

Two-stage automated diagnosis framework for urogenital schistosomiasis in microscopy images from low-resource settings

Oyibo, Prosper; Meulah, Brice; Bengtson, Michel; Lieshout, Lisette van; Oyibo, Wellington; Diehl, Jan-Carel; Vdovine, Gleb; Agbana, Temitope E.

DOI

[10.1117/1.JMI.10.4.044005](https://doi.org/10.1117/1.JMI.10.4.044005)

Publication date

2023

Document Version

Final published version

Published in

Journal of Medical Imaging

Citation (APA)

Oyibo, P., Meulah, B., Bengtson, M., Lieshout, L. V., Oyibo, W., Diehl, J.-C., Vdovine, G., & Agbana, T. E. (2023). Two-stage automated diagnosis framework for urogenital schistosomiasis in microscopy images from low-resource settings. *Journal of Medical Imaging*, 10(4), Article 044005. <https://doi.org/10.1117/1.JMI.10.4.044005>

Important note

To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.

Two-stage automated diagnosis framework for urogenital schistosomiasis in microscopy images from low-resource settings

Prosper Oyibo^{a,b}, Brice Meulah^{c,d}, Michel Bengtson^c, Lisette van Lieshout^c,
Wellington Oyibo^b, Jan-Carel Diehl^{e,*}, Gleb Vdovine^a, and Tope Agbana^a

^aDelft University of Technology, Delft Center for Systems and Control, Faculty of Mechanical, Maritime, and Materials Engineering, Delft, The Netherlands

^bUniversity of Lagos, College of Medicine, Centre for Malaria Diagnosis, NTD Research, Training, and Policy/ANDI Centre of Excellence for Malaria Diagnosis, Lagos, Nigeria

^cLeiden University Medical Centre, Department of Parasitology, Leiden, The Netherlands

^dCentre de Recherches Medicales des Lambaréné, CERMEL, Lambarene, Gabon

^eDelft University of Technology, Department of Sustainable Design Engineering, Faculty of Industrial Design Engineering, Delft, The Netherlands

ABSTRACT. **Purpose:** Automated diagnosis of urogenital schistosomiasis using digital microscopy images of urine slides is an essential step toward the elimination of schistosomiasis as a disease of public health concern in Sub-Saharan African countries. We create a robust image dataset of urine samples obtained from field settings and develop a two-stage diagnosis framework for urogenital schistosomiasis.

Approach: Urine samples obtained from field settings were captured using the Schistoscope device, and *S. haematobium* eggs present in the images were manually annotated by experts to create the SH dataset. Next, we develop a two-stage diagnosis framework, which consists of semantic segmentation of *S. haematobium* eggs using the DeepLabv3-MobileNetV3 deep convolutional neural network and a refined segmentation step using ellipse fitting approach to approximate the eggs with an automatically determined number of ellipses. We defined two linear inequality constraints as a function of the overlap coefficient and area of a fitted ellipses. False positive diagnosis resulting from over-segmentation was further minimized using these constraints. We evaluated the performance of our framework on 7605 images from 65 independent urine samples collected from field settings in Nigeria, by deploying our algorithm on an Edge AI system consisting of Raspberry Pi + Coral USB accelerator.

Result: The SH dataset contains 12,051 images from 103 independent urine samples and the developed urogenital schistosomiasis diagnosis framework achieved clinical sensitivity, specificity, and precision of 93.8%, 93.9%, and 93.8%, respectively, using results from an experienced microscopist as reference.

Conclusion: Our detection framework is a promising tool for the diagnosis of urogenital schistosomiasis as our results meet the World Health Organization target product profile requirements for monitoring and evaluation of schistosomiasis control programs.

© The Authors. Published by SPIE under a Creative Commons Attribution 4.0 International License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JMI.10.4.044005](https://doi.org/10.1117/1.JMI.10.4.044005)]

Keywords: schistosomiasis diagnosis; deep learning; semantic segmentation; ellipse fitting; edge artificial intelligence

Paper 23050GR received Feb. 20, 2023; revised May 23, 2023; accepted Jul. 21, 2023; published Aug. 7, 2023.

*Address all correspondence to Jan-Carel Diehl, j.c.diehl@tudelft.nl

1 Introduction

Schistosomiasis is endemic in 76 countries worldwide with ~252 million people infected and an estimated 779 million people at risk of infection.¹ Schistosomiasis is caused by blood flukes of the genus *Schistosoma* (*S*); both *S. mansoni* (intestinal schistosomiasis) and *S. haematobium* (urogenital schistosomiasis) are endemic in Africa.² Schistosomiasis presents a substantial public health and economic burden as it is a disease of poverty. In the drive to attain the World Health Organization (WHO) control and elimination targets, diagnosis for adequate monitoring of interventions and surveillance is critical.^{2,3} Recently, the WHO published the diagnostic target product profiles (TPP) for monitoring, evaluation, and surveillance of schistosomiasis control programs,⁴ which identifies development of diagnostic tests for *S. haematobium* detection as a high-risk requirement due to lack of its availability. The TPP suggests a semi-quantitative analysis, capable of providing some degree of information regarding intensity of infection, as ideal for a diagnostic test for schistosomiasis to support monitoring and evaluation.⁴ Currently, microscopy is the WHO reference standard for the diagnosis of schistosomiasis in resource-limited settings. For the detection of *S. haematobium* infection, urine samples, after filtration, sedimentation, or centrifugation, are microscopically examined for the presence of eggs.³ This method is operator dependent, costly, laborious, and time-consuming. Furthermore, it requires expertise, which means microscopy skills need to be gained and maintained, which can be an economic challenge, particularly in remote rural communities.³ There is also the risk of visual health complications among microscopists resulting from excessive workload due to the low ratio of trained microscopists to samples for analysis in endemic regions.⁵ Hence, a field adaptable, rapid, and easy-to-use automated diagnosis is relevant for the prompt detection of cases, which will facilitate mapping and monitoring of interventions.⁴ Recent advances in opto-mechanics and opto-electronics have rapidly transformed the field of biomedical optics. Optical imaging technologies, such as conventional light microscopes, are being redesigned to integrate and miniaturize portable light microscopes for use at the point of care.⁶⁻⁹ Although these technologies are readily available in high-income countries, unfortunately, nearly all schistosomiasis cases are seen in low-resource regions of low-income countries, significantly justifying the need for cost-effective and easy-to-use smart diagnostic technologies. In this work, we address these challenges by first increasing the size of the *S. haematobium* (SH) dataset in our previous work¹⁰ from 5198 to 12,051 images of clinical samples.¹¹ We carry out detection and counting of *S. haematobium* eggs present in each image by proposing a two-stage framework consisting of a DeepLabv3 with MobilenetV3 backbone deep convolutional neural network¹² trained on the SH dataset using a transfer learning approach. The second stage of our proposed framework is a refined segmentation and egg counting procedure, which adapts the region-based fitting of overlapping ellipses¹³ to efficiently separate the boundaries of overlapping eggs in the image. Finally, the detected isolated eggs are screened for the presence of an egg, which meets the defined boundary condition before the sample can be determined as positive/negative diagnosis. We further demonstrate the robustness and applicability of the proposed framework for field diagnoses of urogenital schistosomiasis by implementing our framework on an Edge AI system (Raspberry Pi + Coral USB accelerator) and testing 65 clinical urine samples obtained in a field settings in Nigeria. The main contributions of this work can be summarized as follows.

1. A large-scale *S. haematobium* egg dataset of 12,051 images captured in field settings is created with respective manually annotated mask images. The dataset contains images with artifacts, such as crystals, glass debris, air bubbles, and fibres.
2. A *S. haematobium* egg detection framework consisting of the DeepLabv3 with MobileNetV3 backbone deep convolutional neural network, trained using transfer learning approach for semantic segmentation of the eggs. This effectively segments transparent eggs in noisy images taken in the field. The framework also separates overlapping eggs using a refined segmentation algorithm resulting in a more accurate egg count.
3. The implementation and testing of the *S. haematobium* egg on an Edge AI system to demonstrate its field applicability for the diagnosis of schistosomiasis in low-resource settings.

2 Related Work

A pioneering study on the identification and classification of human helminth eggs based on computer vision algorithms was carried out by Ref. 14. However, their focus was on helminth eggs found in microscopic faecal samples. Subsequent works¹⁵⁻¹⁸ included the detection of *S. haematobium* eggs found in urine but only in images pre-captured by professional clinical operators mostly with isolated and non-overlapping eggs in the field of view (FoV) images. Regarding the detection of *S. haematobium* eggs in microscopy images of urine from field settings, these images contain many artifacts with morphological and textural similarity to eggs, such as crystals, glass debris, air bubbles, fabric fibres, and human hair. This makes it difficult to achieve high accuracy using traditional AI methods, which detect objects in the images based on some threshold value or discontinuous local features of an image. The *S. haematobium* eggs are oval-shaped structures (110 to 170 μm long and 40 to 70 μm wide) with a thick transparent capsule and a sword-shaped protrusion known as the terminal spine located at the narrow end of the egg. Detecting an egg is challenging due to its similar appearance to its surroundings. Automated detection of an isolated *S. haematobium* egg by thresholding the cross-correlation coefficient of two sets of invariant moments for both a reference and sample image was performed by Ref. 19. However, this method had poor performance in noisy images and hence cannot be used for *S. haematobium* eggs detection in field settings.

Recently, deep learning algorithms were used by Ref. 20 to solve the challenges of *S. haematobium* egg detection in images captured in field settings. Using transfer learning, they compared RetinaNet,²¹ MobileNet,²² and EfficientDet²³ architectures pre-trained on the COCO 2017 dataset.²⁴ They retained the feature extraction layers and fine-tuned the dense layers of these models to detect *S. haematobium* eggs as a single class. The RetinaNet architecture had improved egg detection performance with egg counts closely related to manual egg counts obtained by a trained user. It was also able to detect isolated eggs and reject other debris from a crowded FoV. However, air bubbles were incorrectly classified as eggs, and the automated detection of eggs aggregated in large clumps with other eggs or debris remained a challenge. In our previous work, we developed a low-cost automated digital microscope (Schistoscope V5.0) with AI for the detection of *S. haematobium* eggs,¹⁰ and we reported the results from a field validation study in Nigeria.¹¹ A U-Net model²⁵ trained with the *S. haematobium* dataset consisting of 5198 images captured from both clinical and spiked urine samples was used for automated egg detection. Although we achieved a high diagnostic sensitivity of 87.3%, the diagnostic specificity was low (48.9%). This was due to the high number of false positives by the U-Net architecture and the inability of the segmented pixel area-based linear model to filter out incorrectly segmented eggs while counting.

All these studies show that deep learning is a promising approach for the automated diagnosis of urogenital schistosomiasis. However, developing a model that is field applicable requires a robust dataset of images with varying degrees of urine artifacts from field settings. Also the separation of overlapping eggs for improved estimation of infection intensity has remained a challenge. This paper proposes a two-stage framework to solve these challenges.

3 Methods

To meet the WHO TPP requirements for a diagnostic test for schistosomiasis, the proposed urogenital schistosomiasis diagnostic framework consists of two stages (Fig. 1). The first stage involves the semantic segmentation of candidate *S. haematobium* eggs in captured images. The segmentation results are further refined in the second stage by ellipse fitting and morphological filtering of the segmented regions. The two-stage framework minimizes false positive detection that enables a high diagnostic specificity, which is a requirement for diagnostic tools for monitoring and evaluation of schistosomiasis control programs and determining transmission interruption.

3.1 Sample Image Capture and *S. haematobium* Egg Annotation

The details of the Schistoscope's mechanical precision and optical quality are described in our previous work.¹⁰ The Schistoscope optical system consists of a 4 \times magnification microscope

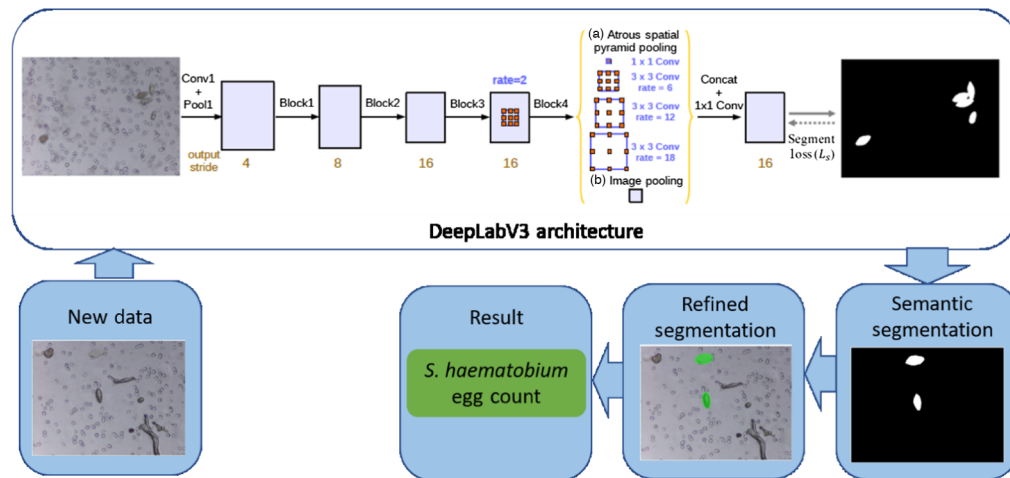


Fig. 1 Schematics of the proposed two-stage diagnosis framework urogenital schistosomiasis with DeepLabV3-MobileNetV3 deep learning architecture for semantic segmentation of eggs and refined segmentation for overlapping eggs separation and count.

objective and a Raspberry Pi High-Quality Camera Module V2.1 equipped with a Sony IMX477R camera sensor. The camera sensor has a pixel-pitch of $1.55 \mu\text{m}$ and registers an image size of 2028×1520 pixels. The device consists of an autofocus and an automated slide scanning system. For urine filtration, we made use of a 13 mm filter membrane, which results in 117 image grid segments per sample when scanned by the device. The *S. haematobium* eggs in the captured images used for training and development testing were manually annotated by an expert parasitologist using the coco annotator tool.²⁶ The annotation process is highlighted as follows.

1. Annotation of the exact boundary pixels of the *S. haematobium* eggs was not strictly enforced due to the limitation posed by the size of the eggs.
2. The pixel values of the filter membrane and artifacts in the ground-truth image were labeled as “0” (background) and the eggs as “1” (foreground).
3. There were few *S. mansoni* eggs found in the images of the clinical urine samples and their pixel values were labeled as “1” (foreground).
4. Pixels of partially cut eggs at the edges of the images were labeled as “1” (foreground).
5. The region of the eggs covered by artifacts was labeled as “0” (background).

3.2 Stage 1: Semantic Segmentation of *S. haematobium* Eggs

3.2.1 Transfer learning using DeepLabv3-MobileNetV3

In transfer learning, a model trained on one task is repurposed to another related task, usually by some adaptation toward the new task. This approach is mainly useful for tasks where enough training samples are not available to train a model from scratch, such as medical image classification for neglected tropical diseases or emerging diseases.²⁷ To overcome the limited data sizes, transfer learning was used to retrain the DeepLabv3-MobileNetV3¹² model for semantic segmentation of candidate *S. haematobium* eggs using the SH dataset. DeepLabv3 is a semantic segmentation architecture that was developed to handle the problem of segmenting objects at multiple scales. Modules are designed, which employ atrous convolution in cascade or in parallel to capture multi-scale context by adopting multiple atrous rates. We initialize the model with weights obtained from the pre-trained model on a subset of COCO train2017, on the 20 categories that are present in the Pascal VOC dataset.²⁸ Since our case consists of two output classes (background and foreground), we replace the 21-output channel convolutional layer with a single output-channel convolutional layer. The weights of all layers of the model are then updated during the training stage.

3.2.2 Loss function

The model was trained using the Dice similarity coefficient (DSC) loss function,²⁹ which is widely used in medical image segmentation tasks to address the data imbalance problem between foreground and background:

$$L_{\text{DSC}} = \frac{2 \sum_{x,y} (S_{i,x,y} \times G_{i,x,y})}{\sum_{x,y} S_{i,x,y}^2 + \sum_{x,y} G_{i,x,y}^2}, \quad (1)$$

where $S_{i,x,y}$ and $G_{i,x,y}$ refer to the value of pixel (x, y) in the segmentation result S_i and ground truth G_i , respectively.

3.3 Stage 2: Refined Segmentation of *S. haematobium* Eggs

To solve the challenge of obtaining accurate egg counts in the occurrence of false positives or overlapping and clustered eggs, we adopted a refined segmentation procedure, which involves fitting ellipses over the region of interest in the binary output image of the semantic segmentation. The refined segmentation algorithm as shown in Algorithm 1 operates in a number of steps, which can be summarized as follows.

3.3.1 Optimization problem formulation

We assume a binary image I that represents the segmentation mask output of the DeepLabV3-MobileNetV3 deep neural network model. The binary image may contain one or more sliced binary region image R , which has the same size as the bounding box. This region image R represents a segmented isolated or overlapping eggs. A pixel p of R belongs to either the foreground FG ($R(p) = 1$) or the background BG ($R(p) = 0$). The area A_R of the segmented egg is given by

$$A_R = \sum_{p \in \text{FG}} R(p). \quad (2)$$

We also assume a set E_R of N_R ellipses are fitted over the region image R . The binary image U_E is defined such that $U_E(p) = 1$ at point p that is inside any of the ellipse $E_{R,i}$; otherwise $U_E(p) = 0$. Also we define the coverage $\alpha(E_r)$ of the segmented eggs by the given set of ellipses E_R as

$$\alpha(E_r) = \frac{1}{A_R} \sum_{p \in \text{FG}} R(p) U_E(p). \quad (3)$$

Essentially, $\alpha(E_r)$ is the percentage of the segmented eggs that are under some of the ellipse in E_R . Let the sum of the areas of all the ellipses be denoted by $|E_R| = \sum (i=1) |E_{R,i}|$ and let $C(E_R)$ denote the coverage area by all the ellipses:

$$C(E_R) = \sum_{p \in R} U_E(p). \quad (4)$$

It should be stressed that $C(E_R) < |E_R|$, with the equality holding in the case that all ellipses are pairwise disjoint. This is because in case of two overlapping ellipses, $|E_R|$ counts the area of their intersection two times, while $C(E_R)$ does not. Similar to the work of Ref. 13, we want to maximize the shape coverage $\alpha(E_R^*)$ with a set of ellipses E_R^* whose covered area by all ellipses $C(E_R^*)$ is as close as possible to A_R :

$$E_R^* = \arg \max_{E_R} \alpha(E_r) \quad \text{s.t.} \quad C(E_R) = A_R. \quad (5)$$

We defined a model complexity measure the ratio of the area A_R of the segmented region to experimentally observed average area of segmented isolated egg A_R^* :

Algorithm 1 The refined segmentation algorithm

```

1 Input: Binary segmentation mask image  $I$ 
2 Output: Set of ellipse  $E_I^*$ , egg count  $N_I^*$ 
3  $N_I^* = 0$ 
4  $E_I^* = \emptyset$ 
5 for each region image  $R \in I$  do
6    $N_R^* = 0$ 
7    $E_R^* = \emptyset$ 
8    $AIC_R^* = \infty$ 
9    $N_{lb}, N_{ub} = \text{ComputeBoundary}(A_R)$ 
10   $N_R = N_{lb}$ 
11  repeat
12     $E_R = \text{FitEllipse}(R, N_R)$ 
13     $AIC_R = \text{ComputeAIC}(R, U_E)$ 
14    if  $AIC_R < AIC_R^*$  then
15       $N_R^* = N_R$ 
16       $AIC_R^* = AIC_R$ 
17       $E_R^* = E_R$ 
18       $N_R = N_R + 1$ 
19    until  $N_R = N_{ub}$ 
20     $E_I^* = \text{union}(E_I^*, E_R^*)$ 
21     $N_I^* = N_I^* + N_R^*$ 
22  end
23 return  $N_I^*, E_I^*$ 

```

Legend I : Binary segmentation mask image E_I^* : Optimal set of ellipses for I N_I^* : Optimal number of ellipses for I R : Segmented egg region image N_R^* : Optimal number of ellipses for R E_R^* : Optimal set of ellipses for R AIC_R^* : Optimal Akaike Information Criterion for R N_{lb} : Lower boundary for N N_{ub} : Upper boundary for N

$$C = \frac{A_R}{A_R^*}. \quad (6)$$

To estimate the optimal number N_R^* of ellipses in a segmented egg region image R , a trade-off between the egg coverage $\alpha(E_r)$ and the model complexity C is optimized by employing

the Akaike information criterion (AIC).³⁰ The AIC-based model selection criterion amounts to the minimization of the quantity:¹³

$$\text{AIC}_R = C \ln(1 - \alpha(E_r)) + 2N_R. \quad (7)$$

This minimizes the error in egg count as intuitively the complexity is proportional to the area of the segmented eggs.

3.3.2 Extracting segmented egg regions

First, connected components in the binary segmentation mask image are extracted and binary region image R , which has the same size as bounding box of the connected component is created. If area A_R of the region image (i.e., the number of pixels in the segmented egg region) is less than the defined area threshold A_r/h , then the detected region is classified as noise; otherwise, we solve for the optimal number of ellipses as described in the next section.

3.3.3 Initializing ellipses solutions

For defined number N_R of ellipses in a segmented egg region image R , we initialize the ellipse hypotheses using k -means clustering this defines a set E_R of clusters, which are circular in shape with hard cut-off borders where each pixel is strictly allocated to one cluster. The cluster centers are the mean vector of the points belonging to the respective cluster, while the diameters are the maximum Euclidian distances of the cluster members from their respective cluster centers.

3.3.4 Optimizing ellipses solutions

To obtain a more complex, ellipsoid shapes with soft cut-off borders (i.e., overlapping ellipses) which closely describes the shape of *S. haematobium* eggs, the ellipse hypotheses is evolved using the Gaussian mixture model expectation maximization (GMM-EM) algorithm to finetune the parameters of the initialized set E_R of clusters with the best coverage $\alpha(E_r)$ of the given segmented egg region. This is achieved by expectation-step and the maximization-step iteratively of the GMM-EM algorithm. The log likelihood function is maximized until the GMM-EM algorithm converges. A detailed explanation of the GMM-EM algorithm can be found in the work of Refs. 31 and 32.

3.3.5 Solving for the optimal number of ellipses

Different models (i.e., solutions involving different numbers N_R of ellipses) for a segmented egg image region are evaluated based on the AIC criterion [defined in Eq. (7)] that balances the trade-off between model complexity and approximation error. To minimize AIC_R , the refined segmentation algorithm increments the number of candidate ellipses N_R starting from a lower boundary, $N_l b = 0.6 \text{ceil}(C)$, with a step size of 1. At each value of N_R , the set E_R of clusters is first initialized by k -means clustering (described in Sec. 3.3.3) and then evolved using the GMM-EM algorithm (described in Sec. 3.3.4). This process continues until N_R is equal to the upper boundary, $N_u b = 1.1 \text{ceil}(C)$. In each iteration, the AIC_R criterion is computed. The lower and upper boundaries of the number of ellipses are formulated using the complexity measure C , derived from prior knowledge about the average pixel area of the *S. haematobium* eggs. This helps to reduce the search space for the optimal number of ellipses. From all possible models (involving from $N_l b$ to $N_u b$), the refined segmentation algorithm reports as the optimal solution as the set of ellipses E_R^* with the minimum AIC_R .

3.3.6 Morphological filtering of detected isolated eggs

To reduce these false positives diagnosis caused by pixels of artifacts, such as crystals wrongly segmented as isolated egg, we introduced two linear inequality constraints, which are functions of the area of the detected ellipse $|E_R|$ and overlap coefficient $\text{OC}(R, U_E)$ defined by the following ratio:

$$OC(R, U_E) = \frac{R \cap U_E}{\min(R, U_E)}. \quad (8)$$

These inequality constraints are derived experimentally and only applied to segmented regions with a single fitted ellipse for determination of diagnosis result. This improves the specificity of the algorithm by accepting only regions that fall within an experimentally defined boundary region as candidate *S. haematobium* eggs while discarding the others as a false positive prediction.

4 Dataset and Implementation Details

4.1 Dataset Description

A total of 103 captured urine samples were used for the creation of the SH dataset. The SH dataset was used for training and development testing of the DeepLabV3-MobileNetV3 deep neural network model, while a separate set of 65 captured urine samples referred to as diagnosis test dataset was used for testing the developed framework for urogenital schistosomiasis diagnosis. The images were captured from urine samples collected in a rural area in central Nigeria with the Schistoscope V5.0.¹¹ The size of the captured images is 1520×2028 pixels. The details of the sample collection and preparation process are described in our previous works.^{10,11} The procedure followed in capturing and annotation of the *S. haematobium* eggs in images is described in Sec. 3. A summary of the SH image dataset is shown in Table 1. It consists of 12,051 images of clinical urine samples and their respective mask images. There are 17,799 annotated *S. haematobium* eggs in 2997 captured FoV images. The dataset consists of images that are easy to identify eggs (Fig. 2) without the presence of artifacts in the background, as well as images that are difficult to analyse (Fig. 3) with backgrounds containing artifacts, such as crystals, glass debris, air bubbles, fabric fibres, and human hair, which makes egg identification difficult.

Table 1 Number of images per category in the SH dataset.

Split	Positive images	Negative images	Total
Training set (80%)	2420	7221	9641
Test set (20%)	577	1833	2410
Total	2997	9054	12,051

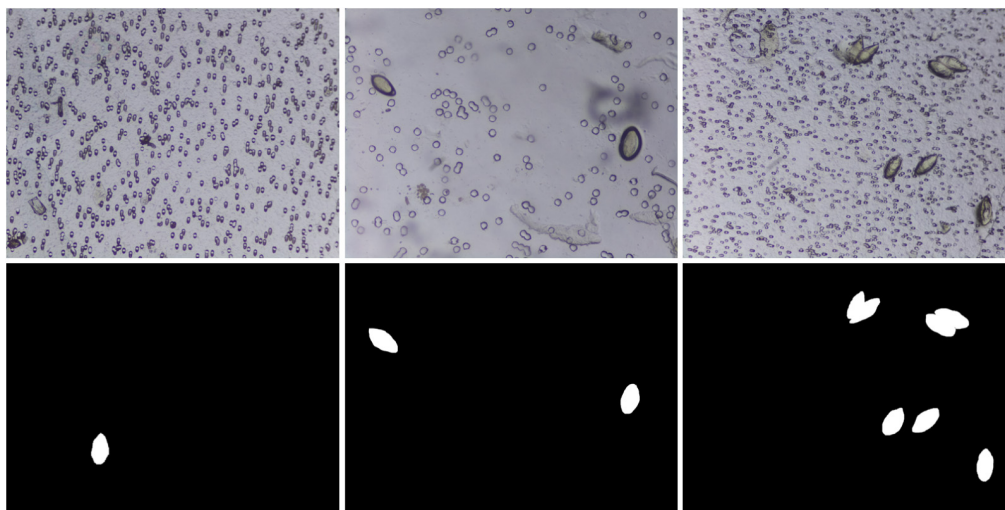


Fig. 2 Example of sample images that are easy-to-identify eggs having glass slides and filter membranes as background and their respective ground truth images.

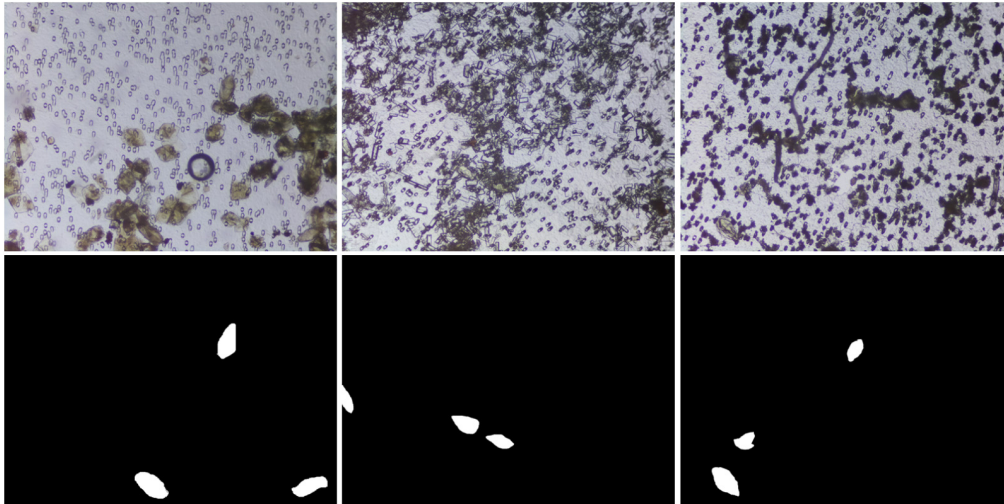


Fig. 3 Example of sample images that are difficult-to-identify eggs with artifacts, such as crystals, glass debris, air bubbles, and fabric fibers in the background, and their respective ground truth images.

The SH dataset is split into 80% (9641 images) and 20% (2410 images) for training and development testing, respectively. To our best knowledge, this SH dataset is the largest robust dataset focused on *S. haematobium* egg images captured in a field setting.

4.2 Implementation Details

The training of the DeepLabv3-MobileNetV3 model was performed using the Pytorch framework³³ on NVIDIA A100-SXM4-40GB GPU. All images were pre-processed by centring and normalizing the pixel density per channel. We fine-tuned the model for 100 epochs. The batch size is set to 8, and ADAM optimizer is used to optimize the Dice loss function, with an initial learning rate of 1×10^{-4} . We employ a “poly” learning rate policy,¹² where the initial learning rate is multiplied by $(1 - \text{iter}/(\text{max_iter}))^{\text{power}}$ with power = 0.9. All images were down sampled to 507×676 before being fed to the neural network.

To demonstrate the field applicability of the two-stage framework in low-resource settings, we performed the development testing and diagnosis testing on a Raspberry Pi 4 model B using a Coral USB accelerator. To perform semantic segmentation on the Edge AI system, we converted the DeepLabv3-MobileNetV3 model from Pytorch to TensorFlow lite.³⁴ This was done by first exporting the Pytorch model in Open Neural Network Exchange (ONNX) format. The ONNX model is then converted to TensorFlow before the final conversion from TensorFlow to TensorFlow Lite. The refined segmentation algorithm was implemented on the Raspberry Pi.

4.3 Evaluation Metrics

We evaluated the performance of semantic segmentation of *S. haematobium* egg by comparing the DeepLabV3-MobileNetV3 segmentation, which are the prediction results with a ground truth (GT) that was manually annotated by a trained parasitologist using the pixel accuracy (PA):

$$PA = \frac{1}{n} \sum_{i=0}^n 1_{(GT_i=DS_i)}. \quad (9)$$

We also compared the semantic segmentation performance using DSC and Jaccard similarity coefficient (JAC), which are widely used in evaluating medical segmentation algorithms:

$$DSC = 2 \frac{|GT \cap DS|}{|GT| + |DS|}, \quad (10)$$

$$JAC = \frac{|GT \cap DS|}{|GT \cup DS|}. \quad (11)$$

Although the diagnostic performance of our two-stage diagnosis framework was evaluated by employing three metrics, precision, sensitivity, and specificity, which are commonly used for evaluating diagnostic devices:

$$\text{precision} = \frac{TP}{TP + FP}, \quad (12)$$

$$\text{sensitivity} = \frac{TP}{(TP + FN)}, \quad (13)$$

$$\text{specificity} = \frac{TN}{(TN + FP)}, \quad (14)$$

where TP, FP, TN, and FN are true positive, false positive, true negative, and false negative samples, respectively.

5 Experiments and Results

5.1 DeepLabV3-MobileNetV3 S. haematobium Egg Semantic Segmentation

To determine the applicability of the framework on the Edge AI system in low-resource settings with no internet connectivity, we implemented and evaluated its performance on a Raspberry Pi 4B with Coral USB accelerator. We evaluate the DeepLabV3-MobileNetV3 deep learning model for the semantic segmentation of *S. haematobium* eggs using the development test dataset. As shown in Table 2, the deep learning model achieved a segmentation accuracy of 99.69%. It is, however, important to note the existence of a very high imbalance between the foreground and background pixels in the images, which could hamper the segmentation accuracy. While using the Jaccard and dice coefficient as performance metric, the model obtained 85.30% and 87.20%, respectively. However, the average inference time per image was 7.13 s with a model size of 7.13 MB. We considered the inference time too high given the need to process 117 images per sample diagnosis. To reduce the processing time on the Edge AI system, we optimized the DeepLabV3-MobileNetV3 deep learning model using post-training quantization on TensorFlow. The optimized model was applied to the development test dataset. We observed a significant reduction in inference time and model size (2× and 4×, respectively) with little effect (about 1% reduction) in the Jaccard and Dice coefficient metric. However, the segmentation accuracy remained the same. All further experiments in the work were carried out using the optimized model.

The visual performance of the segmentation model is shown in Fig. 4(c). We observed that the model detected eggs in images heavily cluttered with artefacts, such as crystals and other particles (sample image 3). It also detected highly transparent *S. haematobium* eggs (sample image 1) present in the captured images. Partially cut eggs on the edge of the images and overlapping eggs were also detected as observed in sample image 2. However, the boundaries in the overlapping eggs are not clearly segmented.

5.2 Refined Segmentation and Egg Count

In the second stage of our framework, we applied a refined segmentation algorithm on the output segmentation mask image of the DeepLabV3-MobileNetV3 deep learning model as described in Sec. 3.2. From Fig. 4(d), we observed that the refined segmentation steps fills-in eggs pixels missed in the deep learning semantic segmentation stage. This improves the visual perceptibility of the eggs in the segmentation mask image especially in regions with overlapping eggs as seen

Table 2 Performance of DeepLabV3-MobileNetV3 for semantic segmentation of *S. haematobium* eggs.

	PA	JAC	DSC	Model size (MB)	Inference time (s)
Base model	99.69	85.30	87.20	42.1	7.13
Optimized model	99.69	84.64	86.55	11.1	4.39

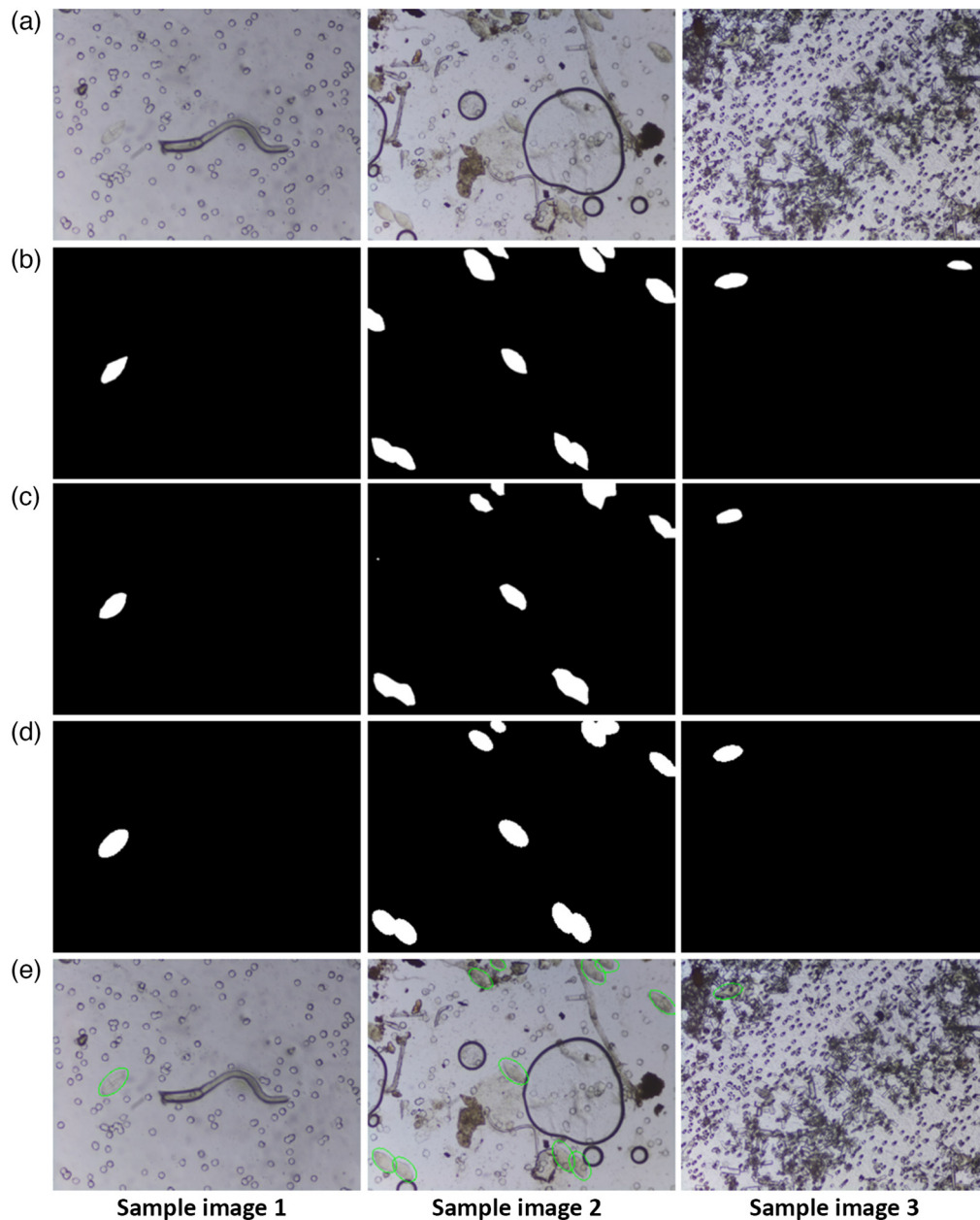


Fig. 4 Visual performance of developed framework on sample images from the dataset. Schistoscope (a) captured and (b) ground truth images. The output mask images of (c) DeepLabV3-MobileNetV3 segmentation and (d) refined segmentation. (e) The result image with detected eggs.

in sample image 2. Figure 5 shows example regions with overlapping eggs in the deep learning segmentation mask image. We observed that the correct number of eggs in Figs. 5(a) and 5(c) are equivalent to the optimal AIC criterion values in Figs. 5(b) and 5(d), respectively. The refined segmentation stage is able to separate overlapping eggs thus improving the accuracy of determining the infection intensity of the sample.

Figure 6 shows the scatter log-scale plots of the automated egg counts versus the manual egg count (i.e., egg count by an experienced microscopist) of samples in the diagnosis test dataset. Although we observed that the predicted egg counts were mostly under the 1:1 line, this signifies underprediction especially in samples with high egg counts. However, the manual and automated egg counts are highly correlated in samples with both low and high egg counts, which indicate the applicability for the developed framework in determining infection intensity of a sample.

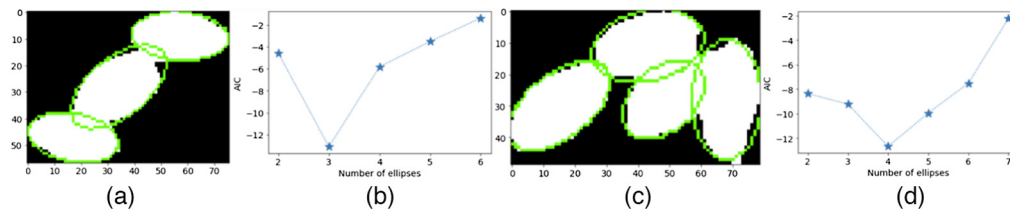


Fig. 5 (b), (d) Estimated AIC criteria values for different number of ellipses fitted on the (a), (c) region images with optimal ellipses are highlighted in green.

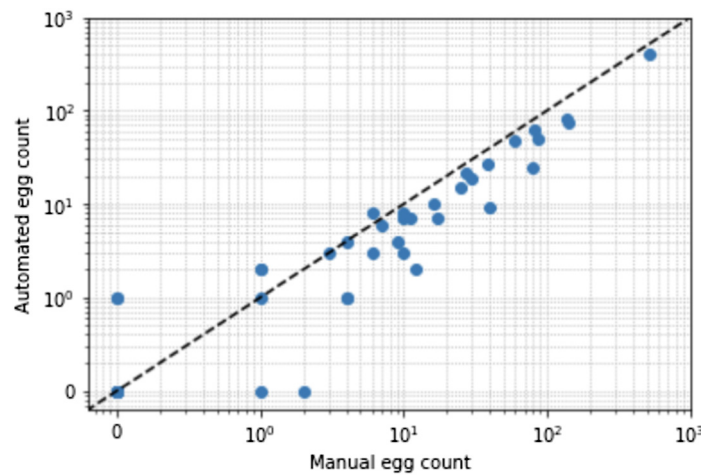


Fig. 6 Logarithmic scale scatter plot of infection intensity per 10 mL urine sample. The manual egg count obtained by a microscopist manually counting the eggs in the diagnosis test image dataset is used as a reference, whereas the automated egg count is obtained using the developed framework.

5.3 Urogenital Schistosomiasis Diagnosis

A 10 mL urine sample consists of 117 FoV images when filtered with a 13 mm membrane and captured by the Schistoscope. For a sample to be determined as true negative diagnosis, the 117 FoV images must contain no false positives. We experimentally defined boundary conditions for the detected isolated eggs using inequality functions, defined by the overlap and area of the fitted ellipse as shown in Fig. 7. The boundary conditions are defined by $OC \geq -4.75 \times 10^{-4}|E_R| + 1.15$ and $OC \leq 3.25 \times 10^{-4}|E_R| + 0.74$, where OC is the overlap and $|E_R|$ is the area of fitted ellipse. The experiment was carried out using images from the development test dataset

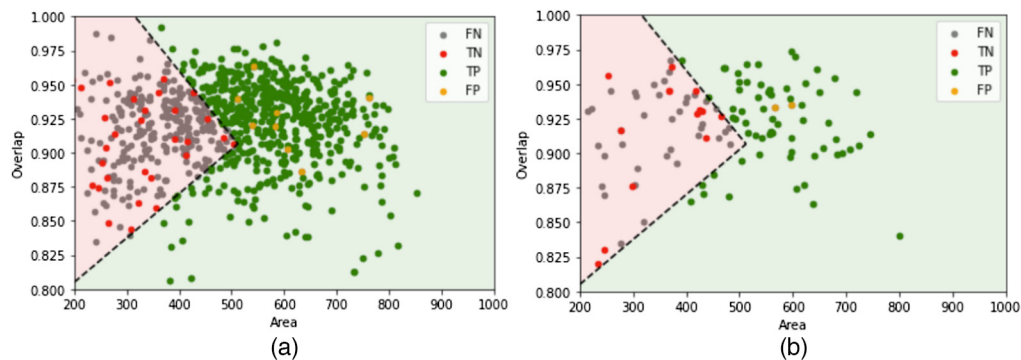


Fig. 7 Boundary conditions to determine a sample as positive or negative diagnosis applied to images from (a) the development test dataset and (b) diagnosis test dataset. Samples are determined as positive if an egg in any of the 177 FoV images is detected in the green region.

Table 3 Diagnostic performance of developed framework on the diagnosis test dataset.

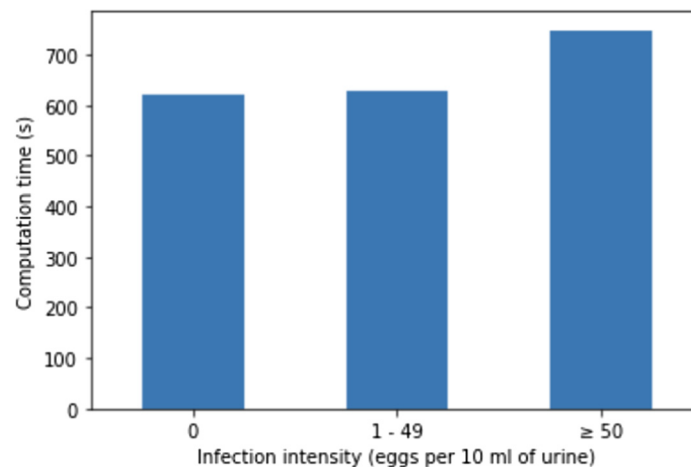
	Sensitivity	Specificity	Precision
Without boundary conditions	96.88	72.73	77.50
With boundary conditions	93.75	93.94	93.75

[Fig. 7(a)], and boundaries were found to hold also in images from the diagnosis test dataset [Fig. 7(b)]. A sample was determined as positive diagnosis if an isolated egg is detected in the set of 117 FoV images, which satisfies the defined boundary conditions (an egg is detected in the green region of Fig. 7). Otherwise, the sample is determined as negative diagnosis. We observed that most of the false negatives in Fig. 7 (gray markers) were broken or partly captured eggs found at the edges of the image, whereas the false positives (yellow markers) are artefacts that are very similar in appearance to a *S. haematobium* egg.

The diagnostic performance of the developed framework is shown in Table 3. We observed a significant improvement in diagnosis specificity (from 72.73% to 93.94%) and precision (from 77.50% to 93.75%) when the boundary conditions are applied in determining the sample diagnosis. However, a reduction in the diagnosis sensitivity is observed. This is due to some samples with very low infection intensities (eggs per 10 mL of urine ≤ 2) not having any detected eggs, which meet the boundary constraint represented by the green region of Fig. 7(b). McNemar's test returned a p -value of 0.008, which indicates a statistically significant difference between both methods (p -value < 0.05). Also we achieved a 7.39% and 92.11% performance improvement in diagnosis sensitivity and specificity, respectively, compared to our previously published work.¹¹

5.4 Computational Time

To evaluate the computational performance of the developed framework, we measured the computational time of both stages of the proposed method as function of infection intensity. *S. haematobium* infection intensity has consistently been characterized by the number of schistosome eggs per 10 mL of urine with 1 to 49 eggs per 10 mL of urine defining a light infection and more than 50 eggs per 10 mL of urine indicating a heavy infection.³⁵ We performed the running time experiments on a Raspberry PI 4B with Coral USB accelerator to study how the infection intensity affects the computational time. The algorithm was applied on images from the diagnosis test image dataset. Figure 8 shows the average computational time from the application of the first (DeeplabV3-MobileNetV3 semantic segmentation) and the second (refined segmentation and separation of overlapping eggs) stages of the developed framework to the diagnosis test image dataset as a function of the infection intensity. From this figure,

**Fig. 8** Average computational time in seconds from the application of framework on the diagnosis test image dataset.

it can be seen that there is little difference between the computational time of negative and light intensity samples (620 and 628 s, respectively). However, processing samples with heavy infection intensity is more time-consuming with an average computational time of 748 s.

6 Discussion

6.1 Impact on Schistosomiasis Control and Elimination

Schistosomiasis affects about 252 million people globally² with ~90% of infections and the vast majority of morbidity occurring in Sub-Saharan Africa. Chronic urogenital schistosomiasis infection can result in bladder fibrosis as well as female and male genital schistosomiasis, which is associated with greater risk of HIV transmission.³⁶ Also the bulk of the more than 1.6 million disability-adjusted life years³⁷ caused by schistosomiasis worldwide affect children, who have the highest prevalence and intensity of infections. Morbidity in children include anaemia, delays in physical and cognitive development, and reduced tolerance to exercise.³⁸ The main strategy for control of schistosomiasis focuses on mass drug administration (MDA) of praziquantel in priority to primary school-aged children because it is more cost-effective to treat all school-aged children in a community above a certain prevalence threshold than to test and treat each individual.⁴ On a population level, higher intensities of infection are associated with higher levels of morbidity, but these relationships are poorly defined, as most control programs monitor only prevalence of infection and not intensity.³⁹ Microscopic examination of urine samples is often a cheap and simple procedure recommended by WHO for the diagnosis of urogenital schistosomiasis. However, it has some critical shortcomings, which include access to microscopes and trained personnel as well as poor sensitivity and reproducibility, and an error-prone manual read-out.⁴⁰ This led to the recent formation of the WHO Diagnostic Technical Advisory Group with the mandate to identify and prioritize diagnostic needs and to subsequently develop TPPs for future diagnostics.^{4,41,42} The TPP requires new diagnostic tools to have high specificity so as reliably measure when prevalence is above or below a cut-off of 10% in school-aged children. This informs decision on the frequency of the MDA. A diagnostic tool with high specificity is also needed to track changes of prevalence, ensuring that MDA is reducing overall prevalence, and to determine if transmission has been interrupted. In this work, we developed a two-stage diagnostic framework, which is a suitable candidate for estimating infection intensity and diagnostic prevalence in urogenital schistosomiasis monitoring and control.

6.2 Limitations

- *Image auto-focusing.* Some of the images in the dataset captured by the Schistoscope were blurry due to sub-optimal autofocusing. Although this had no effect on the diagnostic performance, it did have an effect on the automated egg counts of a few samples in the diagnosis test dataset. This problem has been solved by a more accurate auto-focusing algorithm in subsequent version of the Schistoscope.
- *Annotation problem.* Annotating the exact boundaries of the eggs was difficult due to their small sizes. This may contribute to the difficulty of the model to segment the exact egg boundaries, especially in overlapping eggs.
- *Diagnostic prevalence.* The determining diagnosis of a sample with eggs that are either broken or are at the edges of the images is mostly not considered by the developed framework as they do not meet the boundary requirements. This increases the chances of a false negative diagnosis especially in samples with very low egg counts.
- *Computational time.* On a Raspberry Pi with Coral USB accelerator, the developed framework processes 117 images of the 13 mm urine filter membrane in ~11 min. Therefore, an estimated processing time of 35 min required to process a 25 mm filter membrane with 372 captured FoV images. However, the processing time can halved by the use of 2 Coral USB accelerators for computation through multi-threading.

7 Conclusion

We created a robust dataset of manually annotated *S. haematobium* eggs in microscopy images of urine samples collected from an endemic population, captured by the Schistoscope V5.0 device. We then developed a two-stage diagnosis framework for urogenital schistosomiasis using the SH

dataset. The framework consists of two main stages, the first step involves the semantic segmentation of the eggs using the DeepLabV3 deep learning architecture with a MobileNetV3 backbone. The model effectively segmented the transparent eggs having low contrast with the background, and it also differentiated between eggs and other urine artifacts, such as crystals that have egg-like structures. In the next stage, a refined segmentation algorithm was applied to detect and count the eggs present. The refined segmentation algorithm separates overlapping eggs by fitting the region image with an optimal number of ellipses determined by optimising the AIC criterion. For improved diagnostic performance, we determine a sample as positive only if there is a detected egg present in the sample images that meet a defined boundary requirement, which is a function of the overlap and area of the fitted ellipse. We implemented the developed framework on an Edge AI system consisting of a Raspberry Pi 4B with Coral USB accelerator and applied it to a diagnosis test dataset of 65 samples using results obtained by an expert microscopist as reference. We obtained 93.75%, 93.94%, and 93.75% sensitivity, specificity, and precision, respectively. The automated egg count was also highly correlated with the manual count of the microscopist. The framework also provides causality for its estimated egg counts, which is relevant for diagnosis. From our results, it is evident that our automated framework for urogenital diagnosis combined with the Schistoscope device is a promising diagnostic tool for schistosomiasis. In a future study, the proposed multilayer framework, combined with the Schistoscope, will be validated for the diagnosis of urogenital schistosomiasis by comparing its performance with conventional microscopy as well as more accurate diagnostic methods, such as schistosome circulating antigen detection and DNA-based methods, such as polymerase chain reaction assays.⁴³

Disclosures

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability Statement

Schistosoma haematobium image dataset is available from the Zenodo Repository: <https://doi.org/10.5281/zenodo.6467268>.

Acknowledgments

This work was funded under the INSPIRED (INclusive diagnoStics for Poverty RElated parasitic Diseases) Project (Grant No. W 07.30318.009) by WOTRO (the Science for Global Development department) of NWO (the Dutch Research Council). We would like to acknowledge the Neglected Tropical Disease Team from the Federal Ministry of Health Abuja and Federal Capital Territory Authority Public Health Laboratory, Abuja, Nigeria, and Delft University of Technology Global Initiative for their support toward this study. The study was conducted in accordance with the declaration of Helsinki and approved by the Federal Capital Territory Health Research Ethics Committee in Abuja, Nigeria (Reference No. FHREC/2019/01/73/18-07-19). Informed consent was obtained from all subjects involved in the study.

References

1. WHO, "Schistosomiasis and soil-transmitted helminthiases: treating millions of people, despite the pandemic," 2021, <https://www.who.int/news/item/08-12-2021-schistosomiasis-and-soil-transmitted-helminthiases-treating-millions-of-people-despite-the-pandemic> (accessed 12 May 2022).
2. O. P. Aula et al., "Schistosomiasis with a focus on Africa," *Trop. Med. Infect. Dis.* **6**(3), 109 (2021).
3. J. Utzinger et al., "New diagnostic tools in schistosomiasis," *Clin. Microbiol. Infect.* **21**(6), 529–542 (2015).
4. WHO, *Diagnostic Target Product Profiles for Monitoring, Evaluation and Surveillance of Schistosomiasis Control Programmes*, World Health Organization (2021).
5. N. Rajeshwori et al., "Prevalence of myopia and binocular vision dysfunctions in microscopists," *Int. Eye Sci.* **18**(7), 1180–1183 (2018).
6. J. Rajchgot et al., "Mobile-phone and handheld microscopy for neglected tropical diseases," *PLoS Negl. Trop. Dis.* **11**(7), e0005550 (2017).
7. M. A. Saeed and A. Jabbar, "Smart diagnosis" of parasitic diseases by use of smartphones," *J Clin. Microbiol.* **56**(1), e01469 (2018).

8. A. Vasiman, J. R. Stothard, and I. I. Bogoch, "Mobile phone devices and handheld microscopes as diagnostic platforms for malaria and neglected tropical diseases (NTDs) in low-resource settings: a systematic review, historical perspective and future outlook," *Adv. Parasitol.* **103**, 151–173 (2019).
9. H. Zhu et al., "Optical imaging techniques for point-of-care diagnostics," *Lab Chip* **13**(1), 51–67 (2013).
10. P. Oyibo et al., "Schistoscope: an automated microscope with artificial intelligence for detection of schistosoma haematobium eggs in resource-limited settings," *Micromachines* **13**, 643 (2022).
11. B. Meulah et al., "Performance evaluation of the Schistoscope 5.0 for (semi-) automated digital detection and quantification of schistosoma haematobium eggs in urine: a field-based study in Nigeria," *Am. J. Trop. Med. Hyg.* **107**(5), 1047–1054 (2022).
12. L. C. Chen et al., "Rethinking atrous convolution for semantic image segmentation," arXiv:1706.05587 (2017).
13. C. Panagiotakis and A. Argyros, "Region-based fitting of overlapping ellipses and its application to cells segmentation," *Image Vision Comput.* **93**, 103810 (2020).
14. Y. S. Yang et al., "Automatic identification of human helminth eggs on microscopic fecal specimens using digital image processing and an artificial neural network," *IEEE Trans. Biomed. Eng.* **48**(6), 718–730 (2001).
15. D. Avci and A. Varol, "An expert diagnosis system for classification of human parasite eggs based on multi-class SVM," *Expert Syst. Appl.* **36**(1), 43–48 (2009).
16. C.-C. Lee et al., "Helminth egg analysis platform (HEAP): an opened platform for microscopic helminth egg identification and quantification based on the integration of deep learning architectures," *J. Microbiol. Immunol. Infect.* **55**, 395–404 (2021).
17. K. M. Naing et al., "Automatic recognition of parasitic products in stool examination using object detection approach," *PeerJ Comput. Sci.* **8**, e1065 (2022).
18. G. Sengul, "Classification of parasite egg cells using gray level cooccurrence matrix and kNN," *Biome. Res.* **27**(3), 829–834 (2016).
19. M. O. K. Hassan and K. M. Al-Hity, "Computer-aided diagnosis of schistosomiasis: automated schistosoma egg detection," *J. Clin. Eng.* **37**(1), 29–34 (2012).
20. M. Armstrong et al., "Point-of-care sample preparation and automated quantitative detection of schistosoma haematobium using mobile phone microscopy," *Am. J. Trop. Med. Hyg.* **106**, 1442–1449 (2022).
21. T.-Y. Lin et al., "Focal loss for dense object detection," in *Proc. IEEE Int. Conf. Comput. Vision*, pp. 2980–2988 (2017).
22. A. G. Howard et al., "MobileNets: efficient convolutional neural networks for mobile vision applications," arXiv:1704.04861 (2017).
23. M. Tan, R. Pang, and Q. V. Le, "EfficientDet: scalable and efficient object detection," in *Proc. IEEE/CVF Conf. Comput. Vision and Pattern Recognit.*, pp. 10781–10790 (2020).
24. T.-Y. Lin et al., "Microsoft COCO: common objects in context," *Lect. Notes Comput. Sci.* **8693**, 740–755 (2014).
25. O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," *Lect. Notes Comput. Sci.* **9351**, 234–241 (2015).
26. J. Brooks, "Features," original-date: 2018-09-03T21:38:31Z (2021).
27. S. Minaee et al., "Deep-COVID: predicting COVID-19 from chest X-ray images using deep transfer learning," *Med. Image Anal.* **65**, 101794 (2020).
28. M. Everingham et al., "The Pascal visual object classes (VOC) challenge," *Int. J. Comput. Vision* **88**(2), 303–338 (2010).
29. E. Xie et al., "Segmenting transparent objects in the wild," *Lect. Notes Comput. Sci.* **12358**, 696–711 (2020).
30. H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control* **19**(6), 716–723 (1974).
31. C. M. Bishop and N. M. Nasrabadi, *Pattern Recognition and Machine Learning*, Vol. 4, Springer (2006).
32. T. M. Mitchell and T. M. Mitchell, *Machine Learning*, Vol. 1, McGraw-Hill, New York (1997).
33. A. Paszke et al., "Pytorch: an imperative style, high-performance deep learning library," in *Adv. Neural Inf. Process. Syst.* **32** (2019).
34. Sithu3, "TFLite Conversion," original-date: 2020-09-14T03:53:58Z (2022).
35. R. E. Wiegand et al., "Associations between infection intensity categories and morbidity prevalence in school-age children are much stronger for *Schistosoma haematobium* than for *S. mansoni*," *PLoS Negl. Trop. Dis.* **15**, e0009444 (2021).
36. UNAIDS, "No more neglect—female genital schistosomiasis and HIV—integrating sexual and reproductive health interventions to improve women's lives," 2019, https://www.unaids.org/en/resources/documents/2019/female_genital_schistosomiasis_and_hiv (accessed 23 December 2022).
37. WHO, "Global health estimates: leading causes of death," 2020, <https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/ghe-leading-causes-of-death> (accessed 23 December 2022).
38. D. G. Colley et al., "Human schistosomiasis," *Lancet* **383**(9936), 2253–2264 (2014).

39. R. E. Wiegand et al., “Control and elimination of schistosomiasis as a public health problem: thresholds fail to differentiate schistosomiasis morbidity prevalence in children,” *Open Forum Infect. Dis.* **8**, ofab179 (2021).
40. L. J. Stuyver and B. Levecke, “The role of diagnostic technologies to measure progress toward WHO 2030 targets for soil-transmitted helminth control programs,” *PLoS Negl. Trop. Dis.* **15**, e0009422 (2021).
41. WHO, “Report of the first meeting of the WHO Diagnostic Technical Advisory Group for Neglected Tropical Diseases Geneva, 30-31 October 2019,” 2020, <https://www.who.int/publications-detail-redirect/9789240003590> (accessed 23 December 2022).
42. A. A. Souza et al., “Diagnostics and the neglected tropical diseases roadmap: setting the agenda for 2030,” *Trans. R. Soc. Trop. Med. Hyg.* **115**, 129–135 (2021).
43. P. T. Hoekstra, G. J. van Dam, and L. van Lieshout, “Context-specific procedures for the diagnosis of human schistosomiasis – a mini review,” *Front. Trop. Dis.* **2**, 722438 (2021).

Prosper Oyibo received his bachelor’s degree in electrical engineering and his master’s degree in control engineering from Ahmadu Bello University Zaria in 2014 and 2017, respectively. He is a PhD candidate at Delft Center for Systems and Control of Delft University of Technology. His research focuses on the development of smart optical devices with artificial intelligence for diagnosis of schistosomiasis and soil transmitted helminth infections. He is a Delft Global Fellow.

Jan Carel Diehl is a full professor in design for inclusive sustainable healthcare at the Delft University of Technology. He is currently the co-lead of the NWO project INSPIRED – inclusive diagnostics for poverty related parasitic diseases in Nigeria and Gabon and the NWA-ORC project for sustainable hospitals. His main research aim is to develop design interventions towards a more inclusive and sustainable healthcare system.

Tope Agbana holds a BEng in electronics engineering; an MSc in microwave sensing, signals, and systems; and a PhD in biomedical engineering. Presently, Agbana serves as a postdoctoral researcher at TU Delft and holds the position of CEO at AiDx Medical BV in the Netherlands. He is deeply committed to advancing AI-based optical diagnostic techniques for field deployment, particularly to address diagnostic gaps in resource-limited settings.

Biographies of the other authors are not available.