|  |  |  |
| --- | --- | --- |
| ITU Logo | INTERNATIONAL TELECOMMUNICATION UNION**TELECOMMUNICATIONSTANDARDIZATION SECTOR**STUDY PERIOD 2017-2020 | FG-AI4H-G-019 |
| **ITU-T Focus Group on AI for Health** |
| **Original: English** |
| **WG(s):** | Plenary | New Delhi, 13-15 November 2019 |
| **DOCUMENT** |
| **Source:** | TG-Malaria topic driver |
| **Title:** | Initial TDD: TG-Malaria (Malaria detection) |
| **Purpose:** | Discussion |
| **Contact:** | Rose Nakasi Topic driver Makerere University, Uganda | E-mail: g.nakasirose@gmail.com  |

|  |  |
| --- | --- |
| **Abstract:** | This document contains an outline of a topic description document (TDD) on AI based detection of Malaria (TG-Malaria). |

**Table of Contents**

[1 Introduction 3](#_Toc24214579)

[1.1 Document Structure 3](#_Toc24214580)

[1.2 Topic Description 3](#_Toc24214581)

[1.3 Impact of the benchmarking AI Solution 4](#_Toc24214582)

[1.4 Ethical Considerations 4](#_Toc24214583)

[1.5 Existing AI Solutions 4](#_Toc24214584)

[1.6 Existing work on benchmarking 4](#_Toc24214585)

[2 AI4H Topic Group: Current topic group and its mandate 4](#_Toc24214586)

[2.1 Topic group structure 5](#_Toc24214587)

[2.1.1 Topic group participation 5](#_Toc24214588)

[2.1.2 Tools/process of TG cooperation 5](#_Toc24214589)

[2.1.3 TG interaction with WG and FG 5](#_Toc24214590)

[2.1.4 Next meetings 5](#_Toc24214591)

[2.1.5 Next steps 5](#_Toc24214592)

[3 Method 6](#_Toc24214593)

[3.1 AI Input Data Structure 6](#_Toc24214594)

[3.1.1 Available Data 6](#_Toc24214595)

[3.2 AI Output Data Structure 6](#_Toc24214596)

[3.3 Test Data Labels 6](#_Toc24214597)

[3.4 Score and Metrics 6](#_Toc24214598)

[3.5 Undisclosed Test Data Set Collection 6](#_Toc24214599)

[3.6 Benchmarking Methodology and Architecture 7](#_Toc24214600)

[3.7 Reporting Methodology 7](#_Toc24214601)

[4 Results 7](#_Toc24214602)

[5 Discussion 7](#_Toc24214603)

[6 Declaration of Conflict of Interest 7](#_Toc24214604)

[References 7](#_Toc24214605)

# Introduction

Malaria is one of the largest endemic diseases in the Sub Saharan Africa [5]. In Low developed countries (LDCs), the scourge is further buttressed by the lack of enough skilled lab technologists in health centers to detect the disease using the widely accepted gold standard Microscopy method. Thus, the need for reliable detection interventions. This explains the birth of Automated malaria detection using Artificial Intelligence (AI). The aim is to harness AI to automate the detection of malaria in a fast, accurate and cost-effective manner. Of recent AI and machine learning techniques have been successful in different medical image analysis tasks and have a capability to improve public health.

The aim of this document therefore is to develop a standardised benchmarking approach for AI based detection of Malaria.

## Document Structure

This TDD is dedicated to support standardised bench-marking of AI-Based detection of Malaria. The document will cover all core aspects relevant to topic including data collection and annotation, selection of AI model for Malaria detection, Algorithm performance and evaluation, collection of test datasets and benchmarking methodologies.

## Topic Description

According to the World Health Organization report of 2016, nearly half of the world population is at risk of malaria [5]. Records from the WHO report of 2015 indicate that in 2015, 212 million cases reported, Malaria accounted for over 480,000 deaths, 90% of which were from Africa, 7% from S.E Asia and 2% from Eastern Mediterranean region [6]. Although there were fewer Malaria cases in 2017 than in 2010 according to WHO report of 2017, data for the period 2015-2017 highlighted that no significant progress in reducing global Malaria cases was made in this timeframe [7]. Malaria is thus of major concern to public health and therefore the need for early, fast and accurate diagnosis.

The gold standard method for detection of Malaria is microscopy of blood smear slides. Unlike Rapid Diagnostic Tests (RDTs), microscopy supports direct parasite detection and identification and provides monitoring of systemic inflammation and its response to therapy. Detection of malaria requires examination of thin and thick blood smear images through conventional light microscopy. In general, Malaria parasite detection, species identification, and parasitemia determination requires expertise from trained Microscopists (lab technicians).

Malaria control can effectively be achieved by a fast, consistent and accurate diagnosis. This requires the expertise of Microscopists to operate the gold standard method of microscopy screening of Malaria. Un fortunately, highly Malaria endemic Countries have very few expert Microscopists to diagnose and interprete the results of the huge numbers of malaria patients.

A nationwide study in Ghana, for example, found 1.72 microscopes per 100,000 population, but only 0.85 trained laboratory technicians per 100,000 population [1] which is grossly inadequate. As a result, diagnoses are often made on the basis of clinical signs and symptoms alone, which are error-prone and leads to higher mortality, drug resistance, and the economic burden of buying unnecessary drugs [2].

Computational Microscopy using Artificial Intelligence technologies aims to reduce the need for many human Microscopists by providing a fast, consistent and accurate diagnosis with minimum human intervention. AI models have the capability to learn good representations of image data with reduced turnaround time bridging the gap for lack of enough skilled Micoscopists and significantly improving diagnostic performance and reducing health costs associated to patient care and treatment.

## Impact of the benchmarking AI Solution

The benchmark solution for AI based detection of malaria should focus on developing AI tools that can adequately detect malaria in a fast, accurate, cost effective and reliable manner. The benchmark will be impactful if the selected solutions are robust enough and relevant especially in developing but highly Malaria Endemic Countries. This is envisioned to improve public health and to synergise with goal 3 Target 3.3 of the Sustainable Development Goal (SDG) which aims to have Malaria endemic end by 2030.

## Ethical Considerations

TBC

**Ethical consideration of benchmarking including its data acquisition**:

* Ethical consideration of data collection must follow Ethical practises like acquisition of ethical approvals from Country medical Authorities through working IRBs.
* Data capture and annotation must be done by qualified medical experts.
* Data anonymity must be considered by removing all references to personal identifiers.
* Integration of other datasets must also prevent personal identifiers.

**Ethical consideration on usage of AI**:

* For any testing task, patients must be informed about the use of the AI tool in detection of malaria and consent from them must be sought.
* Best scientific practices and accuracies of the model must be assured.

## Existing AI Solutions

At the AI and Data Science lab of Makerere University, we have deployed both traditional machine learning and deep learning algorithms for pathogen detection in thick blood smear samples and improvements in detection accuracies have been registered. We have also extended this to other related microscopy diagnosis challenges for example in the detection of tuberculosis and intestinal parasites [4].

An extensive study by Rosado et al [3] has been carried out reviewing the various image processing and analysis approaches for the automated detection of Malaria with the conclusion that improvements in accuracy are still needed. Some AI tools tend to fail on the undisclosed data sets due to false alarms. Therefore, there is a need for a standard benchmarking for the training and testing the data sets.

## Existing work on benchmarking

There has been discussion on having a need for a benchmarking in area of AI for health but for the Malaria detection, there is no benchmarking done till now. The tools available have been trained on the data set that was made available to them mostly hospital based.

# AI4H Topic Group: Current topic group and its mandate

The current topic group is specific and relevant to AI4H. The objectives are:

1. to provide a forum for open communication among various stakeholders,
2. to agree upon the benchmarking tasks of this topic and scoring metrics,
3. to facilitate the collection of high-quality labelled test data from different sources,
4. to clarify the input and output format of the test data,
5. to define and set-up the technical benchmarking infrastructure, and
6. to coordinate the benchmarking process in collaboration with the Focus Group management and working groups.

The primary output of a topic group is one document that describes all aspects of how to perform the benchmarking for this topic. (The document will be developed in a cooperative way by suggesting changes as input documents for the next FG-AI4H meeting that will then be discussed and integrated into an official output document of this meeting. The process will continue over several meetings until the topic description document is ready for performing the first benchmarking.)

## Topic group structure

TBC

Topic Groups summarize uses cases of a certain health topic or problem and similar AI benchmarking requirements. However, inside a Topic Group different Sub-topic Groups can be established to pursue different topic-specific specializations. TG-Malaria will start without separate subtopic Groups. However, it is possible that during the process subtopics will be introduced.

### Topic group participation

The topic group on “AI based detection of Malaria” currently involves an interdisciplinary network between the health team at Mulago Referral Hospital with expertise in Microscopy and the technical AI and 3D printing experts from the AI and Data Science research group at Makerere University.

The participation in both the focus and Topic Group is generally open and free of charge. To participate, one can follow the “call for participation” document outlining the process for joining the Focus Group and the Topic Group. For this topic, the corresponding call can be accessed through the official website (<https://itu.int/go/fgai4h>).

### Tools/process of TG cooperation

TBC

As will be regulated by the FG.

### TG interaction with WG and FG

TBC

As will be regulated by the FG.

### Next meetings

The Focus Groups meets about every two months at changing locations. The upcoming meetings are:

* G: New Delhi, India; 11-15 November 2019
* H: Brasilia, Brazil; January 2020
* I: March 2020 (TBC)​
* J: Geneva, Switzerland, 4-8 May 2020 (TBC)

An up to date list can be found at the official [ITU FG AI4H website](https://www.itu.int/en/ITU-T/focusgroups/ai4h/Pages/default.aspx).

### Next steps

We aim to extend the topic of Malaria detection to all Malaria endemic Countries, while bringing together AI solutions and data from different countries. Next steps for the group can be of different forms:

1. Collection of labelled data from different sources. Microscopic Image datasets (both from thin blood smear image and thick blood smear images) and any other data directly linked to malaria endemicity from sources like (environment, clinical records) that could improve the detection confidence is of high value. The ultimate goal is to make a strong malaria detection model given data from different systems and countries.
2. Selection and implementation of AI models and approaches related to malaria detection.
3. Suggestions on scoring metrics.
4. Contributing to the development of a viable and accepted benchmarking framework.
5. Support to the group on different aspects (data, methods, benchmarking, etc.) of this topic
6. Extension of the solution to improve disease surveillance and prediction.

# Method

TBC

## AI Input Data Structure

TBC

Blood smear Images of both thick and thin blood smear slides that have been annotated by laboratory experts from different Health facilities in different Malaria endemic countries would be required and an undisclosed test data for evaluation of the tool.

The labels will depend on the specific attribute to be investigated in the image.

All data will be subject to permissions from the different country authorities.

### Available Data

Currently, we have access to 1000 images of thick blood smear slides that have been annotated by laboratory experts from Mulago referral hospital, however large more amounts of data of both thick blood smear and thin blood smear images from different Health facilities in different malaria endemic countries would be required for machine learning models and an undisclosed test data for evaluation of the tool.

## AI Output Data Structure

TBC

The AI output should include AI tools detecting malaria parasites, species, Test data labels

Confidentiality of gold standard testing data results would be maintained.

## Test Data Labels

TBC

A label/ annotation will be given of the blood smear Image that contains the malaria parasites. The labels will depend upon the specific condition that is being benchmarked and also the type of AI task.

## Score and Metrics

TBC

To evaluate AI tool’s performance, labelled Dataset of blood smear images would be taken and tested against performance of AI. The algorithm evaluation mechanism should include metrics like ROC accuracy, precision, recall, specificity F1 scores, specificity, sensitivity, mean Average Precision (mAP) and the choice will base on the algorithm used and purpose of the task.

## Undisclosed Test Data Set Collection

TBC

1. In order to assess algorithm robustness, sufficient undisclosed image data would be collected. This is envisioned to come from different health facilities both public and private.
2. There is need for examination of the quality of undisclosed dataset by a panel of experienced and skilled lab technicians. Bias in data will be considered.
3. An agreeable number of test data for a benchmarking task will be specified.
4. Annotation process of data with expert labels should also be agreed upon.

## Benchmarking Methodology and Architecture

TBC

* technical architecture
* hosting (IIC, etc.)
* possibility of an online benchmarking on a public test dataset
* protocol for performing the benchmarking (who does what when etc.)
* AI submission procedure including contracts, rights, IP etc. considerations

## Reporting Methodology

TBC

# Results

TBC

# Discussion

TBC

* Discussion of the insights from executing the benchmarking on
	+ external feedback on the whole topic and its benchmarking
	+ technical architecture
	+ data acquisition
	+ benchmarking process
	+ benchmarking results
	+ field implementation success stories

# Declaration of Conflict of Interest

TBC

References

[1] I. Bates, V. Bekoe, and A. Asamoa-Adu. Improving the accuracy of malaria-related laboratory tests in Ghana. Malar. J. Vol.3. No. 38, 2004.

[2] C.A. Petti, C.R. Polage, T.C. Quinn, A.R. Ronald, and M.A. Sande. Laboratory medicine in Africa: A barrier to effective health care. Clinical Infectious Diseases, Vol. 42, N0. 3, PP. 377-382, 2006.

[3] L. Rosado, J.M. Correia da Costa, D. Elias, and J.S. Cardoso. A review of automatic malaria parasites detection and segmentation in microscopic images. Anti-Infective Agents, Vol 14, No. 1, pp 11-22, 2016.

[4] J. A. Quinn, R. Nakasi, P. K. B. Mugagga, P. Byanyima, W. Lubega, and A. Andama. Deep convolutional neural networks for microscopy-based point of care diagnosis. In proceedings of International Conference on Machine Learning for Health Care, Volume 50., 2016.

[5] WHO. World malaria report. Geneva, Switzerland, World Health Organization., 2016.

[6] WHO. World malaria report. Geneva, Switzerland, World Health Organization., 2015.

[7] WHO. World malaria report. Geneva, Switzerland, World Health Organization., 2017.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_