UNIVERSIDAD POLITÉCNICA DE MADRID
Escuela Técnica Superior de Ingenieros de Telecomunicación

Advancing AI Diagnostics for Global Health through Microscopy Image Analysis of Infectious Diseases

D O C T O R A L T H E S I S

Submitted for the degree of Doctor by:

Lin Lin
MSc in Biomedical Engineering

Madrid, 2024


UNIVERSIDAD POLITÉCNICA DE MADRID
Escuela Técnica Superior de Ingenieros de Telecomunicación

Doctoral Degree in Electronic Systems Engineering

Advancing AI Diagnostics for Global Health through Microscopy Image Analysis of Infectious Diseases

D O C T O R A L T H E S I S

Submitted for the degree of Doctor by:

Lin Lin
MSc in Biomedical Engineering

Under the supervision of:
Dr. María Jesús Ledesma Carbayo
Dr. Miguel Ángel Luengo Oroz

Madrid, 2024

Title: Advancing AI Diagnostics for Global Health through Microscopy Image Analysis of Infectious Diseases

Author: Lin Lin

Doctoral Programme: Electronic Systems Engineering

Thesis Supervision:

Dr. Maria Jesús Ledesma Carbayo, Universidad Politécnica de Madrid(Supervisor)

Dr. Miguel Ángel Luengo Oroz, Spotlab(Supervisor)

External Reviewers:

Thesis Defense Committee:

Thesis Defense Date:

This thesis has been partially supported by the predoctoral grant IND2019/TIC-17167 (Comunidad de Madrid)

Agradecimientos

Al iniciar mi carrera, no tenía conocimiento alguno sobre imágenes biomédicas, ni había considerado la posibilidad de dedicarme a la investigación, o más bien, de confiar en que podía hacerlo. Sin embargo, mi tutor de TFG, Carlos Platero, depositó tanta confianza en mí que me introdujo en este fascinante mundo y me animó a comenzar este viaje. Gracias por tu apoyo; como solías decir, "debemos saber, sabremos". También agradezco a José Pineda, ya que si no fuera por ti, no habría conocido a Spotlab y comenzado este doctorado industrial.

A mis directores de tesis, Miguel y Chus, les agradezco por brindarme la oportunidad de formar parte de este maravilloso entorno de colaboración entre la academia y la empresa, donde he podido ver cómo mi investigación se aplica en el mundo real. Gracias por sus valiosos consejos, guías, ideas brillantes y constante apoyo.

A todos los miembros de Spotlab, este proyecto no habría sido posible sin vosotros. Al equipo de IA, David, Roberto y Álvaro, por siempre ofrecerme ideas y discutir los novedades y desafíos relacionados con la inteligencia artificial. Al equipo clínico, Elena, María, Nuria, Lucía, Ramón, David Brau, y aunque ya no estén con nosotros, a Adri y Eli, por gestionar todos los proyectos y conseguir nuevas oportunidades y partners. Al equipo de app, Jaime y Óscar, por vuestra trabajo para implementar algoritmos de IA en dispositivos móviles y hacerlos realidad. Al equipo de cloud, Alex, Carla, Alejandro, y aunque ya no forman parte Ávaro y Endre, por desarrollar Telespot, lo que ha permitido que todo el flujo de desarrollo sea posible. A Dani, por ser el puente entre el ámbito técnico y clínico. A Eva, por tu ayuda en todos los procesos a lo largo de estos años.

También quiero agradecer al grupo BIT, a Dani, Pedro, David, Feli, Benito y Ana, Andrés, Juan y Juanjo, por todo lo que he aprendido de ustedes y por los viajes que hemos compartido.

A mi familia y amigos de siempre, gracias por su apoyo incondicional.

Y, por último, a Dios, por estar siempre a mi lado.

Lin Lin

Septiembre 2024

Abstract

Global health is the study, research, and implementation of efforts aimed at improving health outcomes and achieving equity for all people, regardless of geographic, socioeconomic, or political boundaries. One of the major challenges in global health is the persistent burden of infectious diseases, especially in low- and middle-income countries (LMICs). Despite advancements in medicine and technology, millions of people continue to suffer from preventable and treatable conditions. Infectious diseases such as malaria, tuberculosis, HIV/AIDS, and neglected tropical diseases like schistosomiasis, filariasis, and soil-transmitted helminths (STH) remain leading causes of morbidity and mortality. These diseases disproportionately affect vulnerable populations, perpetuating a cycle of poverty and poor health outcomes. To address these challenges, the World Health Organization has set universal health coverage as one of its key Sustainable Development Goals to promote health equity and combat these diseases on a global scale.

However, LMICs face severe shortages of healthcare professionals, leading to delays in diagnosis and treatment, particularly due to a lack of trained laboratory technicians. While microscopy is widely used for disease diagnosis, it has limitations, including the need for skilled personnel, time consumption, and susceptibility to human error, especially in resource-limited settings. Artificial intelligence (AI) has the potential to revolutionize global health diagnostics, yet its integration in LMICs remains limited due to challenges such as infrastructure limitations and access to data.

The primary objective of this PhD thesis is to design, develop, and validate AI-driven solutions to assist in the diagnosis of global health diseases, with a specific focus on the neglected tropical diseases, starting with the two most prevalent, helminthiasis and filariasis. To achieve this goal, the study involved: 1) establishing a comprehensive workflow for data collection, digitization, and labeling; 2) mitigating the scarcity of labeled data in medical imaging; 3) designing and implementing AI development and deployment workflows; and 4) evaluating AI models and their integration in resource-limited settings.

As a result: 1) we created a dataset for soil-transmitted helminthiasis with over 1,300 samples and approximately 20,000 images, as well as a dataset for filarial samples containing 130 samples and 2,800 images; 2) we explored crowdsourcing for image labeling, demonstrating that labels gathered from the general public can be effectively used to train deep learning models; 3) we developed object detection algorithms to identify soil-transmitted helminthiasis eggs, enabling the integration of the AI model on smartphones. Additionally, we participated in a field study during a deworming campaign in Kenya, assessing the feasibility of deploying AI-driven tools in LMICs and addressing challenges such as infrastructure, data availability, and user training; 4) we applied validated methods from the previous study to filariasis, creating a deep learning algorithm for the detection of microfilariae. This system was validated in a laboratory setting, evaluating not only model performance but also usability and effectiveness in real-world applications; 5) we established a pipeline for generating foundational models for stool samples, which will serve as the building blocks for a system capable of diagnosing multiple parasites in fecal samples and performing various related tasks.

The outcomes of this PhD thesis highlight the effectiveness of AI-assisted diagnosis for microscopy-based images, the feasibility of deploying AI algorithms in resource-constrained settings, and steps toward creating trustworthy AI systems. These techniques can be expanded to other diseases, contributing to the achievement of universal health coverage and improving global health outcomes.

Resumen

La salud global es el estudio, la investigación y la implementación de esfuerzos dirigidos a mejorar los resultados en salud y lograr la equidad para todas las personas, independientemente de las localización geográficas, socioeconómicas o políticas. Uno de los principales desafíos en la salud global es la carga persistente de enfermedades infecciosas, especialmente en los países de ingresos bajos y medianos (PIBM). A pesar de los avances en medicina y tecnología, millones de personas continúan sufriendo de condiciones que son prevenibles y tratables. Enfermedades infecciosas como la malaria, la tuberculosis, el VIH/SIDA y enfermedades tropicales desatendidas como la esquistosomiasis, la filariasis y los helmintos transmitidos por el suelo siguen siendo causas principales de morbilidad y mortalidad. Estas enfermedades afectan desproporcionadamente a poblaciones vulnerables, perpetuando un ciclo de pobreza y malos resultados en salud. Para abordar estos desafíos, la Organización Mundial de la Salud ha establecido la cobertura universal de salud como uno de sus principales Objetivos de Desarrollo Sostenible para promover la equidad en salud y combatir estas enfermedades a nivel global.

Sin embargo, los PIBM enfrentan severas carencias de profesionales de la salud, lo que provoca retrasos en el diagnóstico y tratamiento, particularmente debido a la falta de técnicos de laboratorio capacitados. Aunque la microscopía se utiliza ampliamente para el diagnóstico de enfermedades, tiene limitaciones, incluyendo la necesidad de personal capacitado, el consumo de tiempo y la susceptibilidad al error humano, especialmente en entornos con recursos limitados. La inteligencia artificial (IA) tiene el potencial de revolucionar los diagnósticos de salud global; sin embargo, su integración en los PIBM sigue siendo limitada debido a desafíos como las limitaciones de infraestructura y el acceso a datos.

El objetivo principal de esta tesis doctoral es diseñar, desarrollar y validar soluciones impulsadas por IA para asistir en el diagnóstico de enfermedades de salud global, con un enfoque específico en las enfermedades tropicales desatendidas, empezando con los dos más prevalentes, helmintiasis y filariasis. Para lograr este objetivo, el estudio implicó: 1) establecer un flujo de trabajo integral para la recolección, digitalización y etiquetado de datos; 2) mitigar la escasez de datos etiquetados en imágenes médicas; 3) diseñar e implementar flujos de trabajo para el desarrollo y despliegue de IA; y 4) evaluar modelos de IA y su integración en entornos con recursos limitados.

Como resultados: 1) creamos un conjunto de datos para helmintiasis con más de 1,300 muestras y 20,000 imágenes, y otro de filariasis con 130 muestras y 2,800 imágenes; 2) exploramos el crowdsourcing para el etiquetado de imágenes, demostrando su utilidad para entrenar modelos de aprendizaje profundo; 3) desarrollamos algoritmos de detección de huevos de helmintos y los integramos en teléfonos móviles. Además, en una campaña de desparasitación en Kenia, evaluamos la viabilidad de desplegar estas herramientas en PIBM, enfrentando desafíos como infraestructura y capacitación de usuarios; 4) aplicamos estos métodos a la filariasis, creando un algoritmo para detectar microfilarias, validado en laboratorio tanto en rendimiento como en usabilidad; 5) definimos un flujo de trabajo para desarrollar modelos fundacionales que permitan diagnosticar múltiples parásitos en muestras fecales.

Los resultados de esta tesis doctoral destacan la efectividad del diagnóstico asistido por IA para imágenes basadas en microscopía, la viabilidad de implementar algoritmos de IA en entornos con recursos limitados y los pasos hacia la creación de sistemas de IA confiables. Estas técnicas pueden ser aplicadas a otras enfermedades, contribuyendo al logro de la cobertura sanitaria universal y mejorando los resultados de salud global.

Table of Contents

Acknowledgement . . . iii
Abstract . . . iv
Resumen . . . vi
List of Figures . . . xi
List of Tables . . . xv
Abbreviations and acronyms . . . xix

1 Introduction and objectives **1**
1.1 Introduction . . . 1
1.2 Context of the research . . . 3
1.3 Objectives . . . 3
1.3.1 Technical Methodology Overview . . . 4
1.3.2 AI Development Workflow and Engineering Infrastructure . . . 4
1.4 Document structure . . . 7

2 Clinical context **9**
2.1 Neglected tropical diseases . . . 9
2.1.1 Prevention and control of NTDs . . . 12
2.1.2 Diagnosis of NTDs . . . 12
2.1.3 Soil-transmitted helminth . . . 13
2.1.4 Filariasis . . . 16
Filariasis species . . . 17
Transmission of filariasis . . . 19
Diagnosis of filariasis . . . 19

3 State of the art **23**
3.1 Artificial intelligence in medical image analysis . . . 23
3.2 Diagnosis of microscopical images . . . 24
3.2.1 Early approaches and traditional image processing techniques . . . 25
3.2.2 Deep learning techniques . . . 25
Convolutional neural networks . . . 25
Vision transformers . . . 28
Self supervised learning . . . 28
3.3 AI: from research to clinical deployment . . . 30
3.3.1 Definition of intended use . . . 30

3.3.2 Data collection . . . 30
3.3.3 Data labelling . . . 33
3.3.4 Algorithm development and retrospective validation . . . 34
3.3.5 Prospective validation . . . 35
3.3.6 Regulation . . . 36
3.3.7 Integration into the clinical workflow . . . 38
3.4 Research gaps . . . 38

4 Combining collective and artificial intelligence for global health diseases diagnosis using crowdsourced annotated medical images 39
4.1 Introduction . . . 39
4.2 Methodology . . . 40
4.2.1 Crowdsourcing image annotation . . . 40
4.2.2 AI architecture: Deep learning model . . . 42
4.3 Experiments and Results . . . 42
4.3.1 Quality of crowdsourced annotations . . . 43
4.3.2 DL model: hyperparameter tuning . . . 44
4.3.3 Incremental training . . . 44
4.3.4 Differences between annotator groups . . . 44
4.4 Conclusions . . . 45
4.5 Environmental Impact . . . 46

5 Deployment, adaptation and evaluation in real time and real world settings of mobile AI assistants for soil-transmitted helminths quantification 47
5.1 Introduction . . . 47
5.2 Material and methods . . . 49
5.2.1 Phase 1: Baseline model . . . 50
5.2.2 Phase 2: Pilot study . . . 51
5.2.3 Phase 3: Post pilot study . . . 53
5.3 Experiments and results . . . 55
5.3.1 Phase 1: Baseline model . . . 55
5.3.2 Pilot study . . . 55
Usability feedback . . . 56
5.3.3 Post pilot study . . . 60
5.4 Conclusions . . . 60

6 Edge Artificial Intelligence (AI) for real-time automatic quantification of filariasis in mobile microscopy 65
6.1 Introduction . . . 65
6.2 Material and methods . . . 67
6.2.1 Ethics statement . . . 67
6.2.2 Overview of the methodology . . . 68
6.2.3 Creating a Filariasis differentiation AI model . . . 68
Digitalize samples . . . 68
Labeling data . . . 69

Creating the AI model . . . 70
6.2.4 Validation of the AI model in a lab setting . . . 73
Deployment and integration of technology . . . 74
Experiment- Pilot replicating diagnostic workflow . . . 75
6.3 Results . . . 76
6.3.1 Evaluation of the AI model performance . . . 76
6.3.2 Validation of the AI-assisted mobile app . . . 77
Real-time AI-Performance . . . 78
Inter-observer variability . . . 78
6.3.3 Analysis time . . . 79
6.4 Discussion and Conclusion . . . 79

7 Towards a foundation model for fecal sample analysis **83**
7.1 Introduction . . . 83
7.2 Material and methods . . . 85
7.2.1 Dataset . . . 85
7.2.2 Overview of the proposed foundational method . . . 85
7.2.3 Slide-level parasite detection . . . 87
7.2.4 Patch-level parasite classification . . . 87
7.2.5 Experimental setup . . . 88
7.3 Experiments and results . . . 88
7.3.1 Patch-level parasite classification . . . 89
7.3.2 Evaluation of the proposed system on an independent test set . . . 90
7.4 Conclusions . . . 91

8 Concluding remarks **93**
8.1 Future works . . . 95

9 Publications **97**
9.1 Journal articles . . . 97
9.2 Conference articles and abstracts . . . 99
9.3 Posters and Oral Presentations at Medical Conferences . . . 100

References **103**

List of Figures

1.1 AI use cases in healthcare. Source: (USAID, 2019) 2

1.2 Methodology overview of this PhD thesis.Supervised and self-supervised techniques highlighted in blue and green, respectively. 5

1.3 The AI development workflow, encompassing data acquisition, image visualization, storage, labeling, model design, and development—including data preprocessing, model training, and evaluation—is an iterative process. The best-performing models are exported for cloud or edge inference, followed by prospective validation to assess their real-world effectiveness. 6

2.1 a) Number of people requiring treatment against neglected tropical diseases. Source: (Roser et al., 2024). b) GDP per capital. Data source: World bank . 10

2.2 Combined STH life-cycle. Source (Schlosser-Brandenburg et al., 2023) 14

2.3 Number of children requiring preventive chemotherapy for STH in 2022. Modified from (WHO, n.d.-c) 15

2.4 Soil-transmitted helminth eggs illustration, adapted from (WHO, 2019) . . . 16

2.5 Pathology caused by different fialariasis. a) elephansiasis, caused by lymophatic filariasis; 6) blindness caused by onchocersiasis; c) worms in eyes, caused by loasis. 18

2.6 Lymphatic filariasis lifecycle. Source: (CDC-DPDx, n.d.-a) 19

2.7 Key features of microfilariaes. Source (Sastry & Bhat, 2018) 21

3.1 Current workflow for microscopical image analysis 24

3.2 Example image classification pipeline. First, object of interest is segmented, and then it is cropped into small patches, which are passed through the CNN, and the CNN predicts the corresponding class 26

3.3 Object detection pipeline. The whole image is passed to the CNN and both location and class is predicted 27

3.4 Class imbalance problem in microscopical images. a) Foreground-background imbalance. b) class imbalance 29

3.5 Examples of image acquisition devices used for malaria and NTD diagnosis. A) Ball-lens attached to a smartphone, proposed by Bogoch et al. for the diagnosis of *Schistosoma haematobium* (Bogoch et al., 2013). B) LoaScope, proposed by D'Ambrosio et al. for the diagnosis of *Loa loa* (D'Ambrosio et al., 2015). C) MoMic digital microscope scanner, proposed by Holmström et al. for the diagnosis of *Schistosoma haematobium*. D) Commercial single slide scanner (Grundium), used to scan soil-transmitted helminth (STH) samples (Lundin et al., 2024). E) Schistoscope, proposed by Oyibo et al. for scanning *Schistosoma haematobium* samples. F) Smartphone attached to an optical microscope, employed by various authors for digitizing microscopical images (Dacal et al., 2021; Morais et al., 2022; Quinn et al., 2016; F. Yang et al., 2020) 32

3.6 Comparison of the same sample preparation digitized using three different smartphones, highlighting the potential variability and bias in data collection due to differences in camera quality, image resolution, and color balance. . . 33

3.7 Approved AI by FDA in different medical domains. Data source (FDA, 2021) 37

4.1 Screenshot of the mini-game used to collect data with *Ascaris* spp. (left) and *Trichuris* spp. (right) 41

4.2 Evolution of the model performance (AUC) as the training sample size increases. Results from data annotated by children, adults and experts are presented independently. 45

4.3 Examples of image samples representing all classes under study (upper row) along whit its activation maps generated by the DL algorithm (bottom row). Correct label (GT) as well as predictions (Pred) appear above each image. Activation maps for healthy samples (HEA) focused on the entire image with no significant activation, while activation maps for *Ascaris* spp. (ASC) and *Trichuris* spp. (TRI) focused exactly on the egg location. 46

5.1 Digitisation and labeling system. (left) digitisation system, By aligning the smartphone camera with the microscope's ocular with the smartphone, the conventional microscope is converted to a digital microscope (right) Telemedicine platform for image visualisation and labeling. 51

5.2 Real-time AI-assisted digitization and analysis of Kato-Katz samples by analysts 52

5.3 Workflow of the AI model creation and deployment during the pilot study . 53

5.4 Human in the loop review. Instead of manually replacing the bounding boxes, analyst only has to review the predictions 55

5.5 Correlation of egg counting with conventional microscopy and AI assistance, where R is the correlation coefficient. 59

6.1 Schematic representation of the study design. 68

6.2 The telemedicine platform facilitates image visualization, management, and labeling. When an AI algorithm is deployed, analysts have the option to review the predictions rather than starting the labeling process from scratch. 70

6.3 Simplified SSD MobileNet v2 detection architecture. MobileNet V2 initiates with one convolutional layer, succeeded by a depthwise convolutional layer and another convolutional layer. It is subsequently followed by 16 bottleneck residual blocks (green), each comprising two 2D convolutional layers and one depthwise convolutional layer, concluding with an additional convolutional layer. The SSD feature extractor is further enhanced by integrating five additional feature extractor blocks (purple). The resultant features are passed through a convolutional box predictor block (gray), which is responsible for predicting both the location and class of each detection. 71
6.4 a): example input of the species differentiation algorithm. b): green rectangle represents the sliding windows size 73
6.5 Mosaic augmentation. The original image is 640x640 pixels, for each image, we randomly select an area of 320x320 pixels that contains at least one parasite (green rectangle), using 4 cropped areas we compose a new mosaic image. . . 74
6.6 (a) smartphone attached to the conventional microscope with a 3D printed adapter. (b) screening algorithm working on the smartphone. (c) species differentiation working on the smartphone. 75
6.7 Schema represents the validation workflow of AI assisted filariae detection. At least 3 images of negative fields were acquired for each sample. 76

7.1 Overview of the proposed pipeline comprising: (1) image acquisition using a 3D-printed adapter; (2) field-of-view ROI extraction; (3) parasite detection using YOLOv8; (4) self-supervised pre-training with DINO; (5) species differentiation with our DINO pretrained ViT classifier. 86
7.2 Overview of the proposed pipeline for training a ViT architecture with DINO SSL strategy. 88
7.3 Uniform manifold approximation and projection (UMAP) visualization of the features extracted from our pre-trained ViT architecture. 89

List of Tables

4.1 Distribution of the training, validation and test sets. Training images are annotated by three groups: experts (E), school-age children (SAC) and adults (A). 41
4.2 Mean accuracy and the standard deviation of crowdsourced annotations using different quorum sizes compared to the ones made by experts. 43
4.3 Hyperparameter selection. Mean AUC and standard deviation are shown. . . 44
4.4 Detailed performance of the DL algorithm trained with different annotations. Mean AUC and standard deviation is reported. 45

5.1 Training and validation data distribution for each iteration, where model 0 is the baseline for model 1-4. Model 1-4 were deployed during the pilot study . 54
5.2 Data distribution of the latest AI model 55
5.3 Overview of the results obtained for the detection of both *Trichuris* spp. and *Ascaris* spp. eggs. with the baseline model 56
5.4 STH prevalence in Kwale in 2021. Result obtained with conventional microscopy. AS: *Ascaris lumbricoides*, TR: *Trichuris trichura*, HW: hookworm 57
5.5 Model performance on validation set and on real time during the pilot study. Note that the performance between models are not comparable, as they are evaluated on different dataset. 58
5.6 Confusion matrix of the model 5 on the unique test set 61
5.7 Performance of the model on the unique test set 61

6.1 Cases included in the first phase and the training-validation split. 69
6.2 Label distribution of microfilaria species in the training and validation sets. . 70
6.3 Confusion matrix of the species differentiation algorithm (40x) on the validation set, each row represents the ground truth, and each column represents the prediction. The model may predict the artifacts as a parasite (false positive), but the analysts did not label all artifacts. 77
6.4 Performance of the AI algorithm on pilot study using mobile phone. Each row represents the ground truth and each column represents the AI prediction. . 78
6.5 Inter-observer agreement of detected parasites when analyzing with and without AI assistance of 2 experts and of the AI. Two-tailed t-test indicates that the analysis of analist A and B are significantly correlated.* p-value<0.05. 79
6.6 Average analysis time in seconds for both analysts with and without AI assistance for each image. 79

7.1 Data Distribution: Breakdown of unannotated (SSL) and annotated (Train; Validation; Test) data sets. Expert labels include Ascaris, Trichuris, and hookworm classes, along with artifacts (false positives generated by the detection algorithm). 86
7.2 Parasite classification performance on the validation set, obtained from full fine-tuning and linear probing 90

Abbreviations and acronyms

AI	Artificial intelligence
ACC	Accuracy
AUC	Area Under the Curve
AWS	Amazon Web Services
CNN	Convolutional Neural Networks
FDA	Food and Drug Administration
FOV	Field of View
ISCIII	Instituto de Salud Carlos III
KEMRI	Kenya Medical Research Institute
LAMP	Loop-Mediated Isothermal Amplification
LMICs	Low middle income countries
MAD	Mass Drug Administration
MAP	Mean Average Precision
NLP	Natural Language Processing
NTDs	Neglected tropical diseases
PC	Preventive chemotherapy
PCR	Polymerase Chain Reaction
PRE	Precision
REC	Recall
ROI	Region of Interest
SaMD	Software as a Medical Device
SSL	Self Supervised Learning
SSD	Single Shot Multibox Detector
STH	Soil-transmitted helminths

ViT Vision Transformer

WHO World Health Organization

Chapter 1

Introduction and objectives

1.1 Introduction

Global health is the study, research and implementation efforts focused on enhancing health and attaining equitable health outcomes for all people worldwide, irrespective of geographic, social-economic or political boundaries (Koplan et al., 2009). Global health faces numerous challenges, with infectious diseases remaining a significant burden, particularly in low- and middle-income countries (LMICs). Despite advancements in medicine and technology, millions of people worldwide continue to suffer from preventable and treatable conditions. Infectious diseases such as malaria, tuberculosis, and HIV/AIDS, as well as neglected tropical diseases (NTDs) like schistosomiasis, filariasis, and soil-transmitted helminths (STH), are among the leading causes of morbidity and mortality. These diseases disproportionately affect vulnerable populations, contributing to a cycle of poverty and poor health outcomes. In order to promote health equity, combating infectious diseases (malaria, neglected tropical diseases, HIV, etc), addressing non-communicable diseases, improving access to healthcare services, advancing health systems strengthening, and responding to health emergencies on a global scale, the World Health Organization (WHO) defined included achieving universal health coverage as one of the sustainable development goals (UN, 2015).

A critical challenge facing global health, particularly in low- and middle-income countries (LMICs), is the severe shortage of healthcare professionals. With increasing life expectancy, there is a growing population of older adults, with over 30% of individuals aged 65 or older reported having at least two chronical diseases (OECD, 2020). This demographic shift poses a challenge for healthcare systems as they contend with a larger number of patients with complex medical needs. Nevertheless, the scarcity of healthcare personnel is worsening. The WHO estimates that there is a global shortfall of approximately 18 million health workers, with the most significant gaps found in LMICs (WHO, 2016). This shortage exacerbates the difficulties in diagnosing and managing diseases, as overworked and under-resourced healthcare systems struggle to meet the needs of their populations. In many areas, there is a lack of trained laboratory technicians and specialists capable of performing accurate diagnostics, leading to delayed or missed diagnoses, and contributing to poor health outcomes. This workforce deficit also places immense strain on existing healthcare workers, increasing

the likelihood of errors and reducing the overall quality of care. Addressing this shortage is essential for improving global health (Babalola & Moodley, 2020; Glied & Sacarny, 2018; OECD, 2017).

The diagnosis and management of multiple diseases, such as haematology and infectious diseases often rely on microscopy, which, while effective, have several limitations. Microscopy requires skilled technicians, is time-consuming, and is prone to human error, particularly in resource-limited settings. Additionally, the quality and accuracy of microscopic diagnosis can vary widely depending on the level of training and experience of healthcare workers, as well as the quality of the equipment used. These challenges underscore the need for innovative solutions that can enhance diagnostic accuracy, reduce the burden on healthcare systems, and improve health outcomes in global health.

Artificial intelligence (AI), particularly through advancements in deep learning, has emerged as a transformative technology with the potential to revolutionize healthcare, as depicted on Figure 1.1. AI algorithms have demonstrated exceptional performance in various medical imaging tasks, including the detection of diseases, segmentation of medical images, and classification of pathologies (Rajpurkar & Lungren, 2023; Topol, 2019).

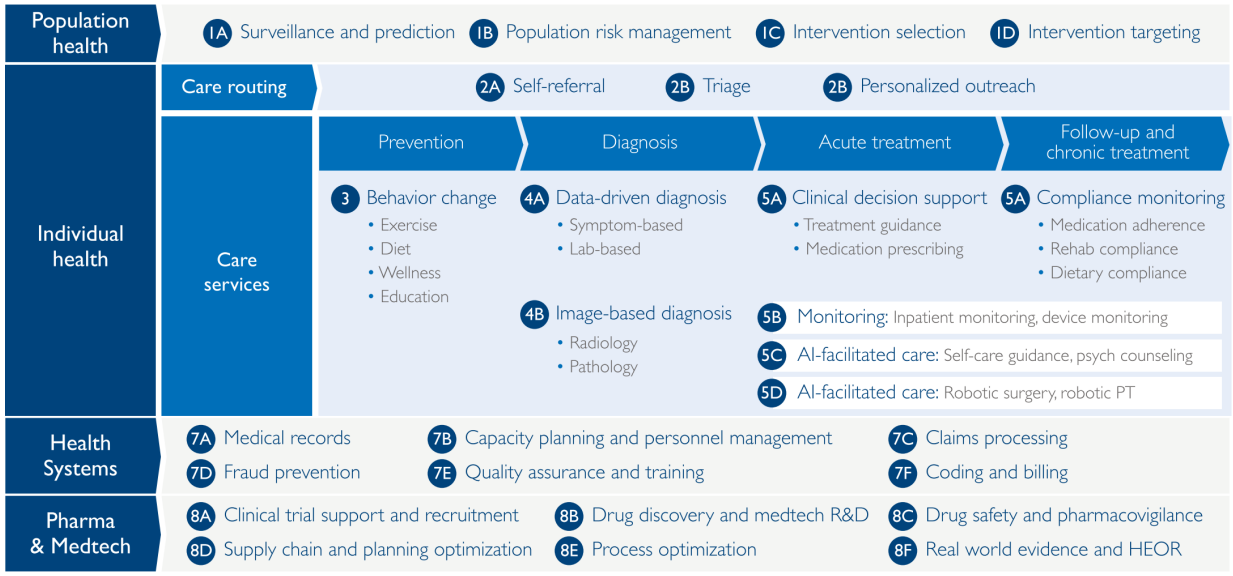


Figure 1.1: AI use cases in healthcare. Source: (USAID, 2019)

The deployment of AI in these areas has shown promising results, leading to a growing number of AI-enabled medical devices being approved by regulatory bodies. Between 2015 and 2020, 222 AI devices received approval in the United States, while 240 were authorized in the European Union, reflecting the rapid adoption of this technology in healthcare. As of 2024, more than 900 AI/ML-enabled medical devices have been approved by the U.S. Food and Drug Administration (FDA), highlighting the accelerating integration of AI into clinical practice. These approvals signify AI's potential to enhance diagnostic accuracy, improve patient outcomes, and streamline healthcare delivery, addressing some of the critical challenges faced by modern healthcare systems (Joshi et al., 2024; Mesko et al., 2020; Muehlematter

et al., 2021).

However, the integration of AI into global health, particularly in LMICs, remains limited. The potential of AI to address some of the key challenges in global health is immense. By automating and enhancing the diagnostic process, AI could significantly reduce the reliance on expert technicians, accelerate diagnosis, and improve the accuracy of results. This could lead to earlier detection and treatment of infectious diseases, ultimately reducing the burden on healthcare systems and improving patient outcomes (Guo & Li, 2018; Schwalbe & Wahl, 2020; USAID, 2019; Wahl et al., 2018).

1.2 Context of the research

This industrial doctoral project is part of a collaboration between the Biomedical Image Technologies research group of the Polytechnic University of Madrid, an international reference in the development of algorithms for the processing of medical images, and Spotlab, a start-up of Internet-of-medical-things recognized by the EU as one of the companies with the most disruptive potential that develops diagnostic solutions using mobile technologies, 3D printing, and artificial intelligence.

We have established collaborations with various national and international health institutions, including:

- The Malaria and Emerging Parasitic Diseases Laboratory at the National Microbiology Centre, Instituto de Salud Carlos III (Spain), which provided parasitic samples, participated in image labeling, validated the technology, and contributed to the creation of AI algorithms for filariasis.
- The Eastern and Southern Africa Center for International Parasite Control (ESACIPAC), Kenya Medical Research Institute (KEMRI) in Nairobi, Kenya, which provided samples of soil-transmitted helminthiasis (STH) and participated in the proof-of-concept study for the STH algorithm.

Since we have established collaboration with these research centers, we have focused on the development and validation of artificial intelligence for helminthiasis and filariasis. However, the ultimate goal is to develop a framework that will allow the design of AI models to assist in the diagnosis of other diseases as well, working in the direction of building a foundation model for all parasitic infections. To this end, and although it falls outside the scope of this thesis, we have also developed AI models for malaria, Chagas disease, and leishmaniasis.

1.3 Objectives

The primary objective of this thesis is to design, develop and validate AI solutions to assist the diagnosis of global health diseases, with a particular focus on neglected tropical diseases.

This overarching goal is supported by the following specific objectives:

1. Creating datasets with microscopy images of helminthiasis and filariasis: To establish a comprehensive workflow for the collection, digitization, and labeling of microscopy images relevant to global health diseases.
2. Mitigating the lack of labeled image datasets for neglected tropical diseases: To address the scarcity of labeled medical images, we explored the use of crowdsourcing and AI with human-in-the-loop and self-supervised learning techniques for data labeling.
3. Designing and implementing an AI development and deployment workflow for microscopy imag analysis: To develop a customizable and systematic workflow for creating AI algorithms, encompassing data preprocessing, AI model training, validation, and deployment, with adaptability for diagnosing new diseases from microscopy imaging.
4. Evaluating AI models and their integration in real clinical workflows in resource-limited settings: To assess the performance of the developed AI models and the feasibility and impact of deploying AI-driven diagnostic tools in LMICs, addressing challenges such as infrastructure limitations, data availability, and user training. Additionally, evaluate the clinical utility and acceptance of the AI technology.

By achieving these objectives, this thesis aims to contribute to the development of AI technologies that can be effectively integrated into healthcare systems, particularly in resource-limited settings, to address critical global health challenges.

1.3.1 Technical Methodology Overview

This thesis focuses on the design and development of deep learning algorithms for NTDs. Over the course of the PhD, our research transitioned from supervised learning to self-supervised learning, reflecting the evolving state of the art in the field.

With the common ground of addressing the specific designs and developments of AI for NTDs, the research began with the training of a classification algorithm using supervised learning on a small dataset of STH images. The focus then shifted to developing object detection models using supervised learning techniques for STH, while also assessing the feasibility of deploying these models on edge devices. After demonstrating technical feasibility, we expanded the technique to detect microfilariae, conducting prospective validation and evaluating human-AI interaction in the process. As the field advanced, self-supervised learning emerged as a powerful alternative, enabling the development of robust models with limited labeled data. This shift prompted the exploration of creating a foundation model for fecal samples, specifically addressing the problem of label scarcity.

Figure 1.2 outlines the methodology overview approached in this PhD thesis for addressing the research gaps in two of the most prevalent NTDs.

1.3.2 AI Development Workflow and Engineering Infrastructure

Building on these advancements, the research in this PhD implemented a structured workflow and infrastructure to streamline the development and deployment of AI models for NTDs. A comprehensive pipeline was established, integrating data preprocessing, model training,

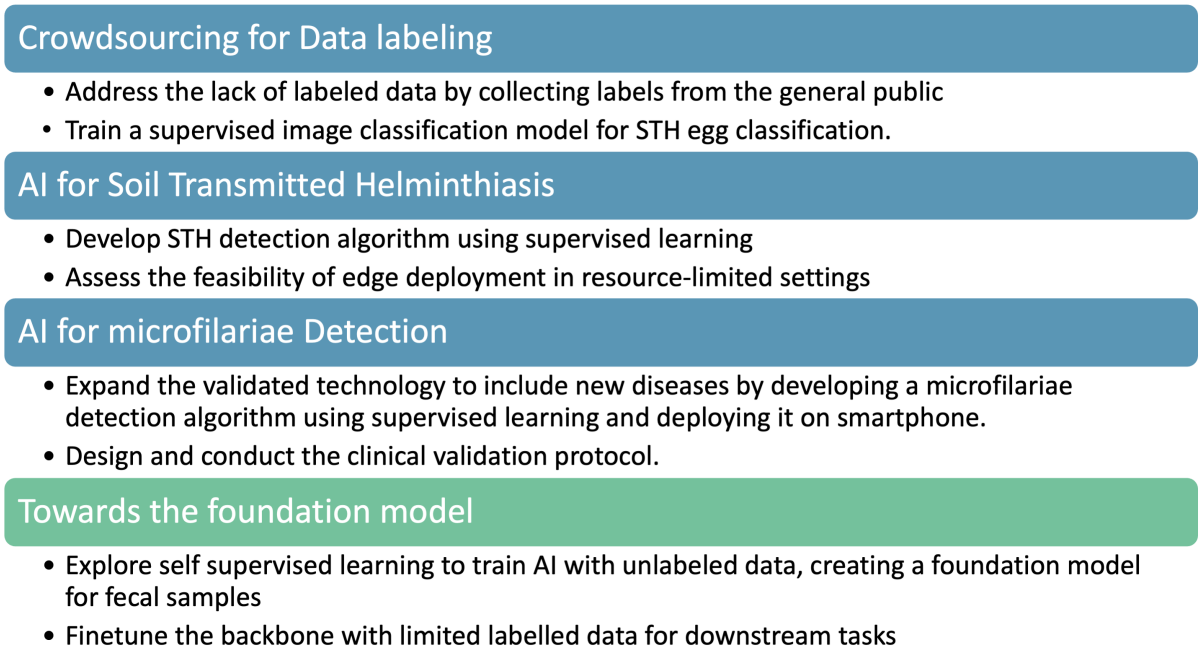


Figure 1.2: Methodology overview of this PhD thesis.Supervised and self-supervised techniques highlighted in blue and green, respectively.

and model deployment. This infrastructure enabled efficient experimentation with both supervised and self-supervised learning approaches, ensuring scalability and adaptability to various NTD-related tasks.

Figure 1.3 illustrates the workflow and infrastructure used throughout this PhD research, following the life cycle of a machine learning project. Including image acquisition,

- Image Acquisition: For this project, images are digitized using Adaptaspot. This system, consists of a 3D adapter and a mobile app, is developed by Spotlab.
- Image Storage, visualization and Labeling: Once digitized, the images were uploaded to the telemedicine platform, Telespot, when an internet connection was available. The platform, hosted on Amazon Web Services (AWS), stores images in AWS S3, with associated metadata saved in MongoDB, a NoSQL database. Through the web interface, users can visualize and label the uploaded images. The platform was also developed by Spotlab.
- Data Preparation. For each project, I downloaded data from S3 and MongoDB, and performed preprocessing on a local computer. After preprocessing, the data was re-uploaded to another S3 bucket and registered in Weights & Biases for dataset versioning.
- AI Model Training and Evaluation: With the preprocessed data, I proceeded to model training. Data was downloaded from S3 for model training and evaluation, which took place on Amazon SageMaker, a cloud-based machine learning platform, and Cesvima, the Supercomputing and Visualization Center of Madrid. The resulting

models were uploaded to S3. Both TensorFlow and PyTorch, two open-source deep learning frameworks, were utilized during this project to develop, train, and deploy AI models, leveraging their unique strengths in research flexibility and production scalability. The training code was uploaded to GitHub, and all experiment-related configurations, including input data, training parameters, outputs, and hardware used, were tracked in Weights & Biases, allowing full traceability and reproducibility.

- Model Deployment: After evaluating candidate models with the validation data, the best-performing model was deployed for prospective validation. For mobile inference, models were exported in TFLite format, optimized for edge devices. Smartphones could directly download the exported model along with the associated configuration file. For cloud inference, a binary file containing the model weights and inference code was generated. The Multi-Model Server framework was utilized, and for each type of task, a single Docker container was created. This container, hosted on AWS Fargate—a serverless compute engine. This serving technique allowed multiple models to be hosted within the same Docker instance. Telespot then made inference requests through API calls.
- Clinical Validation: Once the models are deployed, prospective validation is conducted using a smartphone or the telemedicine platform, with the support of the Spotlab team.

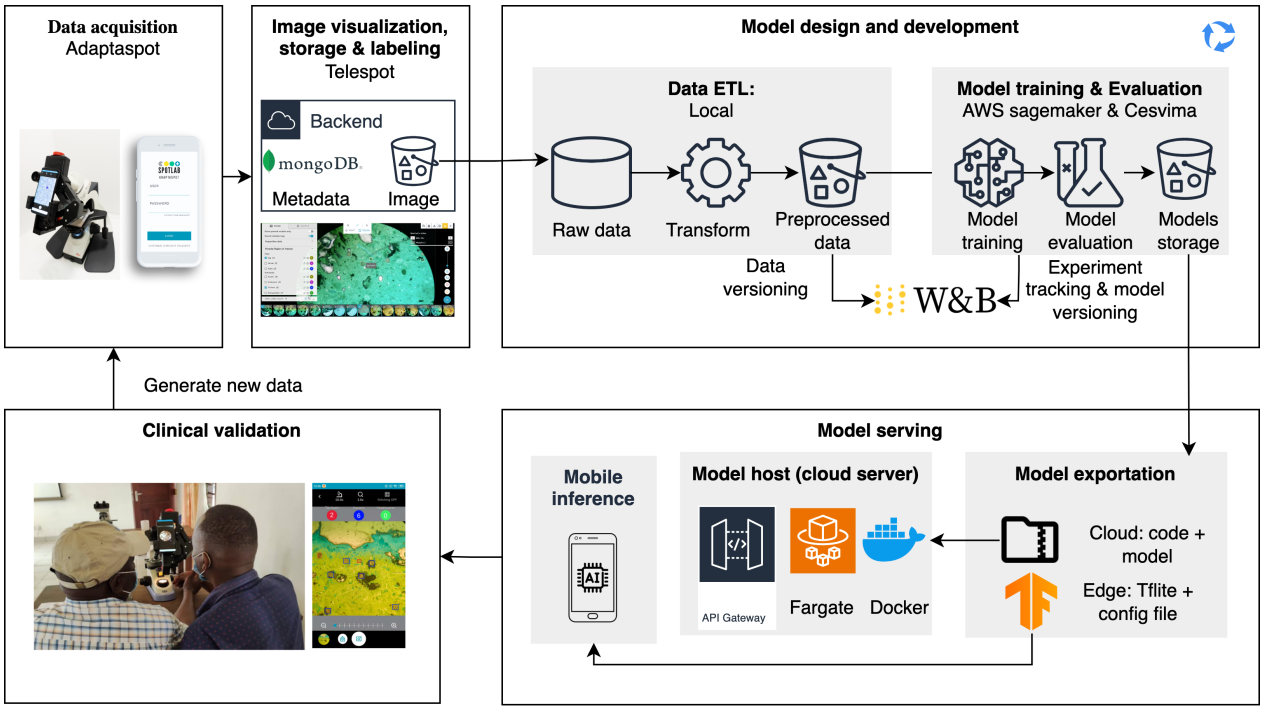


Figure 1.3: The AI development workflow, encompassing data acquisition, image visualization, storage, labeling, model design, and development—including data preprocessing, model training, and evaluation—is an iterative process. The best-performing models are exported for cloud or edge inference, followed by prospective validation to assess their real-world effectiveness.

1.4 Document structure

This thesis is structured as follows:

In Chapter 2, we introduce the clinical context of neglected tropical diseases, detailing their global distribution, prevention strategies, and diagnostic methods, with a specific focus on filariasis and STH. This chapter sets the stage for understanding the significance and challenges of addressing these diseases, particularly in LMICs.

Chapter 3 explore the state of the art of artificial intelligence in medical image analysis, with an emphasis on microscopy imaging. We discuss the workflow from research to deployment, highlighting the importance of creating trustworthy AI systems and the unique challenges faced when implementing AI in LMICs. We also introduced the existing research gaps for the development of AI for NTDs.

In Chapter 4, we explore crowdsourcing as a data labeling method. This chapter examines how leveraging non-expert annotations can help alleviate the burden on specialists and facilitate the development of AI models by providing a scalable solution for generating labeled datasets. As part of this study, we created a game called Spotwarriors, in which players contribute to the diagnosis of real-world samples. Using the annotations generated by players, we trained a deep learning algorithm to classify STH images.

Chapter 5 focuses on the development of an AI-based object detection system for STH egg detection and the assessment of deploying mobile AI models in resource-limited settings. In collaboration with KEMRI, we participated in a deworming campaign in a rural area of Kenya, during which we collected a large dataset of STH samples over two weeks and developed four iterations of AI models. This study allowed us to evaluate the model's performance, the feasibility of deploying mobile AI models in rural areas, and demonstrated the effectiveness of our AI development and deployment workflow.

In Chapter 6, we shift our focus to filariasis. In collaboration with ISCIII, we collected samples and developed an edge AI algorithm for screening and species differentiation of four filariasis species. More importantly, during this study, we designed a protocol to validate AI algorithms in a clinical setting, assessing their usability and effectiveness in real-world applications.

Up until now, we have used supervised learning algorithms. In Chapter 7, we explore the application of self-supervised learning techniques to train AI models with limited labeled data. Using the data collected from Kenya, we investigate how SSL can help overcome the challenges of data scarcity.

Chapter 2

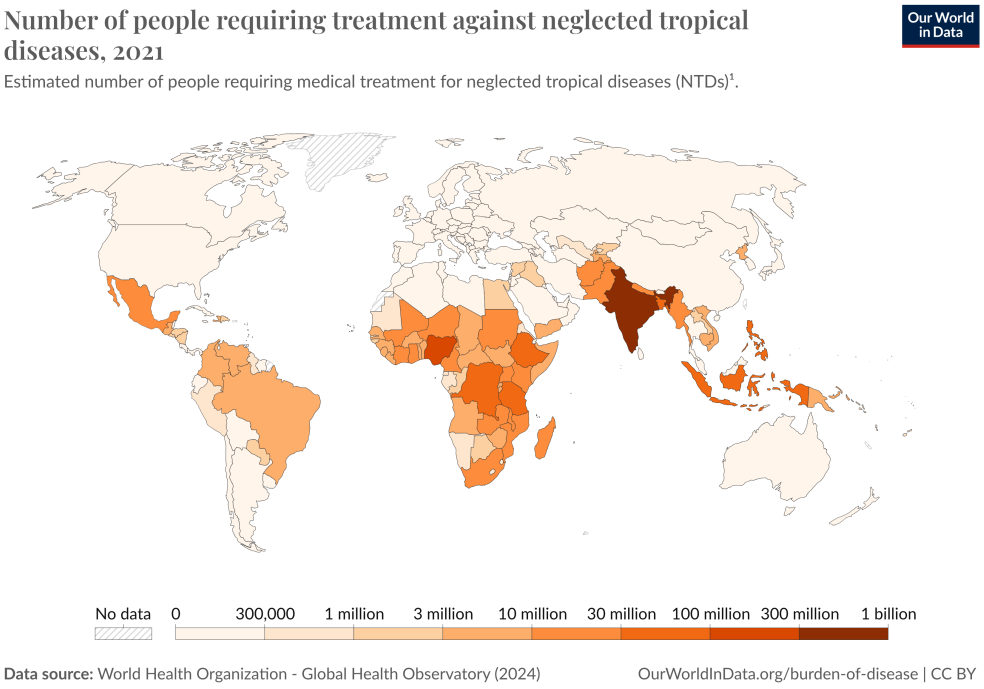
Clinical context

2.1 Neglected tropical diseases

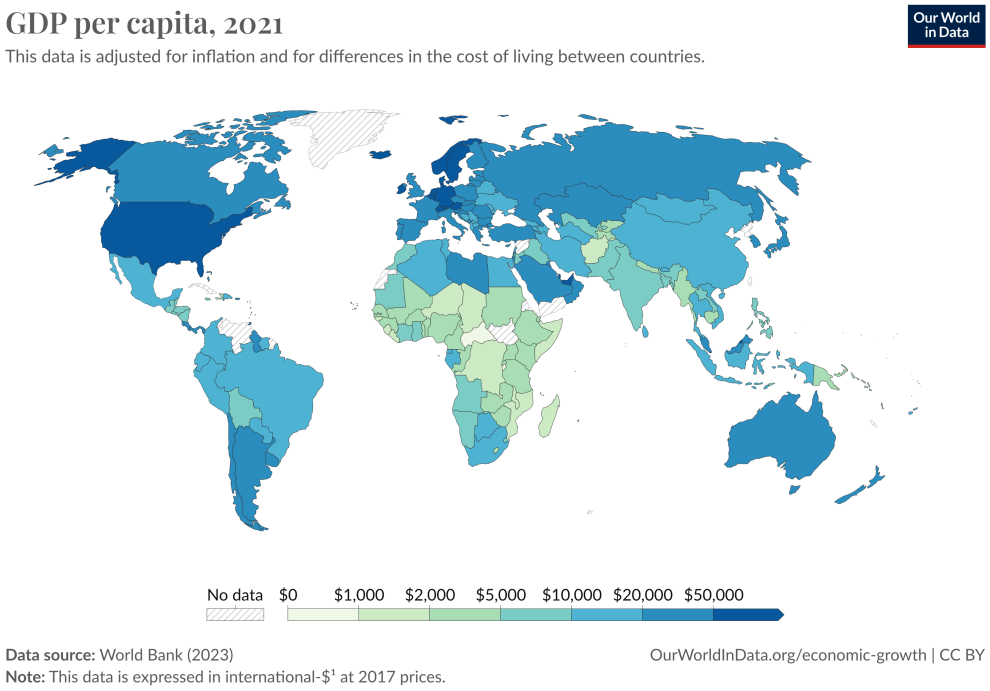
Neglected tropical diseases (NTDs) are a group of infectious diseases that predominantly affect impoverished populations in tropical and subtropical regions (see subfigures 2.1a and 2.1b). Those diseases are caused by a variety of pathogens, including bacteria, protozoa, parasitic worms (helminths), viruses, fungi and toxins, causing devastating health, social and economical consequences. It is estimated that more than 1 billion people are affected by NTDs world wide, and 1.6 billion people require NTD interventions (Feasey et al., 2010; Hotez et al., 2020; WHO, n.d.-b).

The World Health Organization currently classifies 20 diseases as NTDs, including:

1. **Buruli Ulcer**: A chronic, debilitating skin and soft tissue infection caused by *Mycobacterium ulcerans*.
2. **Chagas Disease (American Trypanosomiasis)**: Caused by the protozoan parasite *Trypanosoma cruzi*, transmitted by triatomine bugs.
3. **Dengue and Chikungunya**: Viral diseases transmitted by *Aedes* mosquitoes, causing severe flu-like symptoms and joint pain.
4. **Dracunculiasis (Guinea Worm Disease)**: Caused by the parasitic worm *Dracunculus medinensis*, contracted through drinking contaminated water.
5. **Echinococcosis**: Caused by tapeworms of the genus *Echinococcus*, leading to cysts in the liver, lungs, and other organs.
6. **Foodborne Trematodiases**: Infections caused by trematode worms acquired through consumption of contaminated food.
7. **Human African Trypanosomiasis (Sleeping Sickness)**: Caused by protozoan parasites of the genus *Trypanosoma*, transmitted by tsetse flies.
8. **Leishmaniasis**: Caused by protozoan parasites of the genus *Leishmania*, transmitted by sandflies.



(a)



(b)

Figure 2.1: a) Number of people requiring treatment against neglected tropical diseases. Source: (Roser et al., 2024). b) GDP per capital. Data source: World bank

9. **Leprosy (Hansen's Disease)**: A chronic infectious disease caused by *Mycobacterium leprae*, affecting skin, nerves, and mucous membranes.
10. **Lymphatic Filariasis**: Caused by parasitic worms, leading to severe swelling of limbs and other body parts.
11. **Mycetoma, Chromoblastomycosis, and Other Deep Mycoses**: Chronic, progressive diseases caused by fungi, affecting skin, bone, and tissue.
12. **Onchocerciasis (River Blindness)**: Caused by the parasitic worm *Onchocerca volvulus*, transmitted by blackflies, leading to blindness and skin issues.
13. **Rabies**: A viral disease transmitted through animal bites, leading to severe neurological symptoms and death if untreated.
14. **Scabies and Other Ectoparasites**: Skin infestations caused by mites (scabies) and other ectoparasites.
15. **Schistosomiasis**: Caused by parasitic worms, leading to symptoms like abdominal pain, diarrhea, and liver damage.
16. **Soil-Transmitted Helminthiases**: Caused by parasitic worms such as roundworms, whipworms, and hookworms, leading to malnutrition and anemia.
17. **Snakebite Envenoming**: A significant cause of morbidity and mortality in tropical and subtropical regions.
18. **Taeniasis and Cysticercosis**: Caused by the tapeworm *Taenia solium*, leading to cysts in the brain and other tissues.
19. **Trachoma**: A bacterial infection caused by *Chlamydia trachomatis*, leading to blindness if untreated.
20. **Yaws (Endemic Treponematoses)**: A chronic infection caused by the bacterium *Treponema pallidum pertenue*, affecting skin, bones, and joints.

In order to measure the disease burden, the concept of disability-adjusted life years (DALYs) was developed. DALYs provide a comprehensive measure of overall disease burden, expressed as the number of years of life lost due to premature mortality combined with the years lived with a disability due to prevalent cases of the disease or health conditions in a population. It is estimated that approximately 200,000 deaths and 19 million DALYs are lost annually due to NTDs. These figures highlight the severe impact NTDs have on global health.

In addition to significant mortality and morbidity, NTDs have profound socio-economic impacts. It is estimated that NTDs cost developing communities the equivalent of billions of United States dollars each year. These costs arise from direct health expenses, such as treatment and medical care, as well as indirect costs, including loss of productivity and reduced socioeconomic and educational attainment. Individuals suffering from NTDs often experience prolonged periods of illness that prevent them from working or attending school, thereby limiting their potential to contribute to their communities and economies (Gyapong et al., 2024; WHO et al., 2010). The costs of ongoing medical treatment, coupled with the

loss of income due to illness, can deplete household savings and drive families deeper into poverty.

Moreover, NTDs are responsible for a host of other debilitating consequences. For instance, diseases like lymphatic filariasis can lead to severe swelling and disfigurement, while onchocerciasis can cause blindness. These physical impairments often result in stigmatization, social exclusion, and discrimination.

Despite the large population they are affecting and its socio-economical impact, they are almost absent from global health agenda, the reason that's why it is called neglected (Hotez et al., 2020). Even today, when the focus is on the Universal Health coverage, NTDs have very limited resources and are almost ignored by global funding agencies.

Among the group of NTDs, soil-transmitted helminthiasis and filariasis (including lymphatic filariasis, onchocerciasis, loiasis, and mansonelliosis) are the most prevalent, each affecting more than 1 billion people. As a result, this PhD thesis began by developing AI solutions for these two diseases. A more detailed description is provided in sections 2.1.3 and 2.1.4.

2.1.1 Prevention and control of NTDs

To prevent and control NTDs, the WHO recommends five main strategies: 1) preventive chemotherapy; 2) intensified case-management; 3) vector control; 4) provision of safe water, sanitation, and hygiene; and 5) veterinary public health (WHO, 2020, n.d.-a). Preventive chemotherapy, the primary method for managing lymphatic filariasis, onchocerciasis, schistosomiasis, and soil-transmitted helminthiases, involves the mass distribution of seven broad-spectrum anthelmintic medications. These medicines are effective, safe, and have minimal side effects. Intensified case-management focuses on treating infected individuals and those at risk. This approach includes early diagnosis, treatment to reduce infection and morbidity, and management of complications. It is particularly used for NTDs without available preventive chemotherapy, such as Buruli ulcer, Chagas disease, human African trypanosomiasis, various forms of leishmaniasis, leprosy, and yaws.

Many NTDs are transmitted by vectors: insects spread dengue and other viral diseases, Chagas disease, human African trypanosomiasis, leishmaniasis, lymphatic filariasis, and onchocerciasis. Snails are crucial for transmitting foodborne trematodiasis and schistosomiasis, while crustaceans play a role in dracunculiasis and foodborne paragonimiasis transmission. Controlling these vectors can significantly reduce disease spread. Additionally, improving water quality, sanitation, and hygiene can help eliminate diseases linked to environmental sources like water and soil. Finally, some NTDs are zoonotic, transmitted between animals and humans. People living near infected animals are at risk, so controlling these diseases in livestock is also essential.

2.1.2 Diagnosis of NTDs

Effective diagnostics methods are crucial for the disease mapping, screening, surveillance, monitoring and evaluation (Ngwese et al., 2020; USAID, 2019; WHO et al., 2021a; WHO, 2021). Better diagnostic methods significantly hasten progress toward elimination, by enabling early

detection and appropriate treatment, thereby diminishing the infection source and reducing morbidity. It also reduces and optimizes the cost of the control programmes by reducing the cost of mass drug administration. As the intensity of infection and prevalence of NTDs progressively decrease, current methods may not have the necessary sensitivity or specificity, and new diagnostic methods are required to support decisions including change of frequency of mass drug administration, decision to stop mass drug administration and commence surveillance and validating or verifying elimination or eradication of a disease.

The diagnosis of NTDs involves a variety of approaches, often tailored to the specific disease and the resources available in the affected regions. Common diagnostic methods includes:

Clinical Diagnosis: Clinical diagnosis includes observation of symptoms and physical examination. Many NTDs present with characteristic clinical signs and symptoms. For example, leprosy can be identified by skin lesions and numbness, lymphatic filariasis often presents with swollen limbs, and people with onchocerciasis may have skin nodules and experience blindness.

Microscopy: Examination of blood smears, skin snips, stool and urine samples, and tissue biopsy is essential for some diseases. For example, parasites of lymphatic filariasis and African trypanosomiasis can be observed directly in blood samples. In onchocerciasis, small skin samples are taken and examined for the presence of microfilariae. Schistosomiasis and soil-transmitted helminthiases are often diagnosed by detecting eggs or larvae in stool or urine samples.

Serological Tests: ELISA (Enzyme-Linked Immunosorbent Assay) is based on antigen-antibody interactions. It uses enzymes linked to an antibody or antigen as a marker for the detection of a target molecule. Rapid Diagnostic Tests (RDTs) operate based on capillary action that moves a liquid sample along a solid substrate (such as a test strip) with immobilized antibodies or antigens. There are available RTDs for dengus, chagas, lymphatic filariasis, leishmaniasis, schistosomiasis, african trypanosomiasis, and onchocersiasis.

Molecular Methods: Molecular methods including Polymerase Chain Reaction (PCR) and Loop-Mediated Isothermal Amplification (LAMP) are methods that amplifies DNA or RNA of the pathogens, making it highly sensitive and specific. However, it requires sophisticated equipment, making it less affordable for resource-limited settings.

2.1.3 Soil-transmitted helminth

Soil-transmitted helminth is an infectious disease caused by different parasitic worms, including roundworm (*Ascaris lumbricoides*), whipworm (*Trichuris trichuira*), and hookworms (*Ancylostoma duodenale* and *Necator americanus*). It affects more than 1.5 billion people worldwide, mainly in Southeast Asia and Africa (WHO, 2022). These worms affect health status in multiple ways, such as diarrhoea, abdominal pain, loss of blood and protein, anemia, malnutrition, etc.

Soil-transmitted helminths (STH) are spread through the eggs expelled in the feces of infected individuals. These adult worms reside in the intestines, where they lay eggs. When an infected person defecates outdoors or when their feces are used as fertilizer, the eggs are deposited

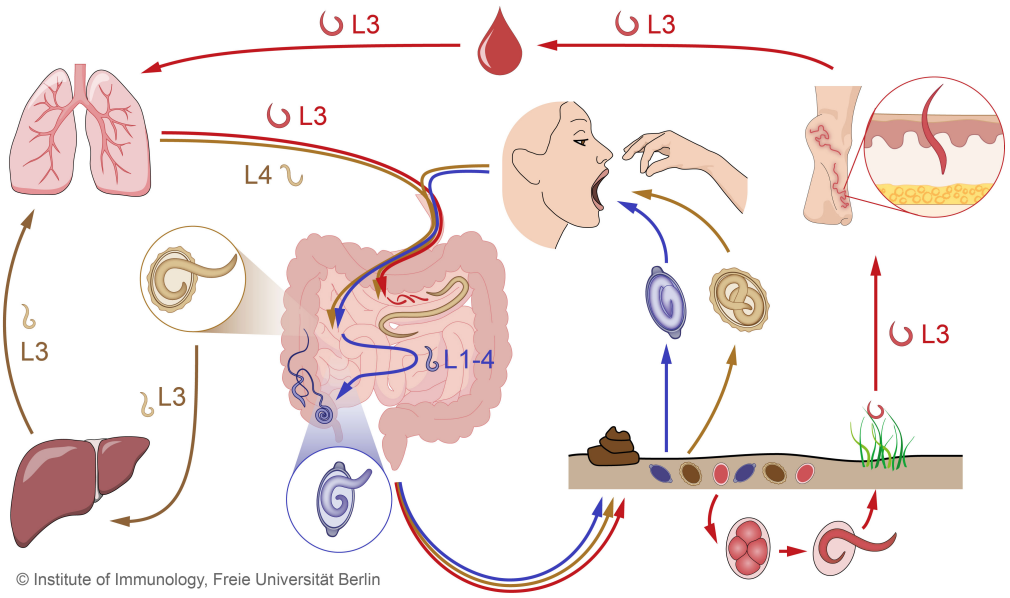


Figure 2.2: Combined STH life-cycle. Source (Schlosser-Brandenburg et al., 2023)

onto the soil. People can become infected with Ascaris and whipworm by ingesting these eggs, often through putting unwashed hands in their mouths or consuming contaminated water, or vegetables and fruits that have not been thoroughly washed, peeled, or cooked. In the case of Ascaris, the eggs hatch in the small intestine, releasing larvae that penetrate the intestinal wall and migrate to the lungs through the bloodstream, then it ascends to respiratory tracts, and are then swallowed, and finally reaching to small intestine. For Trichuris, after ingestion, the eggs hatch in the small intestine, and the larvae migrate to the large intestine, where they embed in the mucosa and mature into adult worms.

In contrast, hookworm transmission operates differently. The eggs themselves are not immediately infective. Instead, they hatch in the soil, releasing larvae. These larvae become infective and can penetrate human skin, typically through the soles of the feet when someone walks barefoot on contaminated soil. The larvae enter the bloodstream, travel to the lungs, ascend the respiratory tract, and are then swallowed, finally reaching the small intestine where they mature into adult worms. This mode of transmission emphasizes the importance of proper sanitation, hygiene practices, and wearing footwear to prevent infections, as detailed in Figure 2.2.

To combat infection, the WHO recommends the administration of preventive chemotherapy using anthelminthic drugs, such as albendazole and mebendazole, to individuals at risk. This includes preschool children, school-age children, women of reproductive age, and adults in high-risk occupations. In 2022, over 897 million individuals received preventive chemotherapy, with distribution primarily concentrated in Southeast Asia and Africa (Figure 2.3). Given the safety of these drugs for healthy individuals, it is deemed more cost-effective to administer treatment to entire populations at risk rather than conducting individual testing and treatment. However, diagnostic tools remain crucial for monitoring the efficacy of deworming programs and informing decision-making strategies. Additionally, in areas nearing elimination, a

test-and-treat strategy is recommended instead of mass drug administration (Turner et al., 2017).

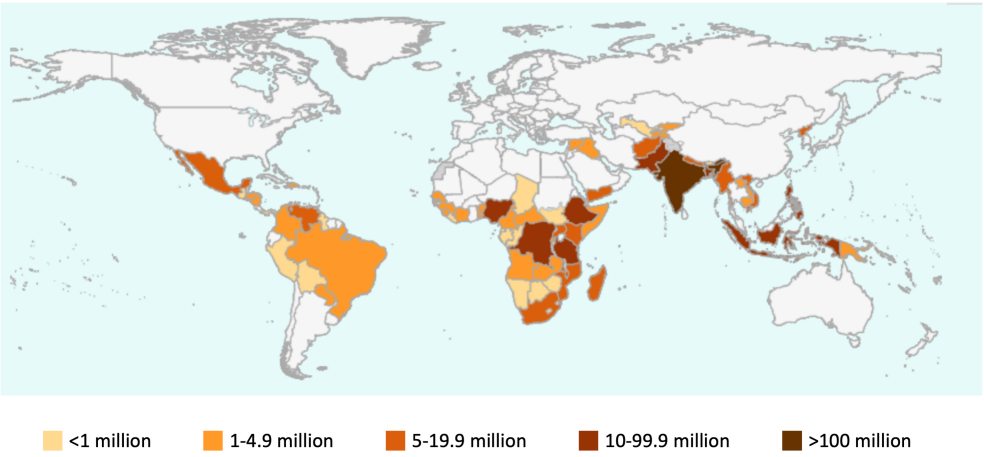


Figure 2.3: Number of children requiring preventive chemotherapy for STH in 2022. Modified from (WHO, n.d.-c)

Various laboratory diagnosis methods exist for soil-transmitted helminth infection. Ngwese et al. provided a comprehensive review and comparison of these diagnostic approaches (Ngwese et al., 2020). These methods can generally be categorized into three groups: 1) methods that detect parasites in fecal samples utilizing microscopy, such as direct examination, Kato-Katz (Katz et al., 1972), formol-ether concentration (FEC) (Ritchie et al., 1948)), McMaster, flotation, translation and centrifugation (FLOTAC) (Utzinger et al., 2008); 2) methods that detect antibodies or antigens in blood samples; and 3) DNA-based techniques, including polymerase chain reaction and loop-mediated isothermal amplification. The selection of a diagnostic technique depends on several factors, including operational cost and infrastructure. DNA-based techniques, for instance, necessitate sophisticated equipment and expensive materials, making them unsuitable for field surveys. Additionally, considerations of sensitivity and specificity are crucial. Depending on whether the objective is to assess prevalence and transmission or monitor the effectiveness of mass drug administration, the need for precise diagnostics varies. For example, while Kato-Katz examines a small sample volume, resulting in lower sensitivity compared to PCR or concentration techniques like FLOTAC, it may suffice for certain surveillance purposes.

Kato-katz Kato-katz is the gold standard method recommended by the WHO and the most widely used diagnostic method (Nikolay et al., 2014; WHO et al., 2021a). This technique is highly valued for its simplicity, cost-effectiveness, and ability to provide quantitative data on parasite load, making it particularly useful in resource-limited settings where intestinal helminth infections are prevalent.

The Kato-Katz method involves the preparation of a stool sample on a microscope slide using a standardized template to ensure a specific quantity of fecal matter is analyzed. For each subject, a fecal sample is collected, and for each sample, 2 slides were prepared. With the assistance of a template, a fix amount of feces is placed on the glass slide. The slide glass is

then covered with cellophane soaked in a glycerol-malachite green solution, which clears the fecal material and stains the helminth eggs, making them easier to identify and count. This preparation is left to sit for a short period, usually 30 to 60 minutes except for hookworm samples to allow the glycerol to clear the fecal debris and enhance the visibility of the eggs under a microscope. Then, the prepared smear should be examined systematically under the microscope, during which eggs are counted. It is recommended to read the prepared sample within 30 minutes after slide preparation, as hookworm eggs decrease with time (Bosch et al., 2021).

In addition to these three common STH eggs, other species could also be found in fecal samples, such as *Strongyloides stercoralis*, *Enterobius vermicularis* or *Schistosoma mansoni*. Figure 2.4 illustrates some common paraiste eggs. Ascaris lumbricoides eggs are typically oval or round with a thick, rough outer shell that is often described as having a mammillated (bumpy) surface. They measure approximately 45-75 micrometers in length and 35-50 micrometers in width. The outer layer is usually brownish-yellow due to bile staining. Trichuris trichiura (whipworm) eggs are barrel-shaped with prominent bipolar plugs at each end, giving them a characteristic lemon like appearance. They measure about 50-55 micrometers in length and 20-25 micrometers in width, with a smooth, thick shell that is yellowish-brown. Hookworm eggs, which include species such as Ancylostoma duodenale and Necator americanus, are oval and measure approximately 60-75 micrometers in length and 35-40 micrometers in width. They have thin, smooth, transparent shells and contain a developing embryo that typically appears as a cluster of cells when freshly passed in feces. These distinct morphological features allow for the differentiation and diagnosis of these helminth infections through fecal microscopy.

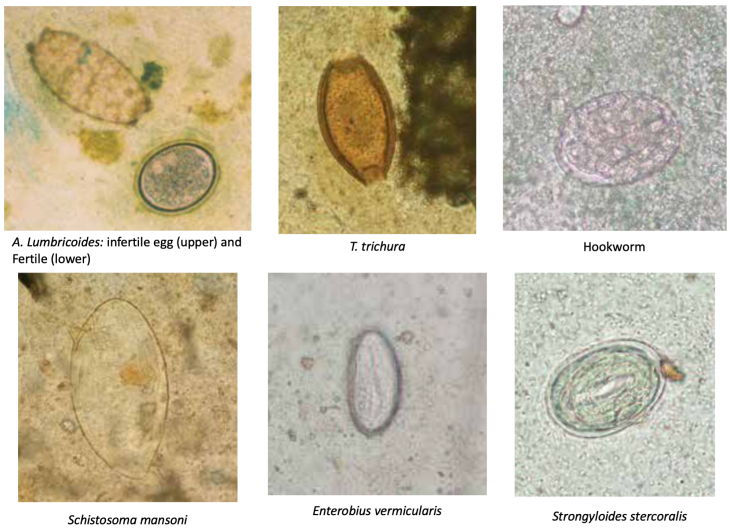


Figure 2.4: Soil-transmitted helminth eggs illustration, adapted from (WHO, 2019)

2.1.4 Filariasis

Filariasis is an tropical infectious disease caused by roundworms of type filarioidea (Phylum Nematoda). It is estimated that more than one billion people require preventive chemotherapy

to stop the spread of this infection. The global burden of filariasis is immense, particularly in tropical and subtropical regions of Africa, Asia, the Americas, and the Pacific, where it contributes to chronic disability, social stigma, and significant economic losses.

There are at least eight filarial worms that hosts in humans, leading to four distinct types of diseases. Lymphatic filariasis is caused by *Wuchereria bancrofti*, *Brugia malayi*, and *Brugia timori*, leading to significant lymphatic dysfunction and swelling, often resulting in elephantiasis. Onchocerciasis, also known as river blindness, is caused by *Onchocerca volvulus* and can result in severe skin and eye damage, potentially leading to blindness. Loiasis, caused by *Loa loa*, is characterized by transient skin swellings and can cause eye discomfort as the worm migrates across the eyeball. Lastly, mansonellosis is caused by *Mansonella perstans*, *Mansonella ozzardi*, and *Mansonella streptocerca*, and although often less symptomatic, it can still lead to skin and systemic symptoms. Sometimes, other filarial parasites like *Dirofilaria immitis* or *Dirofilaria repens*—primarily known for causing infections in animals, particularly dogs—can also affect humans, leading to zoonotic infections. However, in humans, these parasites are generally unable to complete their full life cycle. When they do infect humans, they may cause localized inflammatory reactions or nodules, often in the skin or lungs, but they do not typically result in systemic disease or severe complications as seen with human-specific filarial parasites.

Filariasis species

Lymphatic filariasis, also known as elephantiasis, affects more than 120 million people in 72 counties in tropical and subtropical areas of Asia, Africa, the West Pacific, and parts of the Caribbean and South America, of which 90% are caused by *W. bancrofti*, the remaining 10% are mainly caused by *Brugia malayi* and *Brugia timori* has a low prevalence. People infected by these worms may be asymptomatic, have acute episodes or develop chronical conditions. These worms damage the lymphatic system and alter the immune system. Acute episodes are charecterized by fever, lymphatic inflammation, edema, and dermatolymphangioadenitis. In chronic conditions, it leads to lymphoedema of limbs and hydrocele (accumulation of fluid in the testes cavity). With the aim of eliminating lymphatic filariasis as a public health problem, a preventive chemotherapy (PC) strategy is implemented in endemic areas. Massive drug administration (MDA) of albendazole, ivermectine, diethylcarbamazine (DEC), or a combination of them can eliminate microfilariae on the blood, interrupting the transmission of infection by mosquito.

Onchocerciasis, commonly known as river blindness, is caused *Onchocerca volvulus*. It occurs mainly in sub-Saharan Africa (99%), but also in the Yanimami area of Brazil, Venezuela and Yemen. Infection causes damage to the skin and eye, which can lead to permanent blindness. In 2017, more than 20.9 million people are estimated to be affected by onchocerciasis and at least 220 million people require PC (James et al., 2018). Onchocerciasis is diagnosed by examination of the skin snip under the microscope, serology test, and PCR. To interrupt the transmission of onchocerciasis by 2030, all of the population at risk is treated with ivermectin.

Loiasis, caused by *Loa loa*—also known as the African eye worm—is primarily distributed in West and Central Africa. The infection is characterized by itchy swellings of the body (known as Calabar swellings) and ocular manifestations, including conjunctival granulomas

and eyelid edema, though it does not typically cause blindness. Until recently, *Loa loa* was commonly considered a relatively harmless nematode and was not included on the list of neglected tropical diseases that require control or eradication efforts (Metzger & Mordmüller, 2014; Ramharter et al., 2024). Loasis receives attention mainly because it complicates the elimination of lymphatic filariasis and onchocerciasis, as the use of ivermectin in individuals with high levels of *L. loa* in their blood can lead to severe adverse events (SAEs) (Gardon et al., 1997). However, loiasis is not harmless, it leads to disability adjusted life-years similar to other neglected tropical diseases. Epidemiological studies have revealed that approximately 14.4 million people live in high-risk areas, and 15.2 million people live in intermediate-risk areas (Zouré et al., 2011). Moreover, it is projected that by 2025, around 169 million people will reside in *L. loa* endemic regions (Vinkeles Melchers et al., 2020). Hence, loasis should be treated and eradicated. *L. loa* can be diagnosed in several ways: identification of adult worms under the skin or eyes, identification of microfilariae on a blood smear using a microscope, antibody test or PCR. The treatment of loiasis includes the use of diethylcarbamazine and albendazole to kill microfilariae. In addition to pharmacological treatment, surgical extraction of the adult worm, particularly when it appears in the eye or just beneath the skin, is sometimes necessary to alleviate symptoms and prevent further complications

Mansonellosis is caused by three filarial parasite species from the genus *Mansonella*: *M. perstans*, *M. streptocerca*, and *M. ozzardi*. *M. perstans* is endemic in West, East, and Central Africa, as well as Central and South America; *M. ozzardi* is endemic in Central America, South America, and several Caribbean islands; and *M. streptocerca* is endemic in tropical regions of West and Central Africa (CDC-DPDx, n.d.-b). Most patients infected with *Mansonella* are asymptomatic or exhibit mild symptoms such as urticaria, angioedema, pruritus, or ocular lesions. Studies have reported that *M. perstans* is the most common filariasis in Africa. Despite affecting more than 100 million people globally, with 600 million living in 33 high-risk countries (Simonsen et al., 2011), it remains one of the most neglected filarial diseases (Lima et al., 2016), (Raccurt et al., 2014), (Mediannikov & Ranque, 2018; Ta-Tang et al., 2018), and there are no control programs currently in place. Mansonellosis can be diagnosed by microscopical examination of blood smears (*M. perstans*, *M. ozzardi*) or skin snips (*M. streptocerca*).

Figure 2.5: Pathology caused by different fialariasis. a) elephansiasis, caused by lymophatic filariasis; 6) blindness caused by onchocersiasis; c) worms in eyes, caused by loasis.

Transmission of filariasis

Filariasis is transmitted through the bite of infected blood-feeding insects, primarily mosquitoes and flies. As illustrated in Figure 2.6, when an infected fly or mosquito bites a human, the vector injects filarial larvae into the human's bloodstream. These parasites migrate to specific tissues—such as lymphatic tissues, subcutaneous tissues, or serous cavities—depending on the species, where they mature into adult worms. The adult worms then mate and produce microfilariae, which circulate in the host's blood or migrate to the skin, again depending on the species. When another blood-feeding insect bites an infected person, it ingests the microfilariae. These microfilariae then develop within the insect, and the cycle continues as the insect infects a new host during its next blood meal.

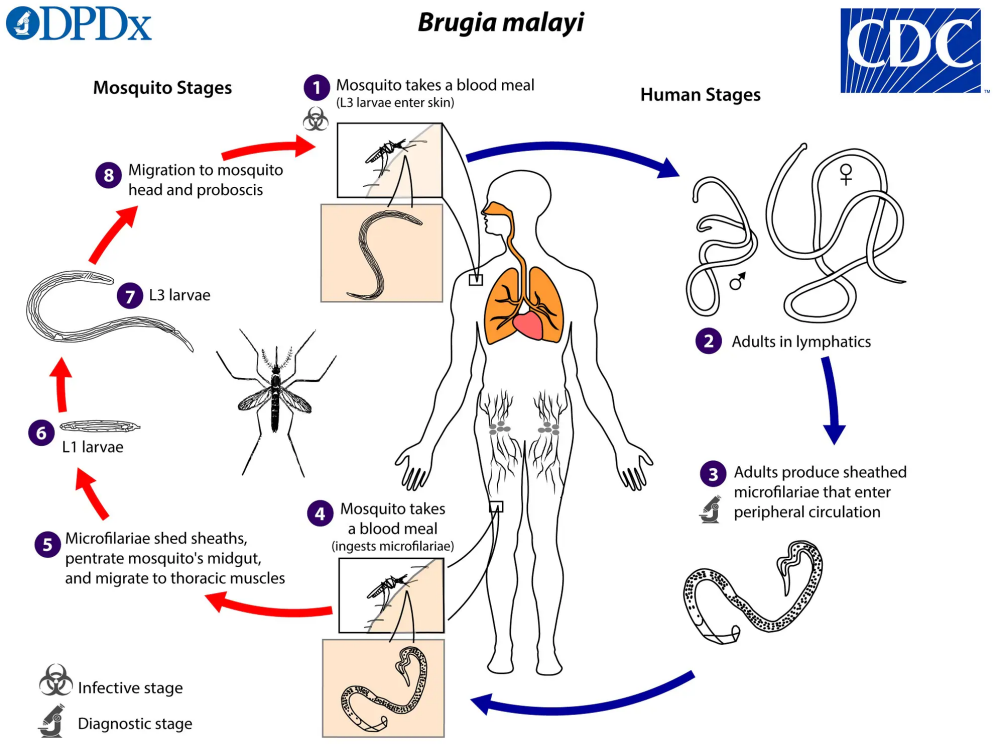


Figure 2.6: Lymphatic filariasis lifecycle. Source: (CDC-DPDx, n.d.-a)

Diagnosis of filariasis

The diagnosis of filariasis involves several approaches, depending on the specific type of filarial infection. The most common method is the microscopic examination of blood samples or skin snip to detect microfilariae. Serological tests, such as ELISA and rapid immunochromatographic tests, are also used to detect filarial antigens or antibodies, offering a more sensitive diagnostic option, especially in cases with low microfilarial loads. In addition, advanced molecular techniques like PCR can be employed to detect filarial DNA in blood or tissue samples, providing highly specific and sensitive identification of the parasite. Ultrasound imaging is sometimes used to visualize adult worms in the lymphatic system, particularly in lymphatic filariasis.

The choice of diagnostic method depends on the type of filariasis, the stage of infection, and the resources available. Due to the fact that these diseases primarily affect resource-constrained settings, the ideal diagnostic test should be one that can be used at the point of care with minimal infrastructure. It should be rapid, highly portable, and easy to administer, requiring little to no specialized training or equipment. Additionally, the test should provide accurate results quickly, enabling immediate diagnosis and treatment to help reduce the spread and impact of the disease in these vulnerable populations(WHO, 2021).

For some diseases, rapid tests are available, such as the Alere Filariasis Test Strip, which detects antigens of *W bancrofti*, the Brugia Rapid point-of-care cassette test, which detects antibodies against *Brugia spp.*, and the OV-16 Rapid Diagnostic Test, which detects antibodies against *O volvulus*. However, for diseases like loasis and mansonellosis, there are no available rapid tests. A significant limitation of these rapid tests is that they do not detect co-infections. In areas where multiple filarial species are endemic, particularly when *L loa* is present, it is crucial to differentiate *L. loa* from other species due to potential treatment complications. This differentiation can be effectively achieved with microscopy, as the microfilariae of different species have distinct morphological characteristics that allow for accurate identification and differentiation.

Diagnosis with microscopy Microscopical examination is a fundamental diagnostic technique for detecting filarial infections, particularly through the identification of microfilariae in blood or skin samples. Most filarial worms are found in blood samples, except *O. volvulus* and *Mansonella streptocerca*, which are diagnosed through skin snips. A thick blood smear is prepared and stained, usually with Giemsa stain, to visualize the microfilariae under a microscope. The timing of blood collection is crucial due to the periodicity of certain microfilariae species: *W. bancrofti* and *Brugia spp.* exhibit nocturnal periodicity, meaning blood samples are typically taken at night when the parasites are present in peripheral circulation, while *L. loa* is diurnal. *O. volvulus* and *Mansonella spp.* do not exhibit periodicity. In cases of onchocerciasis, skin snips are taken from the patient and immersed in saline to encourage the microfilariae to emerge, after which the sample is examined microscopically. The identification of microfilariae is based on their morphology, including size, shape, the presence or absence of a sheath, and the arrangement of nuclei, which can help differentiate between species. Figure 2.7 illustrates some key features of the microfilariaes. It is recommended that the entire blood film is screened with a 10x objective before reporting as negative, and when a filaria worm is found, use a higher magnification to identify the specie.

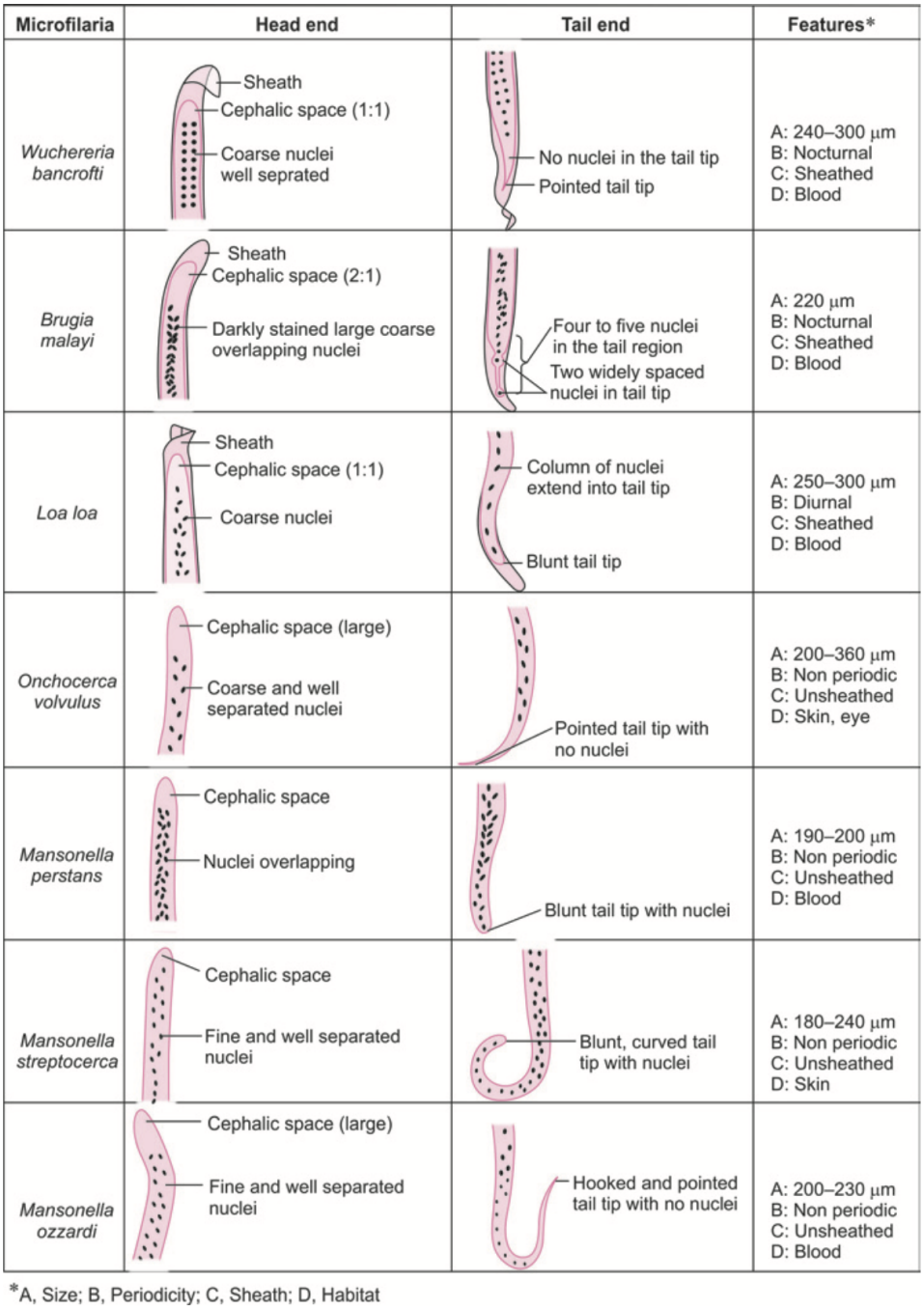


Figure 2.7: Key features of microfilariaes. Source (Sastry & Bhat, 2018)

Chapter 3

State of the art

3.1 Artificial intelligence in medical image analysis

Artificial intelligence (AI) has revolutionized various fields, and its impact on medical image analysis is particularly transformative (Rajpurkar et al., 2022). Medical imaging plays a crucial role in modern healthcare, enabling the visualization of internal structures and functions, aiding in diagnosis, treatment planning, and monitoring of diseases. However, the interpretation of medical images is often complex, time-consuming, and prone to human error. AI, particularly through techniques like machine learning and deep learning, offers powerful tools to enhance the accuracy, efficiency, and accessibility of medical image analysis.

AI algorithms can be trained to recognize patterns and features in medical images that may be subtle or difficult for human eyes to discern. These algorithms are capable of processing vast amounts of data quickly, providing consistent and reproducible results. From detecting tumors in radiological scans to analyzing retinal images for signs of diabetic retinopathy, AI-driven systems are increasingly being integrated into clinical practice to assist radiologists, pathologists, and other healthcare professionals.

Moreover, AI's ability to learn from large datasets allows for continuous improvement in diagnostic accuracy. As more annotated images are fed into these systems, their predictive performance can surpass traditional methods. AI also holds promise in expanding access to high-quality diagnostics in underserved regions, where specialist expertise may be limited.

The integration of AI into medical image analysis is not without challenges. Issues related to data privacy, the need for large and diverse training datasets, and the interpretability of AI models must be carefully addressed. Nonetheless, the potential benefits of AI in improving patient outcomes, reducing diagnostic errors, and streamlining healthcare workflows make it a rapidly evolving and exciting area of research and development in medical imaging.

3.2 Diagnosis of microscopical images

Microscopical images, whether from histology, cytology, blood smears, bone marrow samples or other biological samples, are fundamental in identifying and understanding cellular, parasites and tissue-level changes associated with diseases. The conventional workflow for microscopical image analysis is schematically depicted in Figure 3.1. Initially, the sample is collected and a smear is prepared by placing a small amount of the sample onto a glass slide. This smear is then stained to enhance the visibility of specific structures within the sample, making it easier to identify and differentiate the objects of interest, such as cells or parasites. After the preparation, the slide is systematically examined under a microscope by an expert. During this examination, the objects of interest are counted manually using a counter. Finally, a detailed report is generated based on the observations, summarizing the findings for each preparation. However, the traditional process is labor-intensive, with high inter-observer variability and dependent on the expertise of trained specialists, whose availability can be limited. Moreover, challenges such as the shortage of trained personnel, insufficient monitoring of laboratory testing, lack of essential equipment and supplies (such as microscopes), and difficulties in sample transportation pose barriers to effective healthcare (Petti et al., 2006).

AI-assisted diagnosis presents a promising solution to address some of these challenges, by automating the identification, quantification and classification of pathogens or cells. Recent publications described their utility in histopathology (Niazi et al., 2019; Vorontsov et al., 2024; Xu et al., 2024), haematology (El Alaoui et al., 2022; Fan et al., 2023; Rösler et al., 2023), and infectious diseases (Maturana et al., 2022; P. Ward et al., 2022).

Figure 3.1: Current workflow for microscopical image analysis

Correct diagnosis of parasite infection, including detecting the parasite, identify their specie and compute their parasitemia is crucial to provide effective treatments. In low-resource settings, manual sample analysis with bright-field microscopy is the most important diagnostic tool for parasitic disease. The parasite counting can benefit from AI by the automatizing the task and increase the sensitivity by analysing more samples (de Korne et al., 2023).

Malaria and NTDs primarily affect low-resource and middle-income countries (LMICs), where laboratory infrastructure is often scarce (Nkengasong et al., 2018; Vasiman et al., 2019). To address this challenge, several research groups have developed portable microscopy solutions using mobile phones, by coupling smartphones to light microscopes to convert them into digital microscopes, as well as low-cost scanners (Armstrong et al., 2022; Bogoch et al., 2013;

Coulibaly et al., 2016; Coulibaly et al., 2023; D'Ambrosio et al., 2015; Ephraim et al., 2015; Lundin et al., 2024; Masud et al., 2020; Maturana et al., 2023; Meulah et al., 2022; Meulah et al., 2024; Oyibo et al., 2022; Pattanaik et al., 2022; Pfeil et al., 2022; Shrestha et al., 2020; A. Yang et al., 2019; F. Yang et al., 2020; Yu et al., 2020). Compared to conventional light microscopes, some of these systems also have the capability to digitize images, connect to telemedicine platforms, and integrate AI algorithms, enhancing diagnostic accuracy and accessibility.

3.2.1 Early approaches and traditional image processing techniques

Initial efforts in automatic microscopic image analysis predominantly utilized traditional image processing techniques. These methods involved feature extraction based on the morphology and color characteristics of stained blood smears, followed by classification using algorithms such as k-nearest neighbors (KNN), support vector machines (SVM), and decision trees. For instance, these techniques have been employed for the detection of malaria (F. Yang et al., 2020; Yu et al., 2020), helminths (Inácio et al., 2020; Jiménez et al., 2016; Jiménez et al., 2020; Y. S. Yang et al., 2001), trypanosoma (Soberanis-Mukul et al., 2013; Uc-Cetina et al., 2013), leishmania (Isaza-Jaimes et al., 2020; Zare et al., 2022), and filaria (D'Ambrosio et al., 2015), achieving reasonable accuracy under controlled laboratory conditions. These early methods effectively leveraged predefined features to classify images and provide diagnostic insights. However, their performance was often constrained by their reliance on handcrafted features, which were limited in their ability to generalize across varying staining techniques, different microscope settings, and diverse sample preparations. This lack of adaptability and the need for manual feature engineering led to challenges in achieving consistent accuracy and robustness in real-world applications where image conditions and staining protocols could vary significantly.

3.2.2 Deep learning techniques

Convolutional neural networks

The advent of deep learning has revolutionized image analysis, particularly in computer vision tasks. Convolutional Neural Networks (CNNs) have become the cornerstone of modern approaches, starting with the groundbreaking performance of AlexNet in the ILSVRC-2012 competition, where it achieved a top-5 error rate of 16.4%, significantly outperforming the state-of-the-art Fisher Vector method, which had a top-5 error rate of 25.7% (Krizhevsky et al., 2012). Since then, CNN architectures have evolved significantly. In 2014, Simonyan and Zisserman (Simonyan & Zisserman, 2014) introduced the VGG network, notable for its use of smaller 3x3 filters and deeper architectures (up to 19 layers). This design choice allows the network to have the same effective receptive field as larger 7x7 filters but with more non-linearities and fewer parameters. VGG reduced the top-5 error rate on ImageNet to 7.3%. The same year, Szegedy's GoogleNet (Szegedy et al., 2015) introduced inception modules that apply parallel convolutions with different filter sizes (1x1, 3x3, and 5x5) and utilize bottleneck layers with 1x1 convolutions to compress the depth of feature maps. Despite having 22 layers, GoogleNet had 27 times fewer parameters than VGG-16 and achieved a top-5 error

rate of 6.7%. Deeper networks often face optimization challenges, which He addressed with ResNet, introducing identity connections to mitigate vanishing gradient problems. ResNet's 152 layers achieved an impressive top-5 error rate of 3.6%, surpassing the 5.1% error rate of human classifiers on the same dataset (Russakovsky, Deng, et al., 2015). Following this, research shifted towards developing more efficient networks suitable for mobile devices, such as MobileNet (Howard et al., 2017; Sandler et al., 2018), which employs depthwise separable convolutions to reduce computational costs.

In contrast to image classification, which assigns a single label to an entire image, object detection provides more granular insights, such as counting and localizing multiple objects within a single image. This is particularly important for applications like parasite counting, where accurate object detection is required. Traditional methods often involve extracting patches containing single elements using techniques such as sliding windows or image segmentation, followed by classification of these patches. This approach is illustrated in Figure 3.2. Multiple studies have utilized this method (Abdelmula et al., 2023; Elvana & Suryanto, 2022; Jung et al., 2021; Mehanian et al., 2018; Pereira et al., 2020; Quinn et al., 2016; Rajaraman, Antani, et al., 2018; Rajaraman et al., 2019; Rajaraman, Silamut, et al., 2018; Sadeghi et al., 2024). For instance, Quinn et al. used sliding windows to crop patches from thick blood smears and employed custom CNNs to classify malaria parasites, achieving an average precision of 0.97 (Quinn et al., 2016). Rajaraman et al. used the Laplacian of Gaussian to detect the centroids of red and white blood cells and subsequently classified them using various CNN architectures, including AlexNet, VGG-16, ResNet-50, Xception, DenseNet-121, and a customized CNN. Their approach achieved a sensitivity of 0.981 and a specificity of 0.992, outperforming previous machine learning techniques in terms of specificity and sensitivity, despite a smaller dataset (D. K. Das et al., 2013; Rajaraman, Antani, et al., 2018). Similarly, Sadeghi et al. developed LeishFuNet, which achieved an overall sensitivity of 0.98 and a specificity of 1 (Sadeghi et al., 2024).

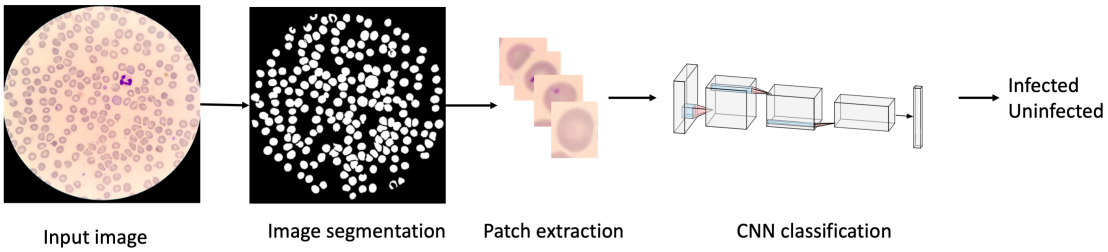


Figure 3.2: Example image classification pipeline. First, object of interest is segmented, and then it is cropped into small patches, which are passed through the CNN, and the CNN predicts the corresponding class

Object detection is designed identify and locate multiple objects within an image, As illustrated on figure 3.3, the entire image is passed to the object detector instead of patches, and the detector outputs both class labels and bounding boxes. Modern object detection algorithms can be grouped into 2 classes: two-stage detection, which separates the region proposal and detection, and one-stage detection, which skips the region proposal stage and directly runs detection over the feature map generated by CNN.

Two-stage detection algorithms first generate region proposals and then classify these proposals and refine their bounding boxes. This approach typically involves a region proposal network (RPN) that suggests potential regions of interest, followed by a second stage that performs object classification and bounding box regression on these proposals. Examples of two-stage detectors include the R-CNN (Region-based CNN) series and its successors like Fast R-CNN and Faster R-CNN (Girshick, 2015; Girshick et al., 2014; Ren et al., 2015).

In contrast, one-stage detection algorithms bypass the region proposal stage and directly perform object detection on the entire feature map produced by the CNN. This approach simplifies the detection pipeline and is generally faster, making it suitable for real-time applications. Examples of one-stage detectors include the YOLO (You Only Look Once) series and SSD (Single Shot MultiBox Detector) and RetinaNet (Ao Wang, 2024; Jocher et al., 2023; Lin, Goyal, et al., 2017; Liu et al., 2016; Redmon et al., 2016; Redmon & Farhadi, 2017, 2018; C.-Y. Wang et al., 2024).

Both approaches have their advantages: two-stage detectors often provide higher accuracy by leveraging the region proposal stage, while one-stage detectors are more efficient and faster, making them ideal for applications requiring real-time processing.

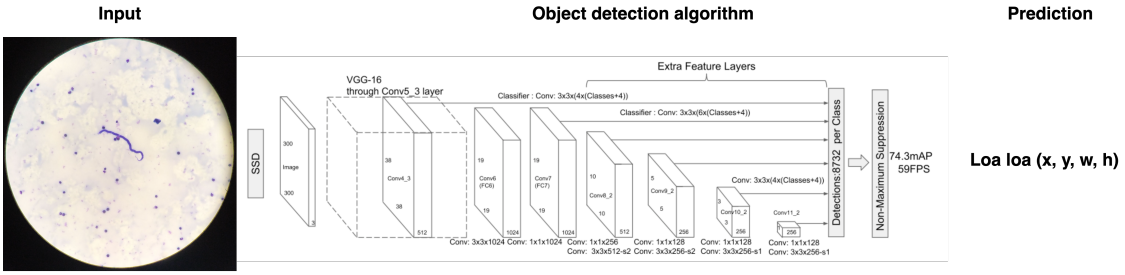


Figure 3.3: Object detection pipeline. The whole image is passed to the CNN and both location and class is predicted

Taking advantage of the strengths of modern object detection models, several researchers have proposed using detection algorithms for parasite detection across various medical applications (Holmström et al., 2017; Lundin et al., 2024; Naing et al., 2022; A. Yang et al., 2019; Y. S. Yang et al., 2001). For instance, Li et al. developed a custom object detector using ResNet152 as the backbone, trained on a dataset comprising 1,122 patient samples. This model achieved a mean average precision (mAP) of 92.16% and an average recall (AR) of 93.56% for detecting parasites in fecal samples (Li et al., 2020). Similarly, Ward et al. employed the R-FCN with ResNet101 as the backbone to detect soil-transmitted helminth eggs, resulting in a weighted average precision of 94.9% and an average recall of 96.1% (P. Ward et al., 2022). Maturana et al. explored multiple object detection architectures, including YOLOv5x, Faster R-CNN, SSD, and RetinaNet, for the detection of malaria parasites in blood smears. Their models achieved an overall precision of 92.10%, recall of 93.50%, F-score of 92.79%, and mAP0.5 of 94.40% for the detection of leukocytes, early-stage, and mature Plasmodium trophozoites (Maturana et al., 2023).

Vision transformers

More recently, Vision Transformers (ViTs) have been gaining significant attention in the field of computer vision. The transformer architecture, originally developed for natural language processing (NLP), has been adapted to handle visual data. In 2020, Dosovitskiy et al. introduced the Vision Transformer (ViT), which marked a departure from traditional convolutional neural networks (CNNs). Instead of using convolutions, ViTs divide an image into smaller 16x16 patches, generate embeddings for each patch, and feed these embeddings into a transformer model. Dosovitskiy and colleagues trained several versions of ViTs, including ViT-Base with 86 million parameters, ViT-Large with 307 million parameters, and ViT-Huge with 632 million parameters. These models were pre-trained on the JFT-300M dataset, which contains 300 million images. Remarkably, ViTs outperformed ResNet-based baselines, demonstrating that transformers could achieve superior results in image classification with substantially fewer computational resources required for pre-training (Dosovitskiy et al., 2020). One of the key advantages of Vision Transformers is their ability to capture long-range dependencies within an image, allowing them to model global context more effectively than CNNs. This is particularly beneficial for tasks that require an understanding of the overall scene, such as image classification and segmentation.

Despite the impressive progress in object detection over recent years, there are still notable limitations, particularly when dealing with small objects. Small objects occupy fewer pixels within an image, which restricts the amount of information available for distinguishing them, making accurate detection challenging. This issue is further exacerbated by class imbalance, a common problem in object detection tasks, and is particularly pronounced in microscopical images with low parasitemia. For instance, as shown in Figure 3.4, the parasite of *Leishmania* occupies a very small portion of the entire field of view, covering less than 5% of the area. Similarly, in malaria detection, the number of uninfected cells (green) significantly exceeds the number of infected cells (red) within the image, further complicating accurate detection and classification.

To address these challenges and achieve fine-grained classification, various strategies have been developed. These include multi-scale feature learning using feature pyramid networks, which improve the detection of objects at different scales by combining low-resolution, semantically strong features with high-resolution, semantically weak features (Lin, Dollár, et al., 2017). Additionally, specific data augmentation techniques have been employed to artificially balance the dataset, thereby enhancing the model's ability to detect underrepresented classes (Kisantal et al., 2019; Tong et al., 2020). Some approaches also involve separating the detection process from the classification task, allowing for more specialized models to handle each step, leading to improved accuracy and robustness (Arshad et al., 2022; Davidson et al., 2021; Klarmann-Schulz et al., 2020; Manku et al., 2020).

Self supervised learning

Supervised learning is the most widely used technique in image analysis, particularly for tasks where labeled data is available. In this approach, a model is trained on a labeled dataset, where the input data is paired with corresponding labels that guide the learning process. Through this, the model learns to recognize patterns and features in the data associated with

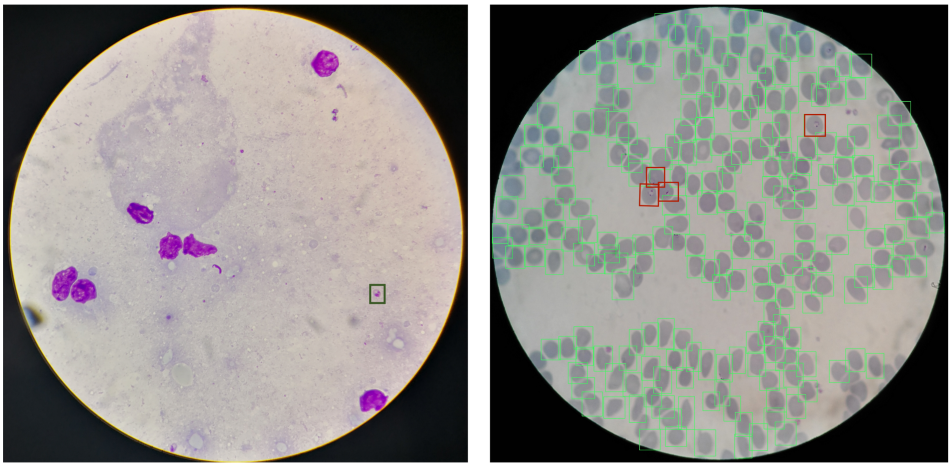


Figure 3.4: Class imbalance problem in microscopical images. a) Foreground-background imbalance. b) class imbalance

specific labels, enabling it to make accurate predictions on new, unseen data.

However, the reliance on large amounts of labeled data presents significant challenges, especially in fields like healthcare, where obtaining expert-annotated data is both expensive and time-consuming. To address these challenges, new learning paradigms such as self-supervised learning, and foundational models have emerged, showing significant potential across various domains, including medicine.

Self-supervised learning, for instance, enables models to learn useful representations from unlabeled data by a by leveraging inherent structures within the data, such as similarities or contrasts. This reduces the dependence on large labeled datasets, which is particularly advantageous in medical imaging, where labeled data is often scarce. (Caron et al., 2021; T. Chen, Kornblith, Norouzi, et al., 2020; T. Chen, Kornblith, Swersky, et al., 2020; X. Chen & He, 2021; Grill et al., 2020; Oquab et al., 2023).

Foundational models, pre-trained on massive datasets across diverse domains, have shown the ability to be fine-tuned for specific tasks with minimal labeled data, making them particularly valuable in medical imaging. These models are praised for their generalizability and robustness, which makes them increasingly useful in assisting with complex medical image analysis tasks (Bommasani et al., 2021; Moor et al., 2023; Rajpurkar & Lungren, 2023). For instance, in 2022, Sellergren et al. introduced CXR Foundation, a foundational model designed to help researchers create customized chest radiography models (Sellergren et al., 2022). Vorontsov et al. developed VIRCHOW, a foundation model trained with 1.5 million whole slide images aimed at clinical-grade computational pathology and rare cancer detection (Vorontsov et al., 2023; Vorontsov et al., 2024). Additionally, Ma et al. proposed MedSAM, a novel foundation model for the segmentation of medical images, further demonstrating the potential of foundational models in healthcare applications (J. Ma et al., 2024).

3.3 AI: from research to clinical deployment

The journey from research to clinical deployment in medical technologies, particularly those involving advanced methodologies like AI, is a complex and multifaceted process. It begins with the development of innovative techniques and algorithms in research settings, where the focus is on achieving proof of concept and demonstrating the potential benefits of these new tools. However, translating these research innovations into practical, reliable clinical applications requires more than just technical proficiency; it involves rigorous validation, adherence to regulatory standards, and the consideration of real-world clinical needs and constraints.

The development and integration of the AI algorithm can be divided into the following stages: definition of the intended use, data collection, data labelling, algorithm development and retrospective validation, prospective validation, and regulation. Once the algorithm is regulated, it can be integrated and used in the clinical workflow.

3.3.1 Definition of intended use

A clear and concise definition of intended use is the first step of the AI development. Instead of focusing on field where public datasets and annotations are available, we have to shift the focus to the clinical relevance and it's impact. The intended use of a medical device refers to the specific purpose for which the product is designed and developed, including the target population, the clinical condition, type of user, clinical environment and operating context.

Healthcare systems in LMICs often face unique challenges, including limited access to advanced diagnostic tools and internet, a shortage of trained healthcare professionals, and variations in disease prevalence compared to high-income settings. As such, the intended use of AI in these regions must consider these factors to ensure the technology is both relevant and feasible for deployment (Nkengasong et al., 2018).

3.3.2 Data collection

Data is a crucial component in the development of AI systems. The first step in generating AI models to assist with microscopical image analysis is the digitization of images. Unlike radiology, which is inherently digital, the majority of traditional microscopes remain analog, presenting a significant barrier to the application of AI. To address this challenge, both large companies and the research community have made strides in developing solutions to digitize microscopy. These solutions range from converting analog microscopes into digital devices by attaching mobile phones to their oculars, to creating low-cost digital microscopes, robotic microscopes, and commercially available scanners.

Commercially available scanners are known for their speed and ease of use, yet their high cost and large dimensions render them less suitable for deployment in LMICs, where resources and infrastructure may be limited. Recognizing this gap, several innovative devices have been developed to provide affordable and practical alternatives. These include mobile phone-based systems and low-cost scanners designed to work with traditional light microscopes, thereby

enabling the digital capture of images that can be analyzed by AI to assist in the diagnosis of diseases such as malaria and NTDs(Meulah et al., 2023; Vasiman et al., 2019).

The evolution of mobile microscopes has been particularly noteworthy. The first generation of these custom devices involved simple modifications, such as mounting a ball lens onto a smartphone. This setup was used effectively for diagnosing helminth infections like *Schistosoma haematobium* (Bogoch et al., 2014; Bogoch et al., 2013). Building on these initial innovations, the second generation of mobile microscopes introduced more advanced designs, utilizing reverse lenses and 3D-printed adapters. These adaptations enabled the diagnosis of infections like *Loa loa* and *Schistosoma haematobium* with improved accuracy and convenience (Armstrong et al., 2022; Coulibaly et al., 2016; Coulibaly et al., 2023; Cybulski et al., 2014; D'Ambrosio et al., 2015; Kamgno et al., 2017; Skandarajah et al., 2014).

Another approach involves using smartphones as digital cameras by aligning them with the microscope's ocular, often using a 3D-printed support for stability. This method combines the high resolution of traditional light microscopes with the ubiquity and processing power of smartphones, making it a practical and accessible solution for many settings (Abdelmula et al., 2023; Dacal et al., 2021; Morais et al., 2022; Quinn et al., 2016; F. Yang et al., 2020). Additionally, there have been efforts to develop low-cost digital scanners, which offer another viable option for digitizing samples in resource-constrained environments (Collins et al., 2020; García-Villena et al., 2021).

Training AI models requires vast amounts of data, and these models often learn both correlations and biases present in the data (Mehrabi et al., 2021). In recent years, numerous studies have explored the fairness of the AI (Caton & Haas, 2024; Celi et al., 2022; Mehrabi et al., 2021). Bias in AI can manifest in various ways, including data bias, algorithmic bias and human bias, leading to unfair, inaccurate, or harmful outcomes, particularly when these systems are deployed in sensitive areas like healthcare or criminal justice. For example, an algorithm designed to diagnose malignant skin lesions was found to perform significantly worse on darker skin tones because the training dataset included few images of dark skin (Daneshjou et al., 2021).

To mitigate the data bias, data collection must include samples from diverse centres, including different ages, ethnicity and gender. In medical imaging, the variability and quality of data depend on factors such as sample preparation, the imaging device used, and the technician collecting the data, all of which can vary widely between institutions. Figure 3.6 illustrates this by showing how a single sample digitized using three different smartphones can result in noticeable differences in data quality. To improve robustness, it is recommended to gather data from multiple institutions. However, despite this recommendation, very few studies report on multi-centric datasets. A review of 130 Food and Drug Administration-approved medical AI devices found that only 37 studies reported using data from multiple organisations (Liang et al., 2022).

In addition to the images, other patient data are often collected, and this data typically contains sensitive information that requires protection against unauthorized access. To ensure data security, several mechanisms can be implemented: encryption in transit (during data upload and download) and at rest (when the data is stored), access control to secure

Figure 3.5: Examples of image acquisition devices used for malaria and NTD diagnosis. A) Ball-lens attached to a smartphone, proposed by Bogoch et al. for the diagnosis of *Schistosoma haematobium* (Bogoch et al., 2013). B) LoaScope, proposed by D'Ambrosio et al. for the diagnosis of *Loa loa* (D'Ambrosio et al., 2015). C) MoMic digital microscope scanner, proposed by Holmström et al. for the diagnosis of *Schistosoma haematobium*. D) Commercial single slide scanner (Grundium), used to scan soil-transmitted helminth (STH) samples (Lundin et al., 2024). E) Schistoscope, proposed by Oyibo et al. for scanning *Schistosoma haematobium* samples. F) Smartphone attached to an optical microscope, employed by various authors for digitizing microscopical images (Dacal et al., 2021; Morais et al., 2022; Quinn et al., 2016; F. Yang et al., 2020)
.

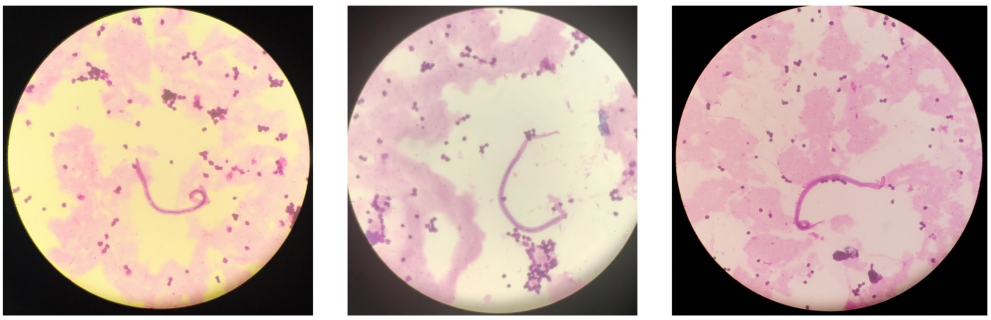


Figure 3.6: Comparison of the same sample preparation digitized using three different smartphones, highlighting the potential variability and bias in data collection due to differences in camera quality, image resolution, and color balance.

storage, and data aggregation, which involves replacing unique values with summary values, or anonymization to protect individual identities. Moreover, medical data often includes personally identifiable information, making it crucial to ensure data privacy and comply with regulations and policies such as the General Data Protection Regulation to prevent patient re-identification.

3.3.3 Data labelling

Data labeling, along with data collection, largely determines the performance of AI models. Currently, annotations often rely on human labeling, which in medical settings can be particularly challenging and expensive due to the need for domain expertise. Several labeling strategies exist:

1. Internal Labeling: This involves using an in-house team to label your dataset. This approach offers full control over the labeling process, which can enhance accuracy and mitigate data security issues. However, the in-house team may lack the specialized knowledge required for certain medical annotations.

2. External Manual Labeling: Collaborating with other research institutions or hospitals can provide access to expert labelers who can both supply and annotate data. However, expert availability is often limited, and studies have reported significant intra- and inter-observer variability, highlighting the challenges in analyzing and labeling medical images (Akintayo et al., 2018; D. Das et al., 2022; Davidson et al., 2021; Karimi et al., 2020; Kim et al., 2019; Matek et al., 2019; Rajaraman et al., 2019; Y. Y. Wu et al., 2020). To address this, some studies use multiple labelers to improve accuracy, although this can increase costs.

3. Crowdsourcing Annotation: This strategy leverages collective intelligence by gathering labels from a large group of non-experts to create a single, unique annotation. While traditionally used in AI for common object labeling (Kovashka et al., 2016; Russakovsky, Deng, et al., 2015), some studies have shown that processing inputs from multiple non-experts can achieve annotation quality comparable to that of medical experts (Keshavan et al., 2019; Linares et al., 2019; Luengo-Oroz et al., 2012).

4. Active Learning: To reduce the time and effort required for annotation, active learning is employed. It starts with labeling a small dataset manually, then training an AI model to generate labels for the next batch. Human labelers review and correct these predictions, which are then used to retrain the model. This process is repeated until the entire dataset is labeled, significantly reducing the annotation burden (Budd et al., 2021; Davidson et al., 2021; Russakovsky, Li, et al., 2015).

3.3.4 Algorithm development and retrospective validation

The selection of model architecture and learning strategy is influenced by the specific use case, the availability of data, and the intended deployment environment—whether on the cloud or at the edge. Each deployment option offers its own set of advantages and drawbacks.

- **Cloud/Server Inference**. Deploying models on the cloud or a server offers several advantages, including virtually unlimited scalability, which allows for easy resource adjustment based on demand. Cloud computing also provides accessibility from any location with an internet connection and simplifies maintenance tasks, such as model updates. However, cloud inference can encounter latency issues due to the physical distance between data centers and end-users, making it less ideal for real-time processing. Additionally, there are concerns about data privacy and security, as data must be transmitted to and stored on remote servers. Deploying a trained model on the cloud using a Docker image grants access to extensive computational resources and scalable hardware, allowing for the deployment of larger models that typically offer superior performance.
- **Edge Device Inference**. Inference on edge devices, such as mobile phones, provides several benefits, with low latency being a key advantage due to the proximity of computation and data storage to the end user, enabling real-time processing. Edge computing also conserves bandwidth by eliminating the need to transfer large volumes of data to cloud data centers. Moreover, edge devices can function offline or with intermittent connectivity, making them suitable for remote use cases. This setup also enhances data privacy and security by keeping sensitive information localized, reducing the risk of data breaches during transmission to data centers.

  Despite these benefits, edge computing is limited by the computational power, memory, and storage capacity of the devices compared to cloud computing. As a result, lightweight models are often preferred for edge deployment. Some applications require offline functionality, making mobile deployment advantageous as it avoids internet dependency, leading to faster and more secure processing.

  To further optimize model size for edge devices, techniques such as model pruning and quantization can be employed. Model quantization enhances model speed by reducing memory usage, such as converting data representation from float32 to int8, which cuts memory requirements by four times. Quantization can be applied during or after model training. Conversely, model pruning involves removing parameters that have minimal impact on prediction accuracy, improving efficiency without significantly reducing model size.

After a model has been trained, it must be evaluated to assess its generalizability and trustworthiness—a process known as retrospective validation. The validation data should be as diverse as possible, representing all potential categories the model might encounter. In medical research, AI models are often trained on data from a single site but are intended for deployment across multiple hospitals. As mentioned earlier, only 37 out of 130 FDA approved medical AI devices reported multi-centric data. In the field of NTDs, as far as we know, no studies have reported the use of multi-centric data, underscoring an area that needs further attention. In contrast, for malaria, several studies have utilized multicentric data, largely due to the availability of existing publicly accessible datasets such as BBBC041v1 and NLM Malaria Screener (Davidson et al., 2021; Manku et al., 2020; Yu et al., 2020).

Rather than reporting an overall metric like accuracy or precision, it's beneficial to compute metrics within subsets of data. These metrics help identify potential biases. If biases are detected, they can be addressed by modifying the input data—such as collecting more diverse data or using data augmentation—adjusting the model by introducing or altering objectives or adding constraints, or by post-processing the results. Additionally, it is important to evaluate performance elements beyond accuracy, including metrics like reliability and applicability.

Cybersecurity is another major concern for AI-driven medical devices (Gordon & Stern, 2019). AI systems may be vulnerable to malicious attacks, which can have harmful consequences for patients, healthcare providers, and insurance companies (Finlayson et al., 2019). For example, subtle modifications to inputs, such as adding random noise that is imperceptible to human eyes, can significantly alter a model's output (Bortsova et al., 2021; X. Ma et al., 2021). Various defense methods have been proposed to protect AI systems from adversarial attacks, including adversarial training and regularization techniques (Ross & Doshi-Velez, 2018; Y. Wang et al., 2021; Yuan et al., 2019).

AI systems are often seen as "black boxes" due to the lack of transparency in their decision-making processes. This lack of explainability poses challenges for integration into clinical workflows, as it hinders trust from both healthcare providers and patients. To address this, different methods have been developed to increase explainability, such as activation maps, saliency maps, or Grad-CAM, are widely used.

Adaptability is another crucial aspect of an AI system. In real clinical workflows, data may change over time—for example, if a hospital upgrades its acquisition systems or alters its staining techniques—resulting in differences between the training and testing data, a phenomenon known as data drift. This can lead to performance degradation. Just as medical experts improve with experience, AI systems should also be capable of continuous learning (Lee & Lee, 2020). Model performance, input data, and changes in labels should be monitored automatically. Clinicians can also contribute to this process by reporting errors and changes in data to the AI team.

3.3.5 Prospective validation

The majority of AI systems are evaluated retrospectively, with very few undergoing prospective evaluation. Wu et al. reported that only 3% of AI algorithms approved by the FDA had undergone prospective validation (E. Wu et al., 2021). Prospective validation involves

integrating the AI system into clinical practice, where clinicians use and validate it according to a predefined protocol. Although this type of validation is rare, it is crucial for assessing the real-world utility of the system. The effectiveness of an AI system is not solely determined by its model performance, but also by how clinicians use and interpret its outputs. To avoid misuse and misinterpretation, clinicians need proper training. Furthermore, the impact of AI integration on both clinical outcomes and workflow efficiency should be evaluated.

Larson et. al proposed a multi-stage evaluation for medical AI (Larson et al., 2021):

1. Phase I - Feasibility Test: This initial phase corresponds to retrospective validation, where developers train an algorithm to assess whether it can compete with the current state of the art, which could be a human expert or an existing algorithm.
2. Phase II - Capability: In this phase, developers refine the algorithm using more data and evaluate its performance on retrospective data in a controlled environment with end-users, simulating real-world conditions. Safety, reliability, and accuracy are assessed, along with the interaction between the algorithm and the operator.
3. Phase III - Effectiveness: The effectiveness of the algorithm is evaluated in two parts: general real-world performance and local validation. The primary objective of general real-world evaluation is to confirm that the algorithm's real-world performance aligns with its performance in the test environment. Feedback from this evaluation should be used to optimize the algorithm, which then undergoes re-evaluation in Phase II. Before clinical implementation, local validation at each site is necessary to identify potential errors due to the algorithm or local quality issues.
4. Phase IV - Durability: AI manufacturers should continuously monitor and evaluate the model's performance, detecting potential data or concept drift and reporting errors. They should also compile user and clinical feedback. Depending on the identified issues, the algorithm may require retraining, updates, and re-evaluation before being redeployed in the clinical workflow.

For algorithms proposed for NTDs, Meulah et al. evaluated their AI algorithm for the detection of *Schistosoma haematobium* eggs in Gabon, focusing on its accuracy and reliability in a real-world setting. Similarly, Ward et al. assessed their AI algorithm for the detection of soil-transmitted helminth eggs, examining key factors such as repeatability, reproducibility, time to result, cost-efficiency, and usability in both laboratory and field environments (Meulah et al., 2023; P. K. Ward et al., 2023).

3.3.6 Regulation

Regulation is crucial for integrating AI into healthcare systems effectively. The term Software as a Medical Device (SaMD) is defined by the International Medical Device Regulators Forum (IMDRF) as "software intended to be used for one or more medical purposes that perform these purposes without being part of a hardware medical device" (IMDRF, 2013). This definition encompasses all software with medical purposes, including AI. While AI is considered SaMD and current approval pathways are applicable, specific regulatory pathways are necessary due to AI's dynamic nature. Unlike traditional software, AI systems require periodic retraining

and updates to maintain optimal performance and adapt to changes in input data.

There is a growing global consensus on the need to update regulatory approvals for medical AI. Various organizations have released new guidelines and best practices to ensure the development of trustworthy AI systems. In the United States, the FDA has published an action plan to update the regulatory framework for AI/ML-based SaMD. This plan emphasizes transparent and real-world monitoring, covering the AI system's journey from pre-market development to post-market performance. The goal is to ensure safety and effectiveness while accommodating AI's iterative improvements (FDA, 2021). The FDA has also issued guidelines outlining good machine learning practices for medical AI development, which include 10 key principles such as the necessity for multidisciplinary expertise throughout the product life cycle, ensuring independent training and test sets, maintaining a representative dataset, and focusing on the performance of the Human-AI team (FDAn, 2021).

In the European Union, the European Parliament has reviewed AI's specific applications in medicine and healthcare. This review identified seven major risks associated with AI in healthcare and provided recommendations for risk assessment. It also outlined seven policy options to enhance the development, evaluation, and deployment of AI solutions in the healthcare sector (EU, 2022).

As of August 2024, there are 950 AI/ML-enabled medical devices approved by the FDA. Of these, 723 systems (76%) are designed for radiology, reflecting the significant focus on imaging applications within AI. In contrast, only 18 devices are approved for hematology, and just 5 for microbiology.

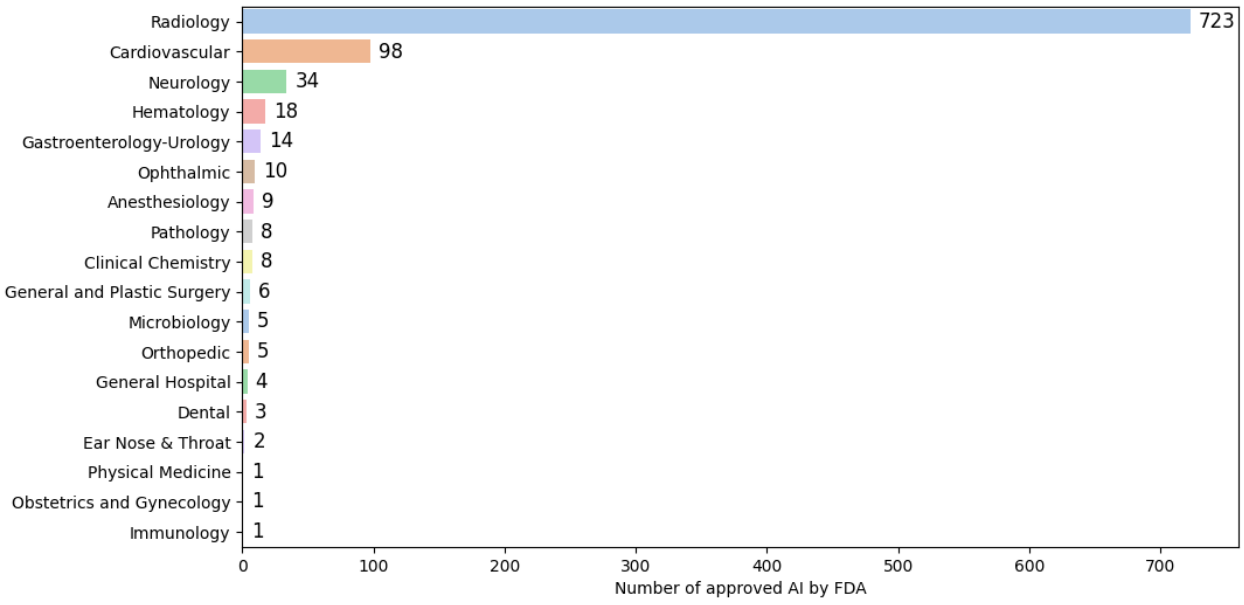


Figure 3.7: Approved AI by FDA in different medical domains. Data source (FDA, 2021)

3.3.7 Integration into the clinical workflow

Regulatory approval allows AI to be integrated into real clinical workflows. However, as mentioned earlier, the performance of AI systems depends on multiple factors, including the quality of input data, the operation, and the interpretation by clinicians.

To ensure optimal performance, it is crucial to provide a clear description of the expected input data, a detailed model description, and an explanation of any limitations. Additionally, continuous monitoring of the AI system's performance is essential. Education programs and awareness campaigns should be implemented to train healthcare professionals on the proper use of AI and to improve the general public's understanding of AI in healthcare, reducing the risk of misuse.

3.4 Research gaps

After reviewing the clinical context and the current state of AI, we will discuss the existing research gaps and challenges.

Overall, there are very few studies focused on the application of AI for NTDs. While some research has addressed STH parasite detection, these studies often employ large architectures that are not suitable for resource-limited settings. Most research efforts culminate in model development and validation using retrospective data, with only a few progressing to deployment and prospective validation. To our knowledge, no AI algorithms for NTDs have been deployed in an edge-computing context, nor have any pilot studies being conducted to assess their technical feasibility. Consequently, there has been no prospective validation of AI for NTDs that evaluates human-AI interaction. Additionally, there are currently no studies investigating the development of foundation models for parasitic diseases using self-supervised learning techniques. Therefore, there is a clear need in the design, development, validation and deployment of specific AI technological solutions for NTDs diagnosis and monitoring to achieve the final goal of translation into the the particular clinical scenarios of the endemic areas.

As previously mentioned, soil-transmitted helminthiasis and filariasis are among the most prevalent neglected tropical diseases. Accordingly, this study prioritizes the development of AI solutions targeting these two conditions, establishing a foundational framework that can be expanded to address other NTDs in future research.