

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.e-jmii.com

Original Article

Helminth egg analysis platform (HEAP): An opened platform for microscopic helminth egg identification and quantification based on the integration of deep learning architectures

Chi-Ching Lee ^{a,b,c}, Po-Jung Huang ^{b,d}, Yuan-Ming Yeh ^b,
Pei-Hsuan Li ^a, Cheng-Hsun Chiu ^{b,e}, Wei-Hung Cheng ^{f,g,**},
Petrus Tang ^{e,f,*}



^a Department of Computer Science and Information Engineering, Chang Gung University, Taoyuan, Taiwan

^b Genomic Medicine Core Laboratory, Chang Gung Memorial Hospital, Linkou, Taiwan

^c Artificial Intelligence Research Center, Chang Gung University, Taoyuan, Taiwan

^d Department of Biomedical Sciences, Chang Gung University, Taoyuan, Taiwan

^e Molecular Infectious Disease Research Center, Chang Gung Memorial Hospital, Linkou, Taiwan

^f Department of Parasitology, College of Medicine, Chang Gung University, Taoyuan, Taiwan

^g Department of Medical Laboratory Science, College of Medicine, I-Shou University, Kaohsiung City, Taiwan

Received 28 September 2020; received in revised form 26 June 2021; accepted 18 July 2021

Available online 2 September 2021

KEYWORDS

Helminth;
Parasite egg
examination;
Deep learning;
Object detection;
Web server;
Database

Abstract *Background:* Millions of people throughout the world suffer from parasite infections. Traditionally, technicians use manual eye inspection of microscopic specimens to perform a parasite examination. However, manual operations have limitations that hinder the ability to obtain precise egg counts and cause inefficient identification of infected parasites on co-infections. The technician requirements for handling a large number of microscopic examinations in countries that have limited medical resources are substantial. We developed the helminth egg analysis platform (HEAP) as a user-friendly microscopic helminth eggs identification and quantification platform to assist medical technicians during parasite infection examination.

* Corresponding author. Molecular Infectious Disease Research Center, Chang Gung Memorial Hospital, Linkou, Taiwan.

** Corresponding author. Department of Parasitology, College of Medicine, Chang Gung University, Taoyuan, Taiwan.

E-mail addresses: chichinglee@cgu.edu.tw (C.-C. Lee), pjhuang@gap.cgu.edu.tw (P.-J. Huang), ymyeh@cgmh.org.tw (Y.-M. Yeh), M0829002@cgu.edu.tw (P.-H. Li), chchiu@adm.cgmh.org.tw (C.-H. Chiu), whcheng@isu.edu.tw (W.-H. Cheng), petang@mail.cgu.edu.tw (P. Tang).

<https://doi.org/10.1016/j.jmii.2021.07.014>

1684-1182/Copyright © 2021, Taiwan Society of Microbiology. Published by Elsevier Taiwan LLC. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Methods: Multiple deep learning strategies including SSD (Single Shot MultiBox Detector), U-net, and Faster R-CNN (Faster Region-based Convolutional Neural Network) are integrated to identify the same specimen allowing users to choose the best predictions. An image binning and egg-in-edge algorithm based on pixel density detection was developed to increase the performance. Computers with different operation systems can be gathered to lower the computation time using our easy-to-deploy software architecture.

Results: A user-friendly interface is provided to substantially increase the efficiency of manual validation. To adapt to low-cost computers, we architected a distributed computing structure with high flexibilities.

Conclusions: HEAP serves not only as a prediction service provider but also as a parasitic egg database of microscopic helminth egg image collection, labeling data and pretrained models. All images and labeling resources are free and accessible at <http://heap.cgu.edu.tw>. HEAP can also be an ideal education and training resource for helminth egg examination.

Copyright © 2021, Taiwan Society of Microbiology. Published by Elsevier Taiwan LLC. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Introduction

More than 800 million people globally have been threatened by helminth infection.¹ Medically important nematodes, trematodes, and cestodes are prevalent among humans in developing countries, including sub-Saharan Africa, South America, and East Asia.^{2–4} These diseases displayed a broad range of clinical manifestations from asymptomatic to malnutrition, anemia, and even cancer.^{5–8} Although these infections rarely cause death, the elimination of pathogenic helminths offers substantial advantages of socioeconomics. Currently, microscopic examination of helminth eggs in fecal samples is the gold standard for clinical diagnosis.^{9–11} A well-trained and experienced technician is responsible for the examination; however, the technician must take courses and comply with retraining protocols to establish and sustain professional skills.

Diagnosis of helminth infections is usually a qualitative analysis. However, eggs per gram (EPG) in fecal samples are the basis to estimate the burden of a parasitic disease, which requires quantitative examination.⁴ Multiple or mixed helminths infection in children is another public health issue in some endemic areas; thus, diagnosis should be more carefully performed to prevent false-negative results.^{12,13} The quantitative and time-consuming determination has further increased the error rate. Therefore, an automatic identification system for helminth eggs would be beneficial in the fields of diagnostics and epidemiology.

Identification and classification of microscopic images of helminth eggs based on computational algorithms have been proven to be workable. A pioneering study was performed by Yang et al., in 2001. This work employed an artificial neural network (ANN) classifier to recognize parasite eggs by shape, shell smoothness, and size in microscope images.¹⁴ Two subsequent studies established systems based on different filtering strategies on objects before their classification by the k-Nearest Neighbors (kNN) algorithm and profiles of grayscale-coloring spaces.^{15,16} A mathematical algorithm proposed that the size, thickness, brightness distribution, circularity, and shapes of eggs were employed as criteria for parasites discrimination.¹⁷

As artificial intelligence (AI) has flourished in recent years, deep-learning architectures have provided a faster and more comprehensive viewpoint. The preassigned criteria in AI modules were the basis for annotating features in images. Methods based on object detection and segmentation is widely used in image analysis and labeling flourished such as SSD,¹⁸ U-net¹⁹ and Faster R-CNN.²⁰ The advanced deep-learning approaches prompted us to develop a faster and automatic helminth eggs identification and quantification platform. Most importantly, the results analyzed by AI are reproducible.

From a practical point of view, a smooth user interface is a critical component for clinical examination in a real-world field domain. A comprehensive workflow, including microscopic image digitalization, computational analysis, and manual validation of post-AI prediction, should be well polished. Notably, previous helminth egg classification systems were designed to recognize individual egg images pre-captured by professional clinical operators.^{14–16} The automatic scanning and classifying platform for clinical samples remains an unachieved goal. Instead of the high cost of complex robotic mechanisms for parasitic egg target finding with a microscope, capturing all areas of the target specimen and performing the target finding *in silico* can be a low-cost and feasible solution. Optical microscopes that are equipped with computer controllable specimen slides are established in most examination facilities.

This work was aimed to build an integrative and interactive cloud-based tool—HEAP (Helminth egg analysis platform)—that provides a total solution for medical technicians in helminth infections diagnosis. HEAP is composed of image processing flows, multiple AI prediction kernels, and highly efficient manual validation interfaces. HEAP serves as a community for helminth images collection and AI prediction model providers. Users can establish their prediction systems by applying pretrained AI models and high-quality microscopic images that are downloaded from HEAP archives. This online tool offers an option for reducing the time required for diagnosis and epidemiological investigations. Because of the well-designed user interface, which is similar to real-world microscopic operation with enhanced target

identification facilitated by AI, HEAP can also serve as an educational tool for medical technician training.

Methods

Image capturing and labeling

The Olympus BX53 microscope is used to capture all helminth egg images. The microscope controlling software is CellSens Dimension which is the official build-in motorized stage and digital camera controlling software. The microscope light source is set to the max level. The exposure time is set to auto exposure. The objective lens is 10X, and the eyepiece is 10X. The images with the same regions were captured from several depth of fields (Fig. 1). This is an essential feature of HEAP to simulate the experience of microscopic observation. For example, in the same specimen slide, we captured 12 images that are the same X and Y position but different depth of fields from 140, 150, 160 ... 250 μm . To ensure that at least one clear focused-on-target image is captured. After the entire specimen is captured, all images within the same depth of field are merged into one piece. The image size is approximately 15 \times 15 thousand pixels. Labellmg which is a widely used image labeling tool is used to create all the targets of eggs for training data preparation. Few image adjustments, such as exposure improving, contrast increasing, and color temperature balancing was performed for the captured images that have low qualities. Table 1 lists the number of helminth eggs of each species used for AI model building.

Image binning and algorithm for egg-in-edge detection

The size of the raw images from the capture system is approximately 10 thousand square pixels, which is too large for the system to use. HEAP splits a large image into hundreds of smaller-sized images that are proper for machine learning architecture usage. In our testing results, the prediction accuracies are substantially lower than those for

Table 1 Number of helminth eggs used for AI model training.

Parasite egg category	Number of eggs used or examined
<i>Ascaris lumbricoides</i>	2089
<i>Clonorchis sinensis</i>	674
<i>Diphyllobothrium latum</i>	334
<i>Echinococcus granulosus</i>	732
<i>Echinostoma</i> spp.	88
<i>Enterobius vermicularis</i>	1118
<i>Fasciola hepatica</i>	850
<i>Fasciolopsis buski</i>	112
Hookworm	80
<i>Hymenolepis diminuta</i>	285
<i>Hymenolepis nana</i>	84
<i>Paragonimus westermani</i>	100
<i>Schistosoma haematobium</i>	100
<i>Schistosoma japonicum</i>	256
<i>Schistosoma mansoni</i>	80
<i>Toxocara canis</i>	56
<i>Trichuris trichiura</i>	1444

intact eggs since the eggs are located in the middle of two binning boxes. In the first step, a sliding window walks through from the upper-left side to the bottom-right side and retrieves the smaller images. In every step, an image adjustment function that converts an image into high contrast and grayscale mode is activated. Ten pixels around the edges of bin boxes are considered. We calculate the pixel density ratio of four edges to decide in which direction to move. The binning box moves 50 percent of the binning box width/height in the highest density edge direction. The edge detection is activated every movement. In the prediction stage, all binning boxes that contain the eggs in the edge are reserved and predicted by the AI model. After predicting all binning boxes, HEAP merges all predicted binning boxes based on the X and Y coordination

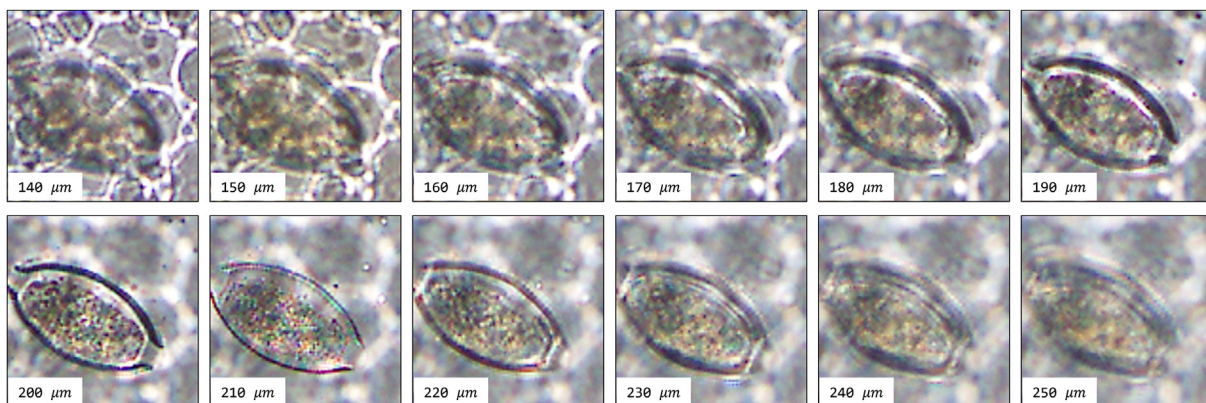


Figure 1. Image from different depth of fields. Each image was captured every 10 μm of the depth of field from 140 to 250 μm of the *Trichuris trichiura* eggs. The detailed pattern of three dimensions of egg shape can only be recognized on some depth of fields. The multiple microscopes within depth of fields extends the AI prediction range. The depth of field is labeled in the left-bottom side of each image.

relative to the raw microscopic image. Fig. 2 shows an example to illustrate the concept. The edge movement process is stopped while immense spots that cover all four edges of binning boxes are detected.

Distributed architecture for GPU-assistant calculating

We built a distributed computing architecture implemented by two core complements: the prediction module and the queuing system. The prediction module is implemented by TensorFlow with GPU (Graphics Processing Unit) and Python codes for image preprocessing. The queuing system is powered by an Apache Web environment and PHP (PHP: Hypertext Preprocessor) scripts, which can be applied to the multiple computing nodes, which can be used with or without GPU accelerators. In our benchmarking test (Table 2), the GPU accelerator can increase the prediction speed by approximately 100X compared with a CPU (by using a PC workstation with Intel Core i5 CPU, 64 GB system memory, and a single nVIDIA Titan Xp GPU graphic card). The system architecture is shown in Fig. 3. This queuing system is based on a shared file system that can be easily implemented on a Linux, Mac or Windows operating system or even mix operating systems (computing node and job scheduler that have different operating systems).

Results

Pretrained models and contents

For the demonstration, we have pretrained and built prediction models of 17 parasite egg samples (*Ascaris lumbricoides*, *Clonorchis sinensis*, *Diphyllobothrium latum*, *Taenia/Echinococcus granulosus*, *Echinostoma* spp., *Enterobius vermicularis*, *Fasciola hepatica*, *Fasciolopsis buski*, *Hookworm*, *Hymenolepis diminuta*, *Hymenolepis nana*, *Paragonimus westermani*, *Schistosoma haematobium*, *Schistosoma japonicum*, *Schistosoma mansoni*, *Toxocara canis*, and *Trichuris trichiura*) applied on AI models with different parameters. The number of the collected models and images were continuously growing. HEAP is an open-platform allowing users to donate their microscopic images and pre-trained models. All pretrained models can be selected and applied to the HEAP web page. The training

images in HEAP can be downloaded for relabeling or combined with other images to generate more comprehensive AI models.

Microscopic images with multiple depth of fields

Medical technicians screened candidates and magnified images to observe the detailed patterns of the target. Compared with traditional manual inspection, HEAP uses an opposite strategy—referred to as capturing before identifying—that captures all images within operational depth of fields and identifies by AI models. A microscope uses approximately 100 to 400X to magnify the target in the specimen. With this magnification rate, three-dimensional structures of helminth eggs varied in different depth of fields. A digital microscopic image capture system has focusing abilities to automatically choose the depth of field on a specific point. However, the eggs are located in all possible positions of the specimen. Autofocusing is not an efficient strategy. To overcome this focusing issue, we capture all images of operational depth of fields. HEAP enables the uploading of multiple images. The image viewing module overlaps all microscopic images within different depth of fields. For example, an egg candidate is identified in one depth of field; the bounding box displays all regardless of which depth of field user is selected. Fig. 1 shows the parasite egg and image qualities in different depth of fields.

On-the-fly analytic workflow

HEAP offers a front-end solution for microscope examination in our cloud. The distributed calculating architecture that allows multiple users submit multiple prediction queries is illustrated in Fig. 3. Users can create an account that has cloud space for microscopic images submission and computing resources for parasite egg identification. Images with different depth of fields can be uploaded in the single space. Currently, HEAP offers a direct file uploading and share URLs, such as Google Drive, OneDrive, and Dropbox URL links, to gather the microscopic images. Users can initiate multiple predictions using different architecture and set different parameters of the same architecture. Typically, an egg candidate prediction using one AI architecture requires approximately 1 min. All predicted results can be downloaded by a flattened image that contains the labeled egg candidates. The web interface of HEAP is shown in Fig. 4.

Integration of multiple AI architectures

HEAP offers users to import their interested parasite egg samples, including microscopy raw images and labeled results, to build a prediction model using our backend training pipelines. The morphologies of eggs are diverse. For example, the eggs of *T. trichiura* have an oval shape with clear borderlines, and the eggs of *C. sinensis* have an oval shape with a visible hook-like protrusion. HEAP equipped three most popular and extensively employed deep learning architectures: Single Shot Detector (SSD), U-net and Faster R-CNN. As an integrated platform, HEAP fulfills the needs of

Table 2 Benchmark of prediction with/without GPU accelerator.

	GPU (seconds)	CPU (seconds)	Speed ratio
U-net	15.8 ± 0.2	662.6 ± 1.0	~41.9X
SSD	30.8 ± 1.4	886.4 ± 0.4	~28.7X
Faster-RCNN	57.9 ± 0.3	5616.0 ± 48.3	~97.0X

The benchmark is achieved by prediction of *Trichuris trichiura* raw images (1152 images for SSD, 392 images for U-net and 1152 images for Faster R-CNN). The total calculation time was considered, including image binning, prediction, and result merging. The testing platform is the same computer using Tensorflow CPU/GPU version. The GPU is nVIDIA Titan Xp, the CPU is Intel Core-i5 7600 and the system memory is 64 GB.

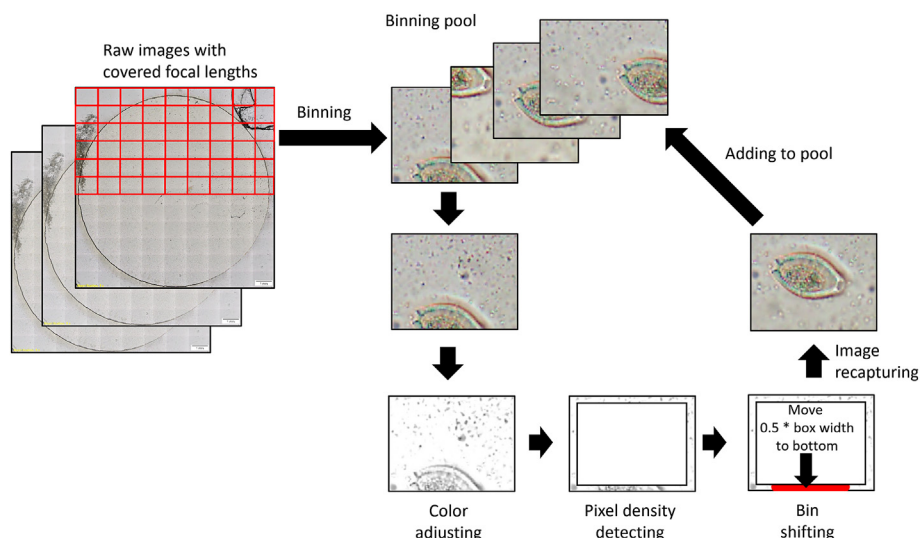


Figure 2. Workflow of image binning and egg-in-edge detection algorithm. To address the vast size of a raw image, image binning is performed using a fixed-sized nonoverlapping sliding window. Each binning image has two destinations: the first destination is part of the binning pool for further egg identification; the second destination triggers eggs in the edge detection protocol, which is based on comparing the pixel densities of four edges of binning boxes. Color adjustments, including grayscale converting and contrast increasing, are performed to emphasize the differences in the backgrounds and egg boundaries. The binning box moves toward the highest pixel-dense direction half-width/height of the binning box. The recaptured images are joined in the binning pool for egg prediction.

possible usage, including the object detection and segmentation method. Users can select SSD, U-net and/or Faster R-CNN as a default search engine(s) for helminth egg identification. HEAP supports the integration of predictions of multiple architectures based on voting from multiple prediction results, including different architectures and detailed programming parameters. Users can set a cut-off value. For instance, egg candidates can be reported by at least two predicting results that are supported.

High efficiency and user-friendly manual validation and confirmation interface

Although HEAP supports voting from multiple object detection methods and variant prediction parameters, human validation represents an unreplaceable rule for parasite identification. In traditional human eye inspection, a well-trained technician can rapidly identify eggs in seconds. To take advantage of human inspection, we designed a simple manual validation and confirmation interface that offers two validation modes: the map view and the list view.

Map view delivers an entire microscopic image with the identified parasite eggs, which are highlighted by the colored square boxes shown in Fig. 5. The results generated from images with different depth of fields or prediction settings were flattened in coordinates within a single microscopic image that were identical to those of the background. Filters such as the confident score can be applied during the validation process to narrow the number of prediction results. The bounding boxes highlighting the candidate eggs display within two styles: the dotted border (represents the software identified potential eggs) and the solid border (shows the human confirmed targets). The

human validation is a user-friendly operation in HEAP that can be achieved by single mouse clicking (a left-mouse click for confirmation; a right-mouse click for the removal of the prediction). An alternative way to perform the validation is referred to as list view, which shows all identified egg candidates in a thumbnail view. Technicians can quickly confirm all AI identified eggs using an efficient way.

The map view is designed to offer the same experiences for microscope usage. In the validation process, technicians can view all microscopic images with different depth of fields, which helps to reduce the false-negative cases (the missing cases of AI predictions can be manually identified). Conversely, the list view delivers all AI identified cases, which enables technicians to remove false-positive cases. Depending on which role of HEAP will be applied, false positive and false negative cases represent the different indexes to evaluate the performance. For example, removing false-positive cases is important in egg counting problems and reducing the false-negative rate is useful in egg identification and co-infection detection problems.

Open platform and community for helminth egg detection

HEAP enables the uploading of raw microscopic images within different depth of fields. Users can use the uploader page of the HEAP website to upload images. Labeled training data of bounding boxes recorded in XML format or pixel-wise masked images can be uploaded into HEAP for model building. Users can also upload their prediction models without uploading their microscopic raw images. With user-friendly interfaces, technicians with non-informatics backgrounds can generate their model by the HEAP backend pipeline. The workflow is shown in Fig. 6. It

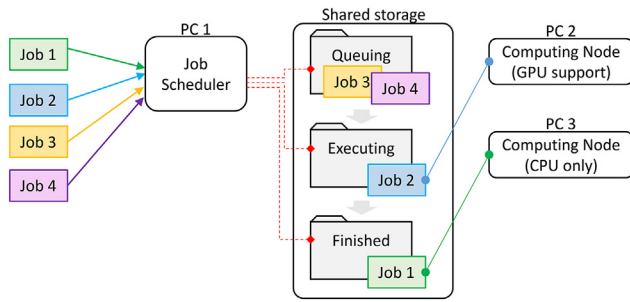


Figure 3. Distributed calculating architecture. This architecture is composed of three elements: the job scheduler (PC1), a shared storage space that is typically shared by a network-attached storage (NAS) protocol and several computing nodes (PC2 and PC3). In this figure, a three-PC system is used to demonstrate the concept. The shared storage can be either an independent device or a simple share folder of the job scheduler node. The job scheduler captured the requests, and then the calculating parameters and models selected by users were passed into a folder named Queuing, which collected all incoming requests. Each computing node has a scheduled script that is triggered every 1 min. Once the script is triggered, it finds the earliest request in the Queuing folder and moves the requested job into the Executing folder. The calculation is initialized. While the calculation is performed, the job is moved into a folder named Finished. The job scheduler node reports the results to users who submitted the job. The environment of each computing node can differ, for instance, different operating systems and different hardware configurations with/without GPU accelerators. A heartbeat subroutine is triggered every few minutes in the job scheduler that is used to detect the status of each computing node. If the computing node crashed or ran out of response time, the job status will be re-initialized and placed in a Queuing folder to wait for another computing node to execute. The red dot lines represent the heartbeat job status checker.

is not easy for a small team to collect all the images of microorganisms. The system requires growing and updating the new AI models and photos. We believe the only way to keep this platform alive is to assemble a community. We define the functional roles of the community by six levels, including: (1) End-user for service usages. (2) Image donor for licensed microscopic image donation. (3) Labeling volunteer providing image labeling. (4) Model donor offers pre-trained AI models. (5) Validation committee keeping the images and labeling qualified. (6) The housekeeper is operating the platform. The schema of the community is shown in Fig. 7.

To keep the community's continuously growth, the basement platform for community usages will be built in the third year, including the user identification system, user role management system, and accuracy evaluation strategies. Only the AI models generated by QC passed high-quality microscopic images are used to offer the services. For labeled images donated by volunteers, the validation committee uses kappa (K coefficient of agreement) to validate the accuracy of each volunteer. P_0 is the accuracy of the labeling; and P_c is the hypothetical probability of

accuracy. A criterion of kappa is set to 0.7 to guarantee the labeling accuracy is acceptable based on the same dataset of each validation volunteer.

$$\kappa = \frac{P_0 - P_c}{1 - P_c}$$

Discussion

A microscopic eye examination is widely used in almost all hospital and inspection laboratories for helminth eggs identification. This type of examination is common in medical examination, such as foreign worker entry inspection. A well-trained medical examiner can identify a single specimen in seconds. However, the limitation of manual inspection exists and cannot be easily overcome. For example, some specimen infected with more than one parasite will extend the manual examination time and reduce the accuracy.

Helminth egg counting (Egg per gram)

A fecal egg count (FEC) is a useful parasite surveillance technique for parasite control and treatment. The count be applied to either human hosts or farm hosts, such as pigs and calves. Eggs per gram (EPG) is used to quantize the parasitological infection. For example, EPG is a primary diagnostic method for schistosomiasis infections by counting the eggs in the feces. Precisely obtaining the EPG value is traditionally a time-consuming task because a technician has to view all specimen slides using a microscope. The viewing area of a microscope is considerably smaller than the area of a specimen slide. Technicians have to constantly move the microscope stage to change the viewing area. HEAP uses an opposite method for capturing all images, including different depth of fields. After all the sample preparations, the specimen slides are scanned and digitalized by an imaging capturing microscope with motorized stage. The egg identifier scans an entire microscopic image. The egg counting operation can be easily and precisely achieved. The workflow used to assistant egg counting is illustrated in Fig. 7.

Parasite co-infection detections and infected ratio estimations

Manual eye-inspection of parasite co-infection is inefficient because technicians have to examine entire specimens. Like egg counting problems, the viewable area of a microscope is considerably smaller than the specimen area, and horizontal movements and focal adjustments are frequently necessary. Most technicians use a lower magnificent rate to compensate for the situation in which possible helminth eggs may be missing some targets. The parasite-infected ratio estimates the relays on accurate counting of the eggs of all possible parasites on specimens. HEAP can be a perfect solution for co-infection detection because all observable regions and operational depth of fields are scanned by HEAP's predictor. The image capturing before identification coupled with high efficiency and user-

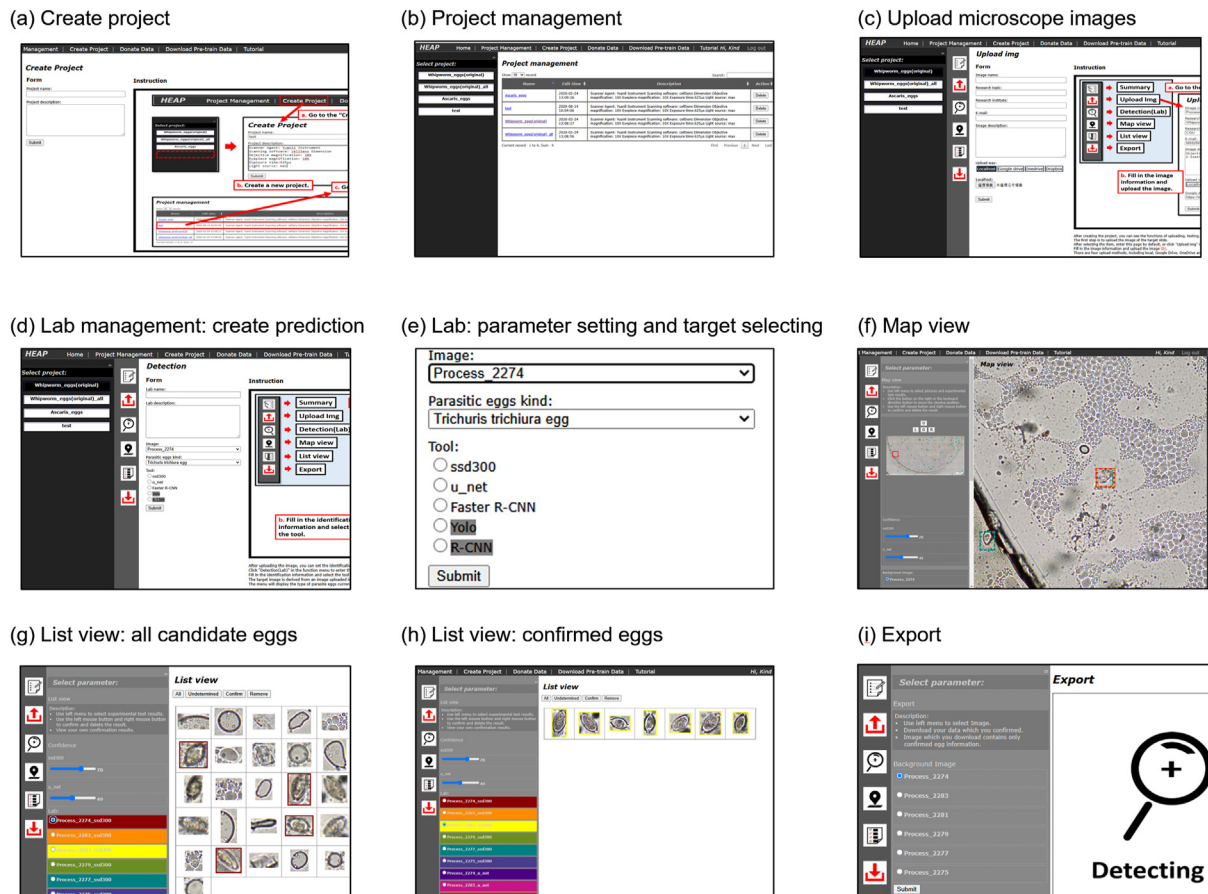


Figure 4. Web interface of HEAP. A web interface for the microscopic image analytic workflow of HEAP is user-friendly and efficient. Project page enables users to (a) create projects, and (b) manage projects. (c) Uploading page accepts microscopic images with multiple depth of fields. (d) Users can create and remove prediction labs. (e) AI architectures and parameters can be applied in different labs. (f) Map view is an integrative interface for prediction results examination. Eggs identified by different labs are labeled in different colors. Users can use mouse clicks to confirm and remove the predicted egg candidates. The mouse wheel can control the images with different depth of fields displayed in the background, which provides a visual assistant for a technician to perform an eye inspection. (g) List view summarizes all cropped identified eggs. Criteria filter and results from different models can be selected in panel of the left-hand side. (h) Tabs on the top of the list view page are used to group the confirmed and unvalidated eggs images, which helps users to efficiently locate images. (i) Results can be downloaded in flattened microscopic images with labeling eggs.

friendly human validation mechanisms guarantee the identification reliability.

Future implementations of helminth identifications protocol

Instead of replacing current manual eye inspection, the designing goal of HEAP served as a microscopic examination assistant system by providing microscope experience coupled with automatically egg identification and quantification abilities. HEAP can be the first step of candidate identifier helping to process a massive amount of microscope examining images per day. Technicians can further validate egg candidates using the list view of HEAP's web page. The manually confirmed egg images can serve as a training dataset to rebuild the AI models. We believe automatic egg identification and an upgradeable recognition kernel can reduce the workload of medical technicians

on the front line. The egg identification AI model can be improved by continuously integrating an increasing number of human validated samples. The types of recognizable helminth eggs are expanded with the growth of the HEAP user's community.

Microscopy observation without a microscope, an ideal tool for education of parasite egg examination and microbiology.

HEAP enables the uploading of multiple microscopic images within different depth of fields. Map view mimics complete microscope experiences, and users can obtain a magnified view field and change the viewing area in a 3-axes approach - horizontal and vertical directions. We believe that the microscopic experiences not only improve productivity but also serve as an ideal tool for educational purposes. Traditional educational material for helminth eggs or microbes

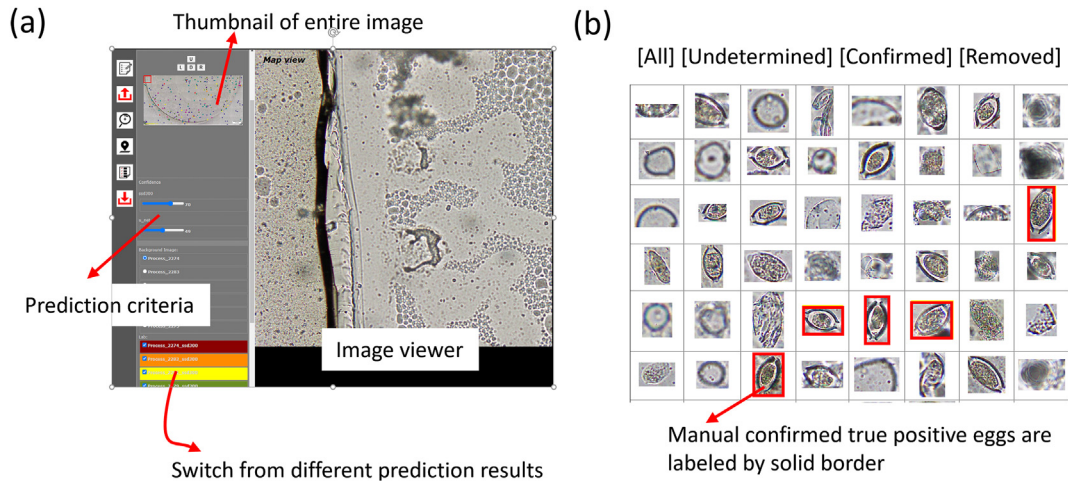


Figure 5. Manual eye-inspection system. (a) Map view offers virtual microscope experiences. Users can move the image horizontally (mimic the X and Y-axis movement) and change images with different depth of fields (Z-axis movement). Because all images are pre-captured, changing the view field from one position to another is considerably easier and smoother than physical microscopic operations. In this viewpoint, technicians can search for false-negative cases, which were not identified by AI. The right-hand side panel shows a small thumbnail that represents the original large-sized microscopic image. The identified eggs are highlighted by a colored dot in the thumbnail view. The left-hand side panel delivered all prediction results and filter parameters, and users can select single prediction or multiple results by a checkbox in the left-hand side list of labs. The prediction results were displayed using a colored box in the middle panel. This interface enables multiple results to be displayed in the same view by a different color. The AI identified un-validated egg candidates, which were displayed within a dotted box. Technicians can click the left mouse button to confirm the positive case and the right mouse button to confirm the false-positive cases. All validated data and operations are synchronized into the HEAP cloud in the background using AJAX. Scrolling up/down with the mouse-wheel can mimic the Z-axis movement, which enables users to change microscopic images with different depth of fields. (b) List view displays all identified egg candidates without other regions. In this view, users can quickly identify false-positive cases. With the same operation behavior as the map view, left-mouse clicking on the egg photo is defined as the true positive confirmation, and right-mouse clicking is the false-positive confirmation. In the top side, four tabs (All, Undetermined, Confirmed, and Removed) help users to quickly switch from cases with/without human validations. In the "All" tab, the validation status (confirmed positive or negative) is displayed using a small icon on the top-right-hand side of each egg target.

provides clear and precise microscope photos. However, in a microscopic examination scenario, the images are not always precisely focused. In training courses, students must learn how to identify targets at a low magnified rate and identify the egg candidates in missing focused images. Instructors and teaching assistants select a single egg under a microscope. HEAP offers not only high-quality microscopic images but also a comprehensive microscope experience. Additionally, the interface enables users to select identified eggs by simple clicking, which is more straightforward than manually searching other objects under a microscope. To train a professional medical examination technician requires

years. Numerous microscope photos must be familiar during the training program. HEAP provides high-quality images and nearly microscopic operating experiences that considerably reduce the cost of training.

The future of microscopic examination

HEAP is a growing platform; in the future, while there are more and more users, more and more microscopic images and prediction models are collected. HEAP can be the educational platform by providing large amounts of

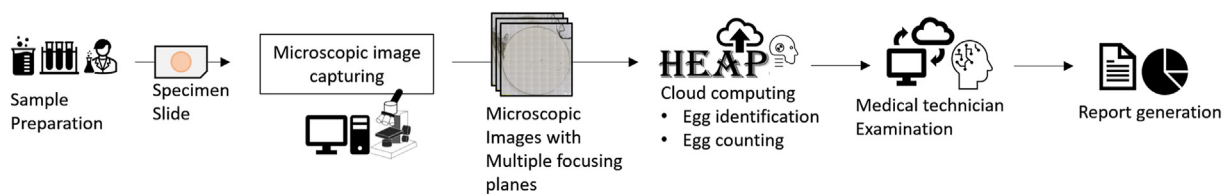


Figure 6. The workflow for HEAP-assisted parasite egg examination. In the first step, all the sample preparation procedures are processed. Then, the specimen slides are digitalized by an automatic microscope image capturing system. To overcome the focusing issue, multiple focusing planes are required to gather all the images. HEAP can do the egg identification and egg counting using cloud computing. Finally, medical technicians can verify the prediction result using any computer client with internet connectivity.

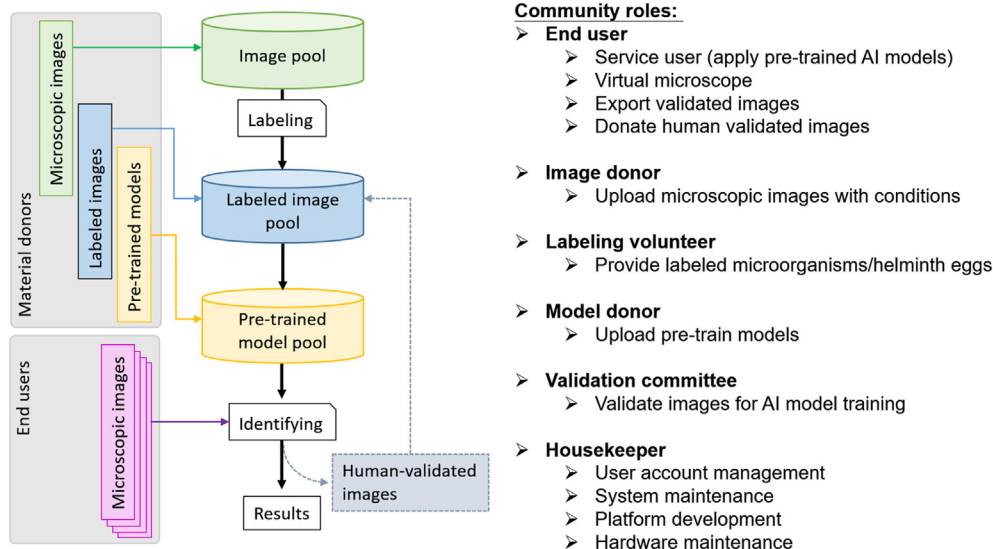


Figure 7. The schema of the microscopic image examination community. The microorganism and helminth eggs microscopic auto-recognition community architecture and roles of members.

microscopic images. Instructors can use HEAP to deliver the morphological features of microorganisms to students. When the AI models are well established, HEAP can help to release workload of medical examinations. The concepts are illustrated in Fig. 6, although HEAP provides only cloud services, it cannot reduce the sample preparation workload. However, HEAP can lower the human workload of medical examination and reduce the technician training time.

Conclusions

HEAP is designed to assist and reduce the workload of medical technicians. A user-friendly eye-inspection interface can increase human validation performance. The flexible backend of HEAP enables future integration of other AI architectures. All prediction results based on different AI models and architectures can be merged and displayed on the same page. The distributed computing architecture enables HEAP to be deployed in personal computers with a low construction cost. HEAP can be deployed in single personal computer or clusters with variant types of operating systems with/without GPU accelerators. Institutes or medical facilities can build their helminth egg identification private system off-grid from the internet to guarantee data security and safety. HEAP also serves as an ideal educational tool for medical technician training. With the continuous collection of helminth egg images and human-validated samples, the prediction rate and covered parasite types in the HEAP can improve and become comprehensive.

Authors' contributions

Website and Software, CC.L., PH.L., PJ.H. and YM.Y.; Data Curation, CC.L. and PH.L.; Data Validation, PH.L.; Program Investigation, CC.L., PH.L., PJ.H. and YM.Y.; Resources,

CC.L. and P.T.; Writing-Original Draft Preparation, CC.L. and WH.C.; Writing-Review & Editing, CH.C., WH.C. and P.T.; Supervision, WH.C. and P.T. All authors have read and approved the final manuscript.

Acknowledgement

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The web site of HEAP is freely accessible at <http://heap.cgu.edu.tw>.

Declaration of competing interest

The authors declare that they have no competing interests.

Funding

This study was supported by grants from the Chang Gung Memorial Hospital Research Funding (CORPD2J0051) and Ministry of Science and Technology, Taiwan (109-2221-E-182-049 and 110-2221-E-182-048). Publication of this article was sponsored by the grant BMRP056.

References

1. Organization WH. Schistosomiasis and soil-transmitted helminthiasis: numbers of people treated in 2017. *WHO* 2018; **50**(2018):681–92.
2. Hotez PJ, Brindley PJ, Bethony JM, King CH, Pearce EJ, Jacobson J. Helminth infections: the great neglected tropical diseases. *J Clin Invest* 2008; **118**(4):1311–21.
3. Olveda DU, Li Y, Olveda RM, Lam AK, Chau TN, Harn DA, et al. Bilharzia: pathology, diagnosis, management and control. *Trop Med Surg* 2013; **1**(4).
4. Pullan RL, Smith JL, Jasrasaria R, Brooker SJ. Global numbers of infection and disease burden of soil transmitted helminth infections in 2010. *Parasites Vectors* 2014; **7**:37.
5. hlaing Thein, Thanetoe Thansaw, Kyin Myat-lay, Lwin Myint. A controlled chemotherapeutic intervention trial on the relationship between *Ascaris-lumbricoides* infection and malnutrition in children. *Trans Roy Soc Trop Med Hyg* 1991; **85**(4): 523–8.
6. Munoz-Antoli C, Perez P, Pavon A, Toledo R, Esteban JG. Soil-Transmitted helminth infections and anemia in schoolchildren from Corn Island Archipelago (RAAS, Nicaragua). *Am J Trop Med Hyg* 2018; **99**(6):1591–7.
7. Dematei A, Fernandes R, Soares R, Alves H, Richter J, Botelho MC. Angiogenesis in *Schistosoma haematobium*-associated urinary bladder cancer. *Apmis* 2017; **125**(12):1056–62.
8. Sripa B, Brindley PJ, Mulvenna J, Laha T, Smout MJ, Mairiang E, et al. The tumorigenic liver fluke *Opisthorchis viverrini* - multiple pathways to cancer. *Trends Parasitol* 2012; **28**(10): 395–407.
9. Speich B, Knopp S, Mohammed KA, Khamis IS, Rinaldi L, Cringoli G, et al. Comparative cost assessment of the Kato-Katz and FLOTAC techniques for soil-transmitted helminth diagnosis in epidemiological surveys. *Parasites Vectors* 2010; **3**.
10. Knopp S, Mgeni AF, Khamis IS, Steinmann P, Stothard JR, Rollinson D, et al. Diagnosis of soil-transmitted helminths in the era of preventive chemotherapy: effect of multiple stool sampling and use of different diagnostic techniques. *PLoS Neglected Trop Dis* 2008; **2**(11).
11. Bogoch II, Andrews JR, Speich B, Ame SM, Ali SM, Stothard JR, et al. Short report: quantitative evaluation of a handheld light microscope for field diagnosis of soil-transmitted helminth infection. *Am J Trop Med Hyg* 2014; **91**(6):1138–41.
12. Drake LJ, Bundy DAP. Multiple helminth infections in children: impact and control. *Parasitology* 2001; **122**:S73–81.
13. Goncalves AQ, Junqueira ACV, Abellana R, del Barrio PC, Terrazas WCM, Sodre FC, et al. Prevalence of intestinal parasites and risk factors for specific and multiple helminth infections in a remote city of the Brazilian Amazon. *Rev Soc Bras Med Trop* 2016; **49**(1):119–24.
14. Yang YS, Park DK, Kim HC, Choi MH, Chai JY. Automatic identification of human helminth eggs on microscopic fecal specimens using digital image processing and an artificial neural network. *Ieee T Bio-Med Eng* 2001; **48**(6):718–30.
15. Jimenez B, Maya C, Velasquez G, Torner F, Arambula E, Barrios JA, et al. Identification and quantification of pathogenic helminth eggs using a digital image system. *Exp Parasitol* 2016; **166**:164–72.
16. Sengul G. Classification of parasite egg cells using gray level cooccurrence matrix and kNN. *Biomed Res* 2016; **27**(3):829–34.
17. Alva A, Cangalaya C, Quiliano M, Krebs C, Gilman RH, Sheen P, et al. Mathematical algorithm for the automatic recognition of intestinal parasites. *PLoS One* 2017; **12**(4).
18. Wei Liu DA, Erhan Dumitru, Szegedy Christian, Scott Reed, Fu Cheng-Yang, Berg Alexander C. SSD: single Shot MultiBox detector. *Proc Eur Conf Comput Vis* 2016:21–37.
19. Falk T, Mai D, Bensch R, Cicek O, Abdulkadir A, Marrakchi Y, et al. U-Net: deep learning for cell counting, detection, and morphometry (vol 16, pg 67, 2019) *Nat Methods* 2019; **16**(4):351.
20. Ren S, He K, Girshick R, Sun J. Faster R-CNN: towards real-time object detection with region proposal networks. *arXiv* 2015:1506.01497.