**Intestinal Worms Final Technical Report**

Global Solutions VIP

Team Members:

Anna Hart

Nathaly Jose-Maria

Shinhaeng Lee

Chenyu Li

ADVISORS:

Dr. James Stubbs, PhD

Dr. Kelsey Kubelick, PhD

Lucia Coco, TA

**Executive Summary**

The purpose of our project is to improve diagnosis of soil transmitted helminth in under-resourced global communities. In doing so, communities will receive more effective treatment and move towards the goal of eliminating child mortality from soil transmitted helminth. To achieve this goal, our team aims to reduce the amount of technical labor needed to diagnose soil transmitted helminth infections. Our solution must detect the presence, species, and number of intestinal worm eggs with sufficient accuracy and efficiency (Children Without Worms, 2021a; Children Without Worms, 2021b). Our team aims to produce a machine learning model which can effectively classify the number and type of worm eggs, run based off of standard microscope images, and is robust to noise (Children Without Worms, 2021a; Children Without Worms, 2021b). This semester, our team worked towards creating a prototype dataset using artificial stool and egg samples with which we can train our machine learning model. Recently, we established the ground truth for our dataset and formatted, split, and exported the dataset to Github which can be immediately used by future teams to train a model. We continued research into various machine learning algorithms and created a diagram containing the pros and cons of each algorithm.

**Introduction**

Our project focuses on mitigating the impact of intestinal worms on children in regions with endemic soil transmitted helminth (STH) infections (Children Without Worms, 2021a; Children Without Worms, 2021b). STH infections pose particular risks to the health of children and pregnant mothers and affects over a billion people worldwide (Children Without Worms, 2021a; Children Without Worms, 2021b). The World Health Organization treats STH as a widespread community issue when more than 2% of the population has a moderate or severe STH infection (Children Without Worms, 2021a; Children Without Worms, 2021b). In order to determine where to allocate resources and administer deworming medications to combat STH, workers need an efficient, effective way to determine which communities meet this benchmark (Children Without Worms, 2021a; Children Without Worms, 2021b). Four major species of STH are of primary concern: roundworms, whipworms, threadworms, and hookworms (Children Without Worms, 2021a). While there are many methods of analyzing stool samples, and methods like PCR have shown a slightly better diagnostic accuracy according to some studies, those methods require increased equipment costs, and the Kato-Katz remains the “WHO golden standard” (Mbong et al., 2020). Currently, the Kato-Katz technique manually processes and counts the number of worm eggs present in stool samples of the community; however, from the standpoint of health professionals, this requires long hours of skilled labor with noxious materials, harming both the worker’s quality of work and the data collection process as human error increases with exhaustion from longer work hours (Children Without Worms, 2021a; Children Without Worms, 2021b; Akintayo et al., 2018). Collecting and processing the data is cited as the most time consuming step, and we seek to minimize this through a more efficient and automated sampling and analysis technique (Stuyver & Levecke, 2021). An automated system would allow for more samples to be processed at a higher accuracy and therefore enhance the aid healthcare professionals can provide to patients as they know about these egg counts in patient stool sooner, and more conveniently (Children Without Worms, 2021a; Children Without Worms, 2021b). Through discussion with Capstone, we learned a previous Capstone team created a 3D-printed device with an installed microscope through which the users can take pictures of stool slides in the box; thus, we aim to further decrease the technical effort required by creating an automated STH diagnosis system with machine learning. Children Without Worms operates in Uganda, Bangladesh, and Kenya (Children Without Worms, 2021a; Children Without Worms, 2021b). The problem space is further complicated by a lack of reliable electricity, transportation, and laboratories in these communities (Jain & Keith, 2021). A reliable, efficient, and standardized method of quantifying STH infection using stool samples is necessary to make critical decisions regarding de-worming medication and allocation of resources in these communities (Children Without Worms, 2021a; Children Without Worms, 2021b; Butploy et al., 2021).

**Description of Project Goals and Scope**

In order for our solution to be successful, the machine learning model needs to accurately count the number of eggs within a stool sample, determine the species of the eggs, and based on that information, classify the severity of the infection through numerical classification by egg count in grouped data sets (Children Without Worms, 2021a; Children Without Worms, 2021b). Our current proposed solution is to process stool samples using Kato Katz, then image the samples with a basic microscope, upload the images to the cloud, analyze the images with a machine learning model, and relay the results back to the team in a user-friendly interface.

Our solution needs to comply with all World Health Organization standards, and we will measure success as such. While we may need to balance sensitivity and specificity, the World Health Organization allows accuracy ranging from 60% sensitivity and 99% specificity to 86% sensitivity and 94% specificity (Children Without Worms, 2021a; Children Without Worms, 2021b). Additionally, we need to maintain a low cost, ideally with equipment costs of less than $1,000 (Children Without Worms, 2021a; Children Without Worms, 2021b). Most importantly, our user requirements are to have a sufficiently accurate model - the details of which are described above- with less technical labor required than the current standard, Kato Katz. Our solution must be simple, robust, and accessible. Our solution will need to identify multiple species of worm eggs: namely, *Ascaris,* Hookworm, and *Trichuris,* as these are common infections in the areas Children Without Worms serves(Children Without Worms, 2021b)*.* Moreover, our solution needs to work in areas with unreliable electricity and internet and must be usable at varying sites (Children Without Worms, 2021b). Our solution must be robust to variations in worm egg characteristics, debris from stool samples, and adapt to thick stool samples (Children Without Worms, 2021b). Ideally, our solution should be able to classify the STH infection based on three or fewer slides (Children Without Worms, 2021b).

Through a meeting coordinated with the Capstone Team, we gained access to 3-D printed prototypes of a manual machine for fecal slide imaging, lighting, microscopy options for slide imaging, sample slides of fake stool, and materials for making fake stool. A recently published video on current research conducted by Johnson & Johnson regarding egg-counting served as inspiration (Greco et al., 2021). From here, we named two specific goals for our machine learning algorithm research: that we can efficiently improve the image quality from noisy microscope images and that our algorithm can effectively count and classify the worm eggs.

This semester, our goal has been to curate a dataset of samples for the machine learning model to learn from due to an unavailability of open-source databases of intestinal worm samples in order for the team to train the machine learning model. Our goal, which we achieved, was to create and image artificial stool samples and curate them in a labeled dataset available for future semesters. We will continue to operate with artificial stool samples until we can safely obtain real stool slides or images of these slides from partners to compile real images into a defined data set in future semesters. The team also held the goal of continuing research into machine learning algorithms and image processing techniques and analyzing which are best suited for our problem space. Our work will help teams in future semesters, as they will have preliminary data to begin building machine learning models using our labeled dataset as training data as well as our research into potential machine learning models to try.

**Project Status**

Upon our return from Spring Break, we received updates from one of our partner labs in Uganda (Do, 2022). We analyzed the images from the partner to further define our goals and problem space and identify challenges to the implementation of our model. Samples are stored with refrigeration (Images 3, 4). PPE protocols and sanitation procedures are implemented in the lab area, particularly with the use of lab coats, gloves, and masks, but also in the addition of newspaper on their tables to make cleanup safer and easier, cans and bottles of disinfectant both sitting on the table and on-hand, the presence of toilet paper for picking up messes, and large yellow jugs assumingly holding water for sanitation (Image 6). Stool slides are prepared on-site with the use of stool samples from containers within refrigeration units, slides with barcodes for scanning with a device, assistive tools like paper to create the ideal sample size of stool for placement on the slide, and more, though these additional tools are unidentifiable for the moment without more information (Images 8, 10). A mobile phone and an Olympus CX21 microscope are available (Images 4,5). However, we did learn that there is not a streamlined process for collecting and sending digitized images of microscopy slides in the lab. This is evident by the lack of slide images received. We did receive one image of a stool sample; this gives us an indication of the quality of photos for our algorithm and the general appearance of an egg; however, it does not suffice to build an adequately sized dataset, and the slide was classified as a non-STH sample (Image 7). Overall, this work was successful in better understanding the project space and collaborating with the Capstone team (Do, 2022). However, we failed to collect adequate numbers of STH samples to build a dataset. In future, it would be useful to collaborate with Capstone to implement a system by which the field team can send us digitized microscopy images of STH samples with little to no additional time required.

Earlier in the semester, we created artificial intestinal worm egg stool samples using glitter as worm eggs and artificial stool adapted from the paper, "Review of synthetic human faeces and faecal sludge for sanitation and wastewater research" (Penn et al., 2021). We obtained approximately 150 samples containing worm eggs and 150 samples free of worm eggs. Next, we worked to assemble these samples into a working dataset. We researched the article, "How I Created a Dataset for Instance Segmentation from Scratch?" to develop our method for labeling the dataset (Leven, 2020). Since our algorithm only needs to identify which and how many eggs there are - not know exactly where they are in the image - we decided to use bounding boxes as opposed to pixel-by-pixel labeling (Pokrel, 2020). We began labeling our dataset with Sense High Performance Data Annotation (Sense High Performance Data Annotation, 2022). We have approximately 100 images in our dataset currently. Since we can effectively use the portions of the slides not containing the glitter as our control, we only used the images that contained the artificial worm eggs. We created one label: ‘Egg’, and drew a bounding box as tightly around each artificial egg as possible in every image. A screenshot of the process is located in Image 1. Additionally, we noticed that some samples had opaque stool sample matter which occluded the worm egg; this will help our model to be more robust to noise as well as teach team members how to train a computer vision model when the objects are difficult to detect. An example can be found in Image 2.

Next, we split our data into three sets: training, testing, and validating. The training dataset will be used to train the machine learning model; the testing dataset will be used to tune the hyperparameters of the model; and the validation dataset will be used to ultimately test the accuracy of our model (Agarwal, 2021). We randomized our data prior to splitting, and due to the small nature of our dataset, we maintained the default 80/10/10 split given by the data labeling application (Sense High Performance Data Annotation, 2022). We exported our dataset in the COCO format. Future teams will benefit from having a correctly formatted dataset with which to learn how to use various machine learning models before using them on the true intestinal worm egg dataset. This will enable team members to learn machine learning techniques while we are finding our true dataset, as well as minimize any wasting of the true dataset due to unfamiliarity with training machine learning models.

Overall, we have created a prototype machine learning dataset of over 100 images. Sample labeled images can be seen in Images 1 and 2. We consider this an overall success, as the ultimate goal of the artificial sample dataset was to create sufficient data for team members to learn the process of training a machine learning model and the mechanics of using the data formats in computer vision. However, we do recognize that our dataset is small and this may limit the ability to create a robust dataset. Future teams may benefit from adding to the dataset to learn the tenets of dataset curation as well as to understand the impact of dataset size on model performance, and we have 50 unlabeled images for this purpose. Furthermore, while our model is an effective prototype for future teams to practice creating machine learning models with, it will not suffice to create a prototype which mimics true STH samples. This is due to the opaqueness and constant size of the glitter, which is fundamentally different from the varied shapes and translucency of STH eggs (Do, 2022).

Finally, a main goal of this semester was to research and understand machine learning models that could be applied to our problem space. After our midterm report, we consolidated some of our research into algorithms and analyzed the pros and cons between different algorithms. For example, we found that both "Towards an automated medical diagnosis system for intestinal parasitosis" and "Circular Object Detection Using a Modified Hough Transform" have robust pre-processing steps that could be valuable for our future use (Beaudelaire et al., 2019; Smereka, 2008). In contrast, "A deep learning framework to discern and count microscopic nematode eggs" uses deep learning techniques - a more modern approach to computer vision - yet deal with data which is not as similar to ours, as they use soil samples whereas we use stool samples (Akintayo et al., 2018).

We also did additional research into machine learning algorithms. We furthered our research on an algorithm noted in the scientific paper “A deep learning framework to discern and count microscopic nematode eggs” to apply in our own algorithm creation (Akintayo et al., 2018). Utilizing a layer system when analyzing images greatly improves accuracy by getting different perspectives of the same image in order to reconstruct and analyze the image. This ties very closely to our goal of minimizing error in our STH diagnosis system.. Next, nematode eggs do not exhibit a large contrast to their environment. Thus, the algorithm will likely transfer well to our STH use-case, where the STH eggs also do not exhibit a large contrast in the slide. One issue however, is a lack of access to the trained model in the paper. We do have a contact for the team, which we can reach out to in the future once we finish our own analysis of the algorithm. Lastly, the paper incorporates a selective autoencoder to improve accuracy (Akintayo et al., 2018). The autoencoder analyzes images by applying mathematical equations and masking filters in such a way that makes it more accurate than normal encoders (Akintayo et al., 2020). It introduces a bias in one of the stability metrics in the mathematics behind the algorithm and masks parts of the image at a time such that the computer analyzes each frame more intently (Akintayo et al., 2020). However, due to our inexperience with the topic, our team needs to conduct further research to better understand the mechanics of how this autoencoder works in application.

We also researched further into the YOLO algorithm. The YOLO algorithm is based on regression and used to classify objects in the images as well as indicate the region of the image where they lie (Skrzydło, 2022). The YOLO algorithm yields high efficiency due to the use of a neural network that predicts both the type and location of objects at once. However, YOLO is limited in accuracy with small, clustered objects (Bandyopadhyay, n.d.).

We did more research into the use of deep learning in image recognition (Image Recognition, 2019). Typically, these algorithms identify each pixel of the image; however, we will need to alter this as we do not need to know the exact locations of the STH eggs and a pixel by pixel analysis would increase our computational overhead. Furthermore, we noted the necessity of a large, pre-labeled dataset which is then used by the model to train and adjust weights in the filters and kernels to develop a detection algorithm. From this, we recognize that finding a large, pre-labeled dataset will be a crucial - albeit difficult - step in our development of the machine learning model for STH (Image Recognition, 2019).

Lastly, we recognize that a major issue with our work may be the lack of STH sample data. As we will likely have a small set due to the unavailability of data, we are looking into ways to handle the small size of a dataset. For example, we can use transfer learning to pre-train a model with a larger dataset not directly relevant to our problem, and then use a smaller dataset tailored to our purpose to fine-tune the model (Leven, 2020). Transfer learning is commonly used in computer vision and begins with general features from a pre-trained model, decreasing the time and expense needed to train a model (Brownlee, 2017). To implement transfer learning, we will need to find a reliable pre-trained model and train the model to suit our unique problem. Future teams will need to find an adequate parent dataset, and analyze which portions of the parent model will need to be fine-tuned to the intestinal worm egg problem (Brownlee, 2017).

**Conclusions**

Our overarching work aims to impact global health by increasing the efficiency of STH diagnosis in the field while minimizing the amount of human labor and time needed to make these diagnoses. Furthermore, by having an artificial intelligence solution to diagnosis, we aim to improve access to high quality diagnosis and treatments in global communities with a shortage of trained pathologists (Children Without Worms, 2021a; Children Without Worms, 2021b). On a near-term level, our work this semester has created a prototype dataset of artificial worm egg stool samples, which has now been properly curated and exported. This will enable us to prototype a machine learning model for STH infections, which will inform us as we begin working with true human STH infection data. Furthermore, we have developed a list of machine learning algorithms, as well as their pros and cons, which will inform future semester members in their choice of machine learning algorithms to apply to the STH diagnosis problem.

Furthermore, we have researched many algorithms for image enhancement and image recognition. We have narrowed down our candidate algorithms and weighted the pros and cons of each algorithm to aid in future decision making. Additionally, we created a dataset of 100 artificial worm egg samples. One shortcoming of our work this semester is that our artificial samples do not closely mimic STH samples. This informs future teams that they must find an accurate dataset of human STH samples in order to train a clinical grade machine learning model. Nonetheless, our prototype dataset will still be crucial for future teams to be able to train new members on the tenets of machine learning and develop a prototype model while waiting for a more robust dataset. Furthermore, future teams are advised to reference our research into existing machine learning models which they can then apply first to the prototype dataset and then to a true clinical dataset in order to create an artificial intelligence solution for STH diagnosis.

One major issue we have identified is access to data. There are two main roadblocks we face in this respect. First, we don't have access to intestinal worm samples from the countries in which Children Without Worms is working. We were hoping to receive intestinal worm egg slides from Ghent, but this unfortunately fell through due to import regulations. We currently have our artificial worm egg samples. This is helpful to our team because it enables members to learn how to curate machine learning datasets, train machine learning models, and ultimately run and test different machine learning models to find ones that work best. However, it will not enable us to create an algorithm that could accurately detect intestinal worm eggs in the field and differentiate between species. This is because for our artificial stool samples, we used glitter for worm eggs - which with its opaque hexagonal shape differs from the variable, translucent morphology of worm eggs (Do, 2022). Furthermore, even if we tried another substance - such as silica particles - the particles still would not replicate the interspecies variations of the STH eggs.

The second roadblock we face to accessing a dataset is the need for accurate labeling of the dataset. The machine learning model can only be as accurate as the ground truth we give it (Leven, 2020). Thus, any dataset we use will need to have an accuracy in labeling worm eggs and species that is greater than the World Health Organization standards. Likely, this will mean finding a pathologist who could read our images - or finding a pre-labeled dataset. While our team members would be eager to learn how to read intestinal worm egg slides, the reality is that we could not replicate the accuracy of a pathologist who has had years of training and experience in the field. This would result in our model not meeting clinical standards for diagnostic tools.

Our recommendations for future work of the teams is, first and foremost, to find an accurate and robust dataset. There are multiple avenues that could be pursued to achieve this. First, we could see if Ghent might be willing to send images of the slides rather than the slides themselves: this may get around import regulations as well as save us time in imaging the slides. Second, we could look online to find an available dataset for intestinal worm egg samples. Unfortunately, our team has currently been unable to find such datasets. However, we have seen an article that does use intestinal worm egg samples (Beaudeclaire et al., 2019). Our best avenue would likely be to reach out to these researchers, and ask whether they would be willing to share their data - or if they would be able to point us in the right direction.

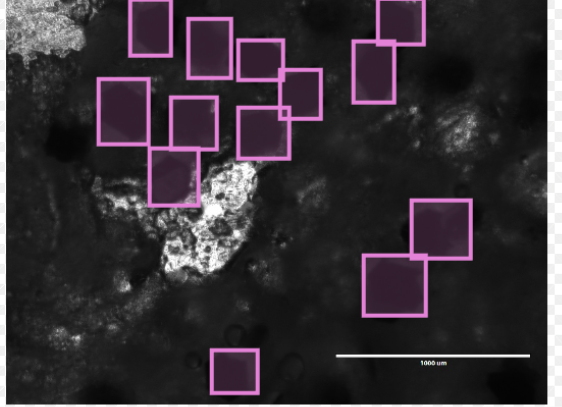
**Addendum**

**Image 1: *Workflow for Image Labeling***

LABEL & TRAIN > DATASETS > DATASET DETAILS 
Artificial Stool Samples 
Activate Pay-as-you-go 
Dashboard 
Appearance 
Egg 
Sources 
Rotate Right 
Label Definitions 
Resize TO Fit 
28 
Labeler 
Undo 
Review 
Red O 
Versions 
Clear All 
O Help 
E XP and 
Bounding Box 
a Egg 
Pending 
(26) 
Olmage Details 
Instructions 
Pending 
Help Topics 
Skip 
Submit 

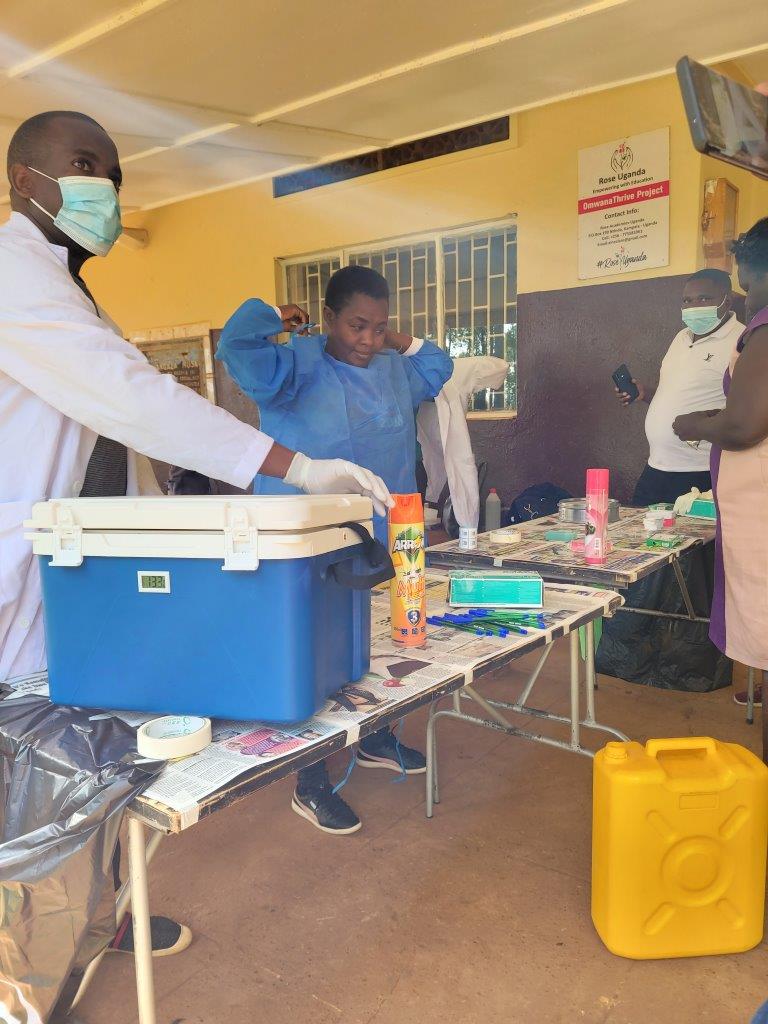
The above image displays the work of labeling our data. We drew one bounding box as tightly as possible around each worm egg - which was defined as having a fairly hexagonal, opaque, shape in the slide. (Screenshot from: Sense High Performance Data Annotation 2022)

**Image 2: *Sample Labeled Image***



The above image shows a sample of a labeled image. Here, the stool sample was fairly thick which diminished the contrast between the worm egg and the slide background. A pink bounding box was drawn tightly around each artificial worm egg. (Screenshot from: Sense High Performance Data Annotation 2022)

**Image 3: *Uganda Lab - Stool Sample Storage***



(Do, 2022). This image exemplifies some of the storage and cleanliness techniques implemented by the team in the Uganda lab to maintain organization and sanitary procedures with the active manipulation of stool samples occurring locally. These protocols include traditional PPE protocols and enhanced sanitation with disinfectants and newspaper covers.

**Image 4: *Uganda Lab - Stool Sample Storage & Available Imaging Technology 1 - Mobile Phone***



(Do, 2022). This image exemplifies some of the organization techniques and imaging technology being utilized by the Uganda lab. Organization techniques include storage of stool samples in individual containers all stored within a smaller refrigerator unit. Imaging technology includes the use of a handheld mobile device of unknown type, brand, and quality. This suggests more technology and electricity available in our experimental areas than previously suggested.

**Image 5: *Uganda Lab - Available Imaging Technology 2 - Microscope***



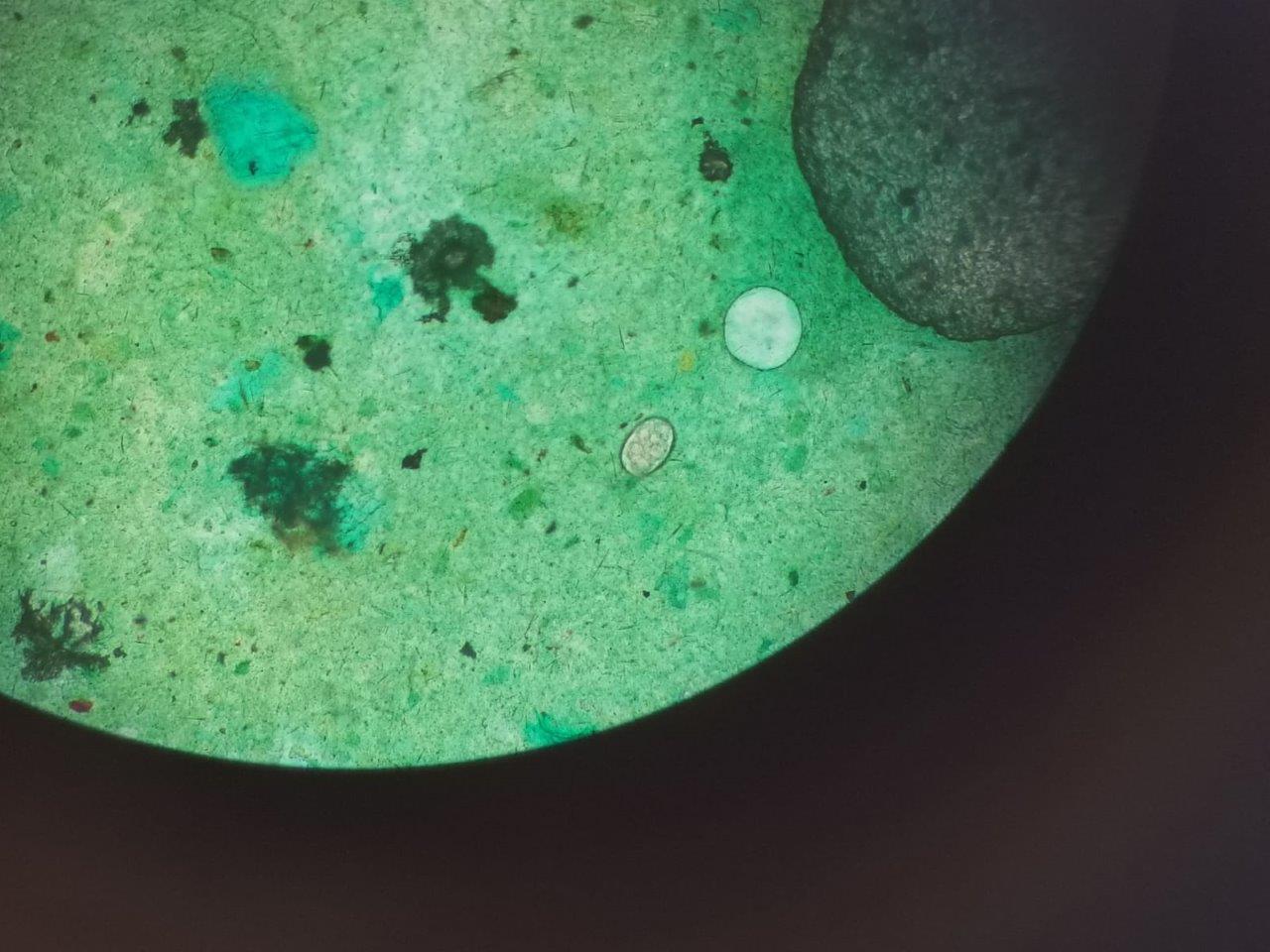
(Do, 2022). This image exemplifies more of the imaging technologies available in the Uganda lab. Pictured to the right is another perspective of the handheld mobile device used to picture stool containers to seemingly match QR code slides to them post-imaging and inspection. Pictured in the center is an Olympus CX21 microscope for looking at fecal slides microscopically. This suggests more advanced technology and electricity is available in our experimental areas than was previously suggested.

**Image 6: *Uganda Lab - PPE Protocols and Sanitation & Worker Density***



(Do, 2022). This image demonstrates the average worker density and spacing in the Uganda lab while also exemplifying some of the sanitation techniques necessary in such noxious lab conditions, including gloves, masks, disinfectant spray, and water.

**Image 7: *Uganda Lab - Stool Sample through Microscopic Lens***



(Do, 2022). This image exemplifies what the slides prepared on-slide look microscopically. The image demonstrated the green background found in the physical slides while also showing stool particles, air bubbles, and non-STH worm eggs.

**Image 8: *Uganda Lab - On-Site Slide***



(Do, 2022). This image shows what slides prepared on-site look like after creation before microscopy. The slides exhibit a blue hue to provide a colored background under the microscope for easier object detection while also having an associated barcode tag for easier sample and slide correlation.

**Image 9: *Uganda Lab - On-Site Slide Preparation 1***



(Do, 2022). This image shows a worker of the Uganda lab preparing a fecal slide by picking up stool from a container with a small wand and spreading excess stool on a pan-like object to seemingly apply the remaining smaller concentration on the slide for microscopy.

**Image 10: *Uganda Lab - On-Site Slide Preparation 2***



(Do, 2022). This image shows the same worker from Image 9 preparing a slide with the smaller amount of stool picked up on the wand and being placed on a slide with an assistive paper to help better approximate the stool concentration being applied.

**References**

Agarwal, R (5 October 2020). "How I created a Dataset for Instance Segmentation from Scratch?" Medium. <https://medium.com/mlwhiz/how-i-created-a-dataset-for-instance-segmentation-from-scratch-1cbf1a771a03>

Agarwal, S (17 May 2021). "How to split data into three sets (train, validation, and test). And Why?" TowardsDataScience. <https://towardsdatascience.com/how-to-split-data-into-three-sets-train-validation-and-test-and-why-e50d22d3e54c>

Akintayo, A., Lore, K.G., Sarkar, S., & Sarkar, S. (2020). Prognostics of Combustion Instabilities from Hi-Speed Flame Video Using a Deep Convolutional Selective Encoder. *IJPHM Special Issue on Big Data and Advanced Analytics for PHM,* 7(4). <https://doi.org/10.36001/ijphm.2016.v7i4.2461>.

Akintayo, A., Tylka, G.L., Singh, A.K., Ganapathysubramian, B., Singh, A., & Sarkar, S. (2018). A deep learning framework to discern and count microscopic nematode eggs. *Sci Rep* 8, 9145. <https://doi.org/10.1038/s41598-018-27272-w>

Bandyopadhay, H. (n.d.). *Yolo: Real-time object detection explained*. V7. Retrieved March 17, 2022, from https://www.v7labs.com/blog/yolo-object-detection

Beaudelaire Saha Tchinda, Michel Noubom, Daniel Tchiotsop, Valerie Louis-Dorr, Didier Wolf (2019). Towards an automated medical diagnosis system for intestinal parasitosis. Informatics in Medicine Unlocked, Volume 16, 2019, 100238, ISSN 2352-9148, https://doi.org/10.1016/j.imu.2019.100238.

Brownlee, J. (3 May 2019). "Best Practices for Preparing and Augmenting Image Data for CNNs." Machine Learning Mastery. https://machinelearningmastery.com/best-practices-for-preparing-and-augmenting-image-data-for-convolutional-neural-networks/

Do, Jennifer, personal communication, 1 April 2022.

Fish, D. A., Brinicombe, A. M., Pike, E. R., & Walker, J. G. (1995). "Blind deconvolution by means of the Richardson–Lucy algorithm," J. Opt. Soc. Am. A 12, 58-65.

Greco, B., Nogaro, S., & Stuyver, L. [The International Society for Neglected Tropical Diseases]. (2021, Jan 29). Access to schistosomiasis innovation: the next frontier to reach elimination by 2030. [Video]. YouTube.<https://www.youtube.com/watch?v=jguh-vzdkQM>.

*Image recognition : A complete guide*. Deepomatic. (2019, January 8). Retrieved April 21, 2022, from <https://deepomatic.com/what-is-image-recognition>. <http://shabdbooks.com/gallery/34-june2020.pdf>

Jain, Manav, Keith, Emily, et, al. (28 September 2021) “Design Inputs Report, Paradigm” Georgia Institute of Technology.

Leven, S. (13 August 2020). "Machine Learning's Secret Sauce: Curation" Towards Data Science. <https://towardsdatascience.com/machine-learnings-secret-source-curation-e8c3107dcc13>

Mbong Ngwese, M., Prince Manouana, G., Nguema Moure, P. A., Ramharter, M., Esen, M., & Adégnika, A. A. (2020). Diagnostic techniques of soil-transmitted helminths: Impact on control measures. *Tropical Medicine and Infectious Disease*, *5*(2), 93.

Narut Butploy, Wanida Kanarkard, Pewpan Maleewong Intapan, "Deep Learning Approach for Ascaris lumbricoides Parasite Egg Classification", *Journal of Parasitology Research*, vol. 2021, Article ID 6648038, 8 pages, 2021.<https://doi.org/10.1155/2021/6648038>

Pokrel, Sabrina (11 March 2020). "Image Data Labelling and Annotation - Everything You Need TokKnow."k<https://towardsdatascience.com/image-data-labelling-and-annotation-everything-you-need-to-know-86ede6c684b1>

Roni Penn, Barbara J. Ward, Linda Strande, Max Maurer, Review of synthetic human faeces and faecal sludge for sanitation and wastewater research,Water Research,Volume 132, 2018, Pages 222-240

*Sense High Performance Data Annotation*. (2022). [Software]. Plainsight.ai. <https://plainsight.ai/sense-platform-sign-up/>.

Skrzydło, A. (2022, March 17). *Yolo algorithm and Yolo Object Detection*. R Shiny. Retrieved April 23, 2022, from https://appsilon.com/object-detection-yolo-algorithm/

Stuyver LJ, Levecke B (2021) The role of diagnostic technologies to measure progress toward WHO 2030 targets for soil-transmitted helminth control programs. PLOS Neglected Tropical Diseases 15(6): e0009422.<https://doi.org/10.1371/journal.pntd.0009422>