Uganda	NR	NR	Microscope + smartphone	NR	NR	3925	FoVs	Yes
D65 (Muhammad *et*	Nigeria	NR	Thin	Microscope +	Giemsa, Field	100	24,712	FoVs	Yes

al., 2025)				smartphone					
D66 (Adeleke *et al.,* 2024)	Nigeria	Pm, Po, Pf	Thick & Thin	Microscope	Giemsa	NR	881	FoVs	Yes
D67 (Sukumarran *et al.,* 2024)	Malaysia	Pf,Pv,Po,Pm,Pk	Thin	Microscope + camera	Giemsa	NR	472	FoVs	No
D68 (Li & Yang, 2020)	NR	NR	NR	NR	NR	NR	843	Cell images	Yes
D69 (da Silva Ramos *et al.,* 2024)	Brazil	Pv	Thin	Microscope	Giemsa	NR	27	FoVs	Yes
D70 (Chamidah *et al.,* 2024)	Indonesia	NR	NR	NR	NR	NR	100	FoVs	No
D71 (Ozbilge *et al.,* 2024)	Cyprus	Pf,Pv,Po	Thin	Microscope + smartphone	Giemsa	25	1081	FoVs	No

[4]Table 3: Imaging Techniques used in the selected studies

Identifier	Optical Train	Microscope type	Magnification	Image Size
D1	Microscope + smartphone	Olympus C×23 microscope	NR	NR
D2	Microscope + Camera	Brunel SP150 microscope	1000×	1024 × 768
D3	Microscope	A×io Zeiss Scope A1 Optical microscope	NR	2056 × 2452
D4	Microscope	ZEISS Axio Imager M2 microscope	100×	1388 ×1040
D5	Microscope	×SZ-107 series microscope	100×	NR
D6	Microscope + smartphone	Olympus microscope	100×	3264 × 2448
D7	Microscope	Brightfield microscope (Olympus B×63)	100×	NR
D8	Microscope + Camera	Olympus B×43	100×	2400×1800
D9	Microscope	Leica DM2000 optical laboratory microscope	100×	2592 × 1944
D10	Microscope + smartphone	NR	NR	NR
D11	Microscope + smartphone	Light microscope	100x	3024×4032
D12	Microscope + Camera	Olympus B×41 microscope	100×	574 × 764
D13	Microscope	VS120, Olympus Corp	100×	2048× 2048
D14	Microscope	Leica DLMA microscope	100×	1200 × 900
D15	Microscope	NR	100×	2048 ×1536
D16	Microscope	Light microscope	100×	1500 × 2100 × 3
D17	Microscope + Camera	Olympus C×41 Microscope	100×	3136 × 2352

[4] NR – Not reported

D18	Microscope + Camera	NR	100×	2592 × 1944
D19	Microscope + Camera	Light microscope	100x	640 × 480
D20	Microscope	NR	100×	1500 × 1000
D21	mSmartScope prototype + Smartphone	NR	NR	1944 x 2592 - 1840 x 3264 pixels
D22	Microscope	NR	100x	NR
D23	Microscope	NR	100X	352 x 288
D24	Microscope + Camera	Olympus BX51 microscope	NR	1360 x 1024
D25	Microscope + Camera	OMAX MD827S30 Binocular microscope	100x	NR
D26	Microscope + Camera	Leica DLMA microscope.	100x	800 x 600
D27	Microscope	Leica (DM750) microscope	100x	NR
D28	Microscope + Camera	Olympus BX43 microscope	100x	2400×1800
D29				
D30	Microscope + Camera	Axio Imager Z2, Carl Zeiss Microscopy AG, Jena, Germany	100x	1280×1024
D31	NR	NR	NR	NR
D32	Microscope + smartphone	NR	NR	NR
D33	Microscope + Camera	Leica ICC50 W microscope	NR	NR
D34	Microscope + smartphone	Olympus BX51 optical microscope	500 ×	NR
D35	Microscope	NR	100×	1280 x 960
D36	Microscope	XSZ-107 series microscope	100x	NR

D37	Microscope	NR	100×	NR
D38	Microscope, Microscope + Camera	NR	NR	NR
D39	Microscope	NR	400x	540 x 960
D40	Microscope + smartphone	Mobile Microscope Model	100x	NR
D41	Microscope + camera	Olympus CX31 microscope	NR	1920 × 1200
D42	NR		NR	2560 x 1920
D43	Microscope + camera	Olympus BX51 microscope	NR	1360×1024
D44	Microscope + camera	Olympus (BX51) Microscope	NR	NR
D45	NR	Optilab camera	100x	1600x1200
D46	Microscope + camera	Olympus BX51 microscope	100x	4080 x 3072
D47	NR	NR	NR	4032 × 3024
D48	Microscope + camera	Olympus CX21i bright field microscopic stage	100x	NR
D49	NR	NR		NR
D50	Microscope + camera	NR	NR	1600x1200
D51	Microscope	Olympus BX41 Microscope	100x	
D52	NR	NR	100x	1280 x 960
D53	NR	NR	100x	
D54	Microscope + camera	NR	100x	400 x 320
D55	Microscope + camera	NR	NR	NR
D56	Microscope + camera	Olympus BX51 microscope	100x	NR
D57	NR	NR	NR	256 x 256

D58	Microscope	NR	NR	NR
D59	NR	NR	100x	1500×1000
D60	Microscope	Optilab digital microscope camera	100x	960x1280
D61	Microscope	Motic EasyScan-Go	NR	2048×1536
D62	Microscope	Benchtop microscope, a portable stand-alone 3D-printed microscope and a smartphone-based microscope	NR	NR
D63	Microscope + smartphone	Olympus CX 21 microscope	100x	NR
D64	Microscope + smartphone	NR	NR	NR
D65	Microscope + smartphone	NR	NR	750×750
D66	Microscope	NR	100x	NR
D67	Microscope + camera	Olympus BX53 microscope	100x	NR
D68	NR	NR	400x	NR
D69	Microscope	NR	100x	1240×1645
D70	Microscope	NR	NR	600×512
D71	Microscope + smartphone	Light microscope	100x	640×640

Table 4: Dataset Availability[5]

Identifier	Data Availability	Dataset Link	Doi/Title
D1	Publicly Available	https://data.lhncbc.nlm.nih.gov/public/Malaria/MalariaScreener/index.html	https://doi.org/10.1186/s12936-023-04446-0
D2	Publicly Available	https://air.ug/microscopy_dataset/	10.1201/9781315215723-10
D3	Publicly Available	https://www.researchgate.net/publication/359064205_Image_features_for_quality_analysis_of_thick_blood_smears_employed_in_malaria_diagnosis	https://doi.org/10.1186/s12936-022-04064-2
D4	AOR		https://doi.org/10.1016/j.bspc.2022.103931
D5	AOR		https://doi.org/10.1007/s00521-021-06602-6
D6	Publicly Available	https://drive.google.com/drive/folders/1p45Dt-BJy8hhoI-rYnhcaL6IMl5FsFL-	https://doi.org/10.3390/a14010017
D7	Publicly Available	https://rdr.ucl.ac.uk/articles/dataset/Giemsa_Stained_Thick_Blood_Films_for_Clinical_Microscopy_Malaria_Diagnosis_with_Deep_Neural_Networks_Dataset_/12173568	10.1002/ajh.25827
D8	Publicly Available	Uninfected: https://data.mendeley.com/datasets/c37wnbbd3c/1 Infected: https://data.mendeley.com/datasets/2v6h4j48cx/1	doi:10.3390/e22060657
D9	Publicly Available	https://github.com/andrealoddo/MP-IDB-The-Malaria-Parasite-Image-Database-for-Image-Processing-and-Analysis	10.1007/978-3-030-13835-6_7
D10	Publicly Available	https://ceb.nlm.nih.gov/repositories/malaria-datasets/	10.7717/peerj.4568
D11	Publicly Available	https://data.lhncbc.nlm.nih.gov/public/Malaria/NIH-NLM-ThickBloodSmearsPV/NIH-NLM-ThickBloodionsPV.zip Thin - TBD	https://doi.org/10.3390/diagnostics11111994

[5] NR – Not reported, AOR- Available on Request

D12	AOR		10.32604/iasc.2022.024361
D13	Publicly Available	https://ai.cdc.gov.tw/datasets/	10.1001/jamanetworkopen.2020.0206
D14	NR		: 10.1007/978-981-15-5281-6_57
D15	NR		https://doi.org/10.1016/j.patrec.2020.07.002
D16	NR		10.1109/AFRICON46755.2019.9133937
D17	NR		10.1109/WiSPNET.2017.8299974
D18	NR		10.1007/978-3-319-75420-8_45
D19	NR		10.1007/978-3-319-75193-1_23
D20	NR		10.1007/s00521-017-2937-4
D21	NR		10.3390/s17102167
D22	NR		Identification Of Malaria Disease and Its Stadium Based On Digital Image Processing
D23	NR		Unsupervised identification of malaria parasites using computer vision
D24	NR		10.1109/KST.2016.7440501
D25	NR		APDS (automated parasite detection system) for field malaria diagnosis in the Philippines
D26	NR		10.1063/1.4915828
D27	NR		10.1111/jmi.12206
D28	NR		10.1136/jclinpath-2019-206419
D29	NR		10.3389/fmed.2023.1117192

D30	Publicly Available	http://fimm.webmicroscope.net/ research/momic/mamic	10.1371/journal.pone.0104855
D31	NR		10.1109/icodse.2014.7062676
D32	Publicly Available	https://doi.org/10.6084/m9.figshare.22679839	https://doi.org/10.1016/j.patter.2023.100806
D33	NR		https://doi.org/10.1016/j.compmedimag.2020.101845
D34	NR		https://doi.org/10.1016/j.imu.2022.101132
D35	NR		10.18517/ijaseit.9.4.4843
D36	NR		https://doi.org/10.1016/j.cmpb.2022.106996
D37	NR		10.3390/diagnostics11030527
D38	NR		10.1017/S2633903X21000015
D39	NR		10.1063/1.4976918
D40	NR		10.1109/OJEMB.2023.3328435
D41	Not publicly available to be shared		https://doi.org/10.1093/ofid/ofad469
D42	NR		https://doi.org/10.1007/978-981-13-3441-2_9
D43	NR		10.1109/KST.2018.8426105
D44	NR		10.4066/biomedicalresearch.29-18-970
D45	NR		http://doi.org/10.11591/ijeecs.v13.i2.pp721-728
D46	NR		10.1109/ACCESS.2020.2990497

D47	NR		https://doi.org/10.1145/3399637.3399641
D48	NR		https://doi.org/10.1007/s11042-019-7162-y
D49	Publicly Available	http://peir-vm.path.uab.edu/wsi.php?slide=IPLab11Malaria	Pathology Education Informational Resource (PEIR) Digital Library
D50	NR		https://doi.org/10.1051/bioconf/20214104002
D51	NR		10.3109/00365548.2014.880186
D52	NR		https://doi.org/10.1145/3429789.3429825
D53	NR		10.1007/s40846-015-0101-0
D54	NR		Adi *et al.,* 2016
D55	NR		10.1109/ICICI-BME.2015.7401339
D56	NR		10.1002/9781118715321.ch8
D57	NR		10.1109/SIET.2017.8304114
D58	NR		https://doi.org/10.1117/12.2249845
D59	NR		https://doi.org/10.1080/03772063.2020.1787238
D60	NR		10.48550/arXiv.2211.15105
D61	Not Available		https://doi.org/10.48550/arXiv.1908.01901
D62	Publicly Available	https://osf.io/3kc2d/	10.1038/s41551-021-00733-w
D63	Publicly Available	https://zenodo.org/records/4314960	https://doi.org/10.5281/zenodo.4314960
D64	Publicly Available	https://drive.google.com/file/d/16T40TdpaB8VXohm50SySREwrzbuPcJBC/view?usp=s	-

		haring	
D65	Publicly Available	https://zenodo.org/doi/10.5281/zenodo.13763938	https://doi.org/10.1016/j.bspc.2024.106869
D66	Publicly Available	https://data.mendeley.com/datasets/6zpxnhjxzz/1	10.1016/j.dib.2024.110950
D67	AOR		10.1186/s13071-024-06215-7
D68	Publicly Available	https://data.mendeley.com/datasets/38jtn4nzs6/2	10.17632/38jtn4nzs6.2
D69	Publicly Available	https://github.com/JonathanRamos/PlasmodiumAI/tree/master/Datasets/Exams/FIOCRUZ	https://doi.org/10.1371/journal.pcbi.1012327
D70	NR		https://doi.org/10.28919/cmbn/8578
D71	NR		10.1109/ACCESS.2024.3393410

Table 5: Geographical information of the selected studies[6]

Identifier	Country	City	Study area
D1	Sudan	Alsororab (SOR) area & Slanj (GS) area	Institute of Endemic Diseases, University of Khartoum, Sudan (IEND)
D2	Uganda	Kampala	Makerere AI Lab, Makerere University
D3	Columbia	NR	NR
D4	Brazil	NR	Instituto Nacional de Pesquisa da Amazˆonia– INPA
D5	Pakistan	Lahore	NR
D6	Uganda	Kampala	Mulago National referral hospital
D7	Nigeria	Ibadan	University College Hospital (UCH)
D8	Spain	Barcelona	Hospital Clínic of Barcelona
D9	Switzerland	Lausanne	Centre Hospitalier Universitaire Vaudois (CHUV)
D10	Bangladesh	Chittagong	Chittagong Medical College Hospital
D11	Bangladesh	Chittagong	Chittagong Medical College Hospital
D12	Malaysia	Kelantan	Hospital Universiti Sains Malaysia, Kubang Kerian
D13	Taiwan	NR	Taiwan CDC Biobank
D14	Malaysia	NR	Medical Microbiology & Parasitology Department, Hospital Universiti Sains Malaysia
D15	India	West Bengal	Midnapore Medical College and Hospital (MMCH), West Bengal
D16	Ghana	Accra	Noguchi Memorial Institute for Medical Research
D17	India	Surat	NR
D18	Philippines	Palawan	College of Public Health, University of the Philippines
D19	Spain	Barcelona	Microbiology department, Vall d'Hebron

[6] NR – Not Reported

			Hospital, Barcelona
D20	India	Silchar	Silchar Medical College and Hospital
D21	Portugal	NR	Instituto Nacional de Saúde Dr. Ricardo Jorge, Portugal
D22	Indonesia	NR	Central Health Laboratory North Sumatera Province
D23	NR	NR	NR
D24	Thailand	NR	NR
D25	Philippines	NR	Brgy. Luzviminda, San Rafael, and Puerto Princesa City Proper.
D26	Malaysia	NR	Department of Microbiology and Parasitology, Hospital Universiti Sains Malaysia (HUSM) and Hospital Universiti Kebangsaan Malaysia (HUKM).
D27	India	West Bengal	Midnapur Medical College and Hospital; and Medipath Laboratory, Midnapur, West Bengal, India
D28	NR	NR	NR
D29	China	Wuhan	Department of Schistosomiasis and Endemic Diseases, Wuhan City Center
D30	Finland	Helsinki	Helsinki University Central Hospital Laboratory
D31	Indonesia	NR	Eijkman Institute for Molecular Biology, Indonesia
D32	Sierra Leone	NR	Sierra Leone-China Friendship Hospital and the Rokupa Government Hospital
D33	NR	NR	Interstate blood bank (USA)
D34	France	NR	Parasitology lab of the Piti´e-Salpˆetri`ere Malaria NRC (National Reference Center) and the hematology lab of Saint Antoine Hospital

D35	Indonesia	NR	Eijkman Institute for Molecular Biology
D36	Pakistan	Punjab	Province of Punjab
D37	Korea	NR	Korea University Guro Hospital
D38	NR	NR	NR
D39	NR	NR	Hydas World Health
D40	Sudan	White Nile State	Alkawa Hospital, Alkawa, White Nile State, Sudan
D41	China	NR	Peking Union Medical College Hospital and Yunnan Institute of Parasite Diseases
D42	Indonesia	Jakarta	Bina Medical Support Services (BPPM)
D43	Thailand	NR	National Center for Genetic Engineering and Biotechnology (BIOTEC)
D44	India	Manipal	Haematology Lab, Kasturba Medical College (KMC), Manipal Academy of Higher Education (MAHE), Manipal
D45	Indonesia	NR	Department of Parasitology, Faculty of Medicine, Universitas Gadjah Mada
D46	Thailand	Pathum Thani	Protein-Ligand Engineering and Molecular Biology laboratory at the National Center for Genetic Engineering and Biotechnology (BIOTEC)
D47	China	Wuhan	Central South Hospital of Wuhan University
D48	India	Chennai	Tagore Medical College & Hospital, Chennai, India
D49	NR	NR	University of Alabama at Birmingham, Department of Pathology
D50	Indonesia	NR	Parasitology Laboratory, Faculty of Medicine, Universitas Gadjah Mada
D51	Iran	NR	NR
D52	Indonesia	NR	Eijkman Institute of Molecular Biolog
D53	NR	NR	China Medical University, China Medical

			University Hospital, and Taichung Veterans General Hospital in Taiwan
D54	NR	NR	NR
D55	Indonesia	NR	Dinas Kesehatan Provinsi Jawa Timur (East Java Health Department)
D56	Singapore	NR	NR
D57	Indonesia	NR	Balai Besar Laboratorium Kesehatan (Center for Health Laboratory) Indonesian Ministry of Health Surabaya
D58	NR	NR	Bernhard Nocht Institute for Tropical Medicine (BNITM)
D59	India	Assam	Pathological clinic of Cachar district, Assam, India
D60	Indonesia	NR	Eijkman Institute and the medical faculty of UGM
D61	South America, Africa, Asia, and London	NR	
D62	NR	NR	Massachusetts General Hospital (MGH) IVF laboratory and the MGH clinical pathology laboratory
D63	Tanzania	White Nile State	The University of Dodoma and Benjamin Mkapa Hospital Research Center
D64	Uganda	NR	NR
D65	Nigeria	Kano State	Asiya Bayero pediatric hospital, kano state, Nigeria & Murtala Muhammad specialist hospital
D66	Nigeria	Osun State, Oyo State and Cross River	Bowen University Hospital Iwo and Bowen University Teaching Hospital Ogbomoso and Cross river state
D67	Malaysia	NR	Malaria Research Centre, UNIMAS, Sarawak

			(MRC-UNIMAS)
D68	NR	NR	NR
D69	Brazil	Porto Velho, Rondônia	Fundação Oswaldo Cruz (FIOCRUZ)
D70	Indonesia	Surabaya	Department of Clinical Pathology, Faculty of Medicine, Airlangga University
D71	Cyprus	NR	Department of Parasitology, Manisa Celal Bayar University, and the Microbiology Laboratory of Near East University Hospital

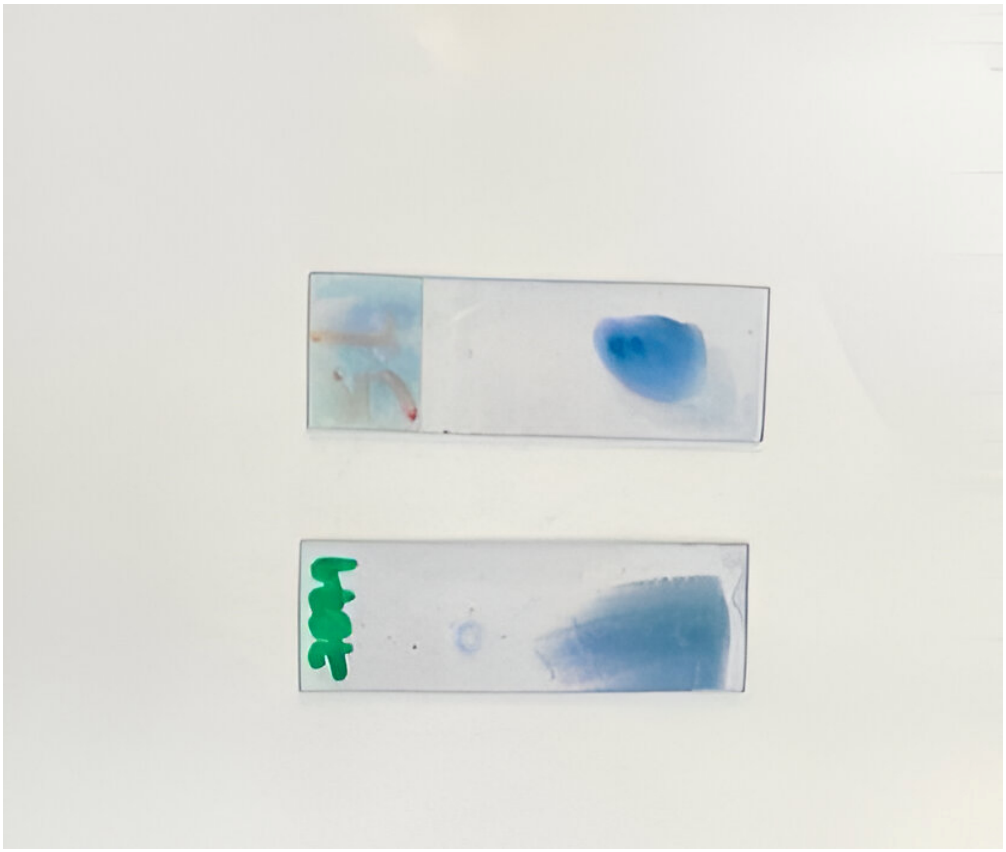

Figure 1

153x129mm (138 x 138 DPI)

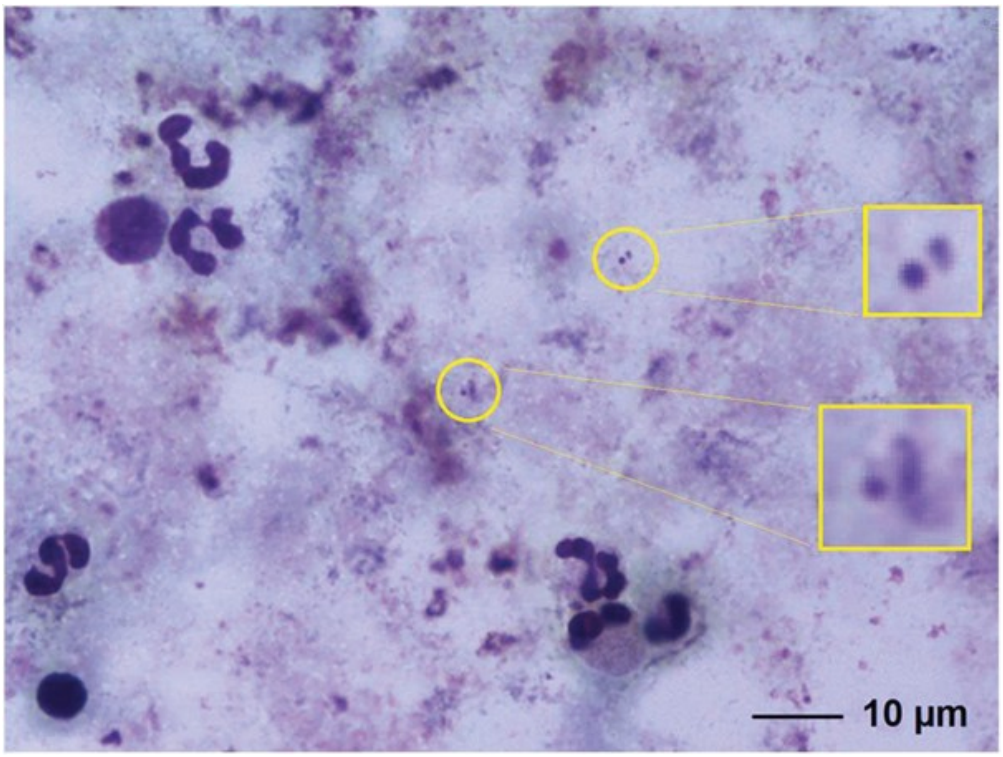

Figure 2

111x83mm (138 x 138 DPI)

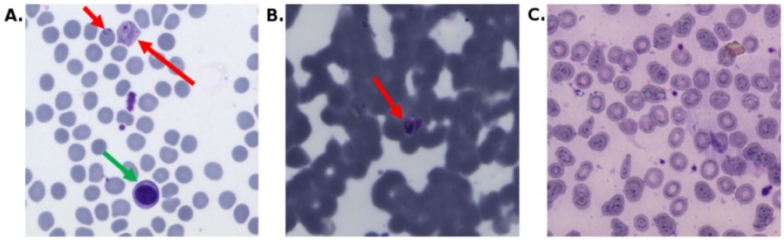

Figure 3

263x69mm (138 x 138 DPI)

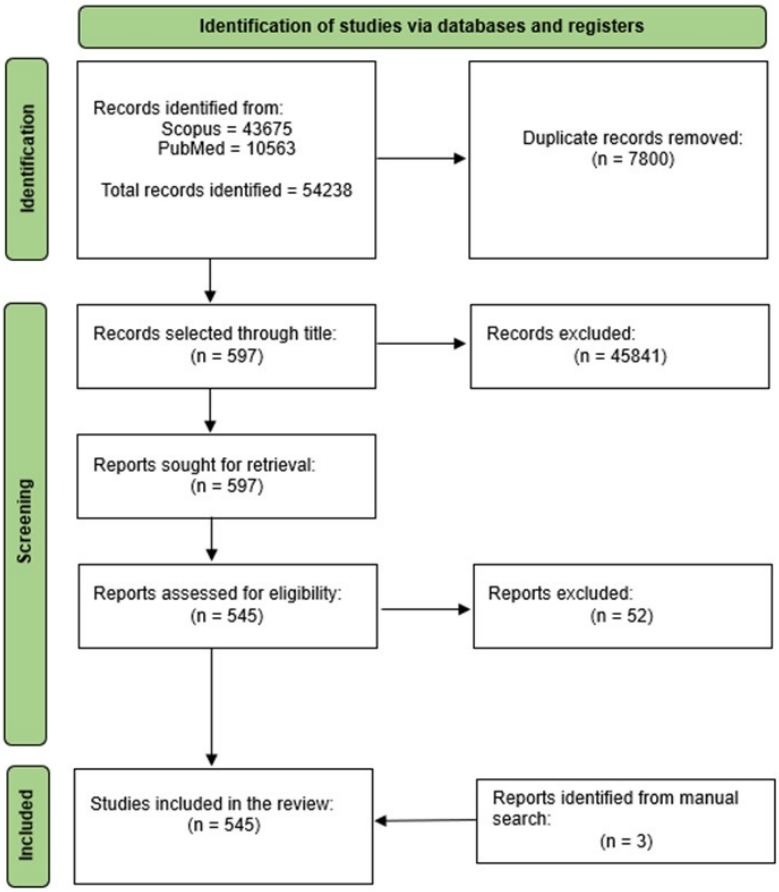

Figure 4

162x150mm (138 x 138 DPI)

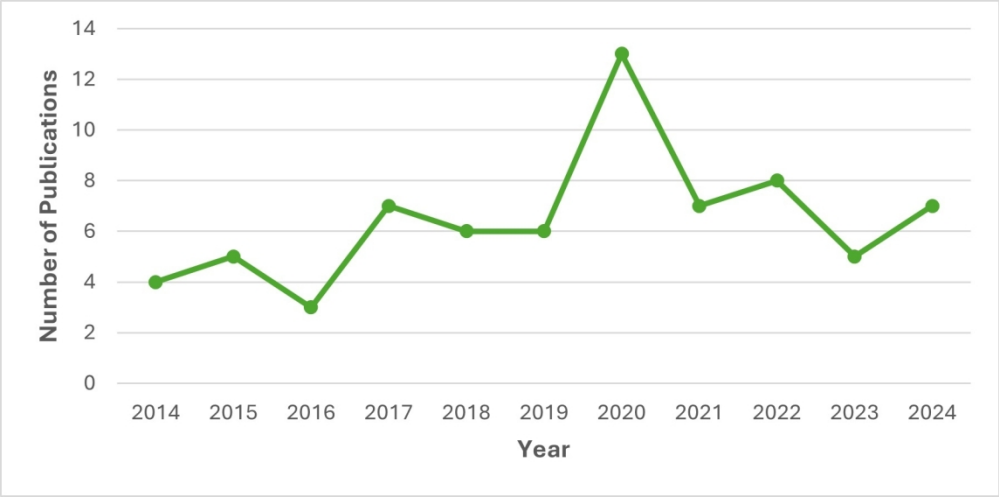

Figure 5: Summary of all publications included in the study by year of publication

138x68mm (330 x 330 DPI)

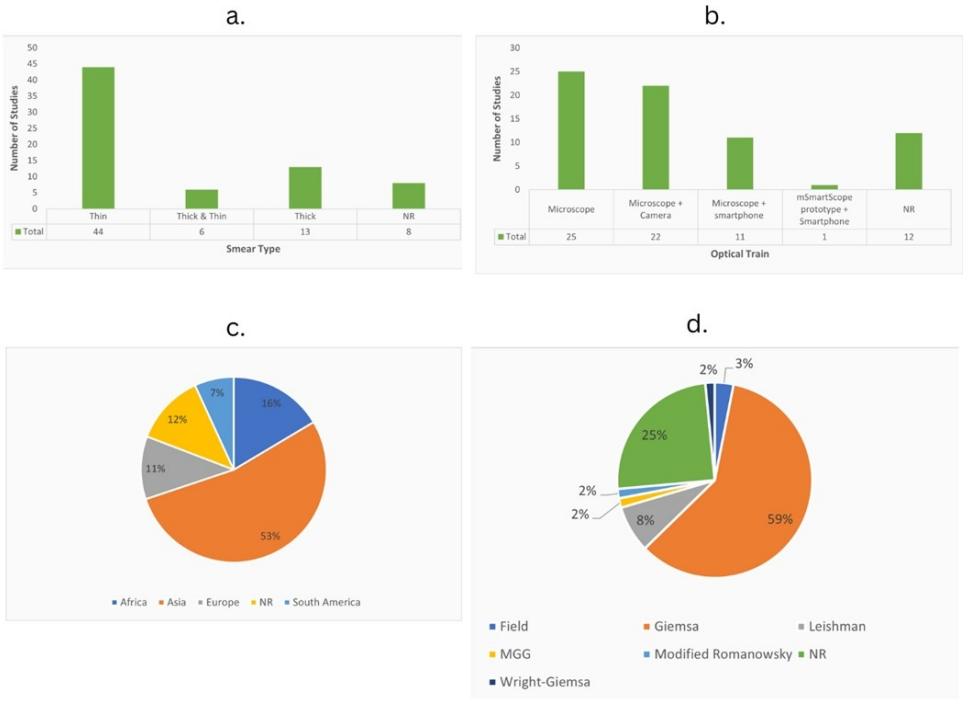

Figure 6: (a) Type of smear used in selected studies (b) Imaging techniques used in selected studies (c) Distribution of datasets by continent (d) Staining techniques used in selected studies

161x129mm (171 x 171 DPI)