Water Practice & Technology

IWA PUBLISHING

Water Practice & Technology Vol 19 No 4, 1542 doi: 10.2166/wpt.2024.087

Reducing costs of *Giardia* spp. cyst enumeration using machine learning-based systems

Valdinei Luís Belini [a,*], Natália de Melo Nasser Fava[b] and Lyda Patrícia Sabogal-Paz[b]

[a] Department of Electrical Engineering, Universidade Federal de São Carlos, UFSCar Rodovia Washington Luís, km 235, São Carlos CEP 13565–905, SP, Brazil
[b] Department of Hydraulics and Sanitation, São Carlos School of Engineering, University of São Paulo, Avenida Trabalhador São-Carlense 400, São Carlos 13566-590, SP, Brazil
*Corresponding author. E-mail: belini@ufscar.br

VLB, 0000-0002-8894-7519

ABSTRACT

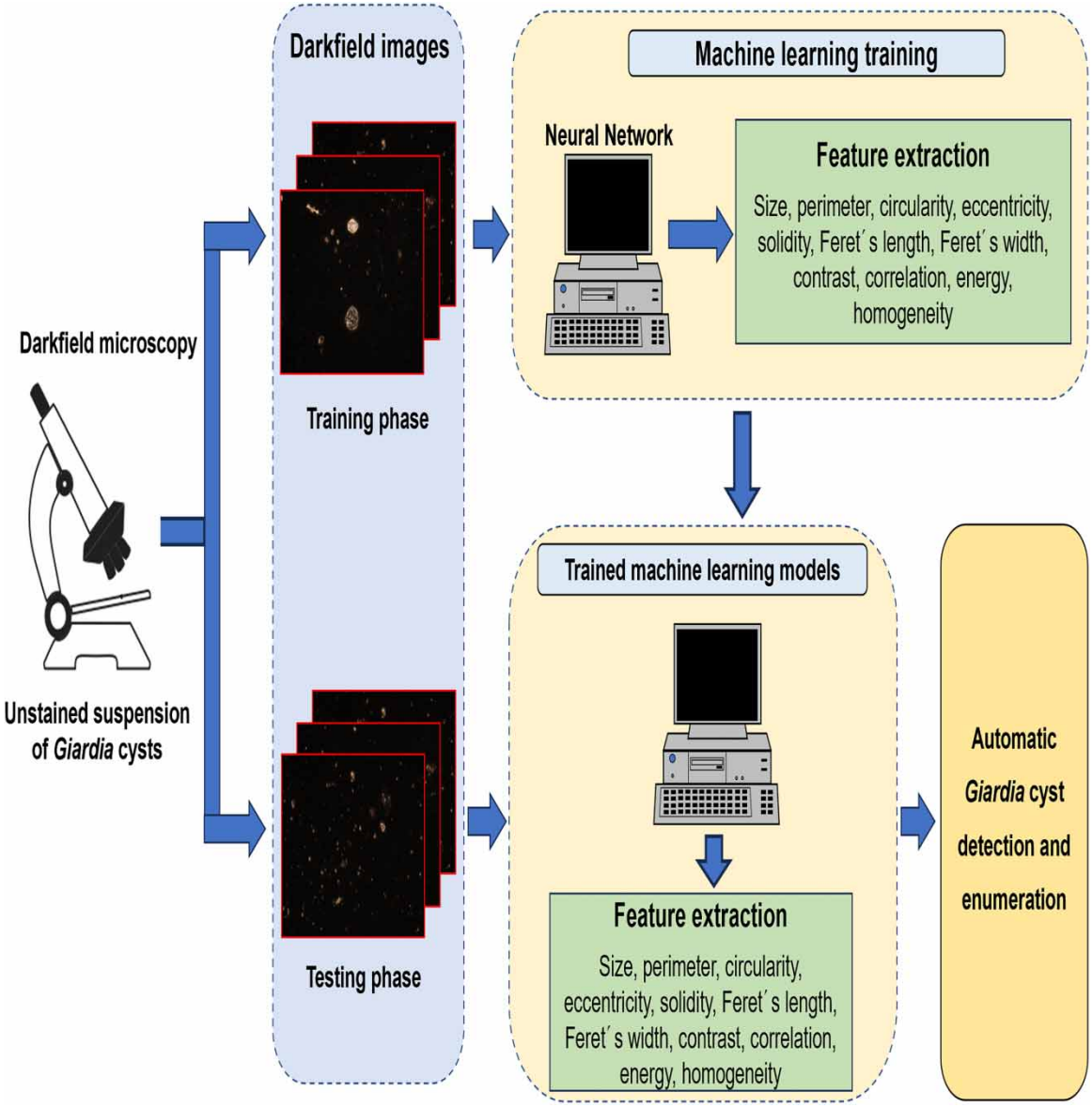
Giardia spp. cyst enumeration is a laboratory procedure that is frequently required in water treatment-related research. Currently, detection conducted by experts using fluorescence microscopy on samples stained with specific markers for *Giardia* spp. cysts is still the standard method, despite its high costs limiting its usage worldwide and, ultimately, hindering waterborne analyses in low-income countries. We present an approach based on darkfield imaging and machine learning to reduce costs associated with *Giardia* spp. cyst enumeration and the lack of experts. Automated counts were compared to manual counts, achieving an average sensitivity (SE) rate of 88%, specificity (SP) of 100% and accuracy of 88% across a wide range of cyst concentrations. By using machine learning in conjunction with darkfield microscopy, a low-cost illumination technique that can be easily integrated into standard laboratory microscopes, we have significantly reduced the costs associated with *Giardia* spp. cyst detection, all while still maintaining the SE and SP of fluorescence microscopy. Based on the findings, the proposed system has the potential to be a useful tool to enumerate *Giardia* spp. cyst suspensions. It can be accessed by virtually any microbiology laboratory as it is consumable-free and expert-independent.

Key words: drinking water, fluorescence, image analysis, protozoa, sustainable development goal 6 (SDG6), water treatment

HIGHLIGHTS

- First application of darkfield microscopy for automatic *Giardia* cyst detection and enumeration.
- We have developed a machine learning-based system to overcome cost limitations and lack of trained personnel.
- Our system allows for label-free establishment of *Giardia* cyst concentrations, which can be used in routine laboratory procedures.

GRAPHICAL ABSTRACT



1. INTRODUCTION

Giardia duodenalis is one of the main intestinal parasites in humans worldwide and among the most common reasons for diarrheal diseases (Keserue *et al.* 2012). Each year, an estimated 1.5 million deaths globally result from diarrheal diseases, and most of these deaths occur in low-income settings (Daniels *et al.* 2018). Drinking water is the major transmission route for infections, thus needing proper water treatment and monitoring to avoid human infection (Feng *et al.* 2021).

Standard methods for *Giardia* spp. cyst detection in water samples, such as the U.S. Environmental Protection Agency (EPA) 1623.1 (U.S. EPA 2012), require (1) concentration of samples (filtration) and (2) parasite purification through immunomagnetic separation, followed by (3) staining and (4) visualization using fluorescence microscopy. After preparing water samples through the filtration of large volumes of water (minimum of 10 L), the concentrate undergoes purification using immunomagnetic beads to separate target organisms from process-related impurities. Afterwards, the sample receives fluorescent dyes specific for *Giardia* spp. cysts (fluorescein isothiocyanate – FITC) following the manufacturer's recommendations. For internal structure visualization, the sample also receives 4′,6-diamidino-2-phenylindole (DAPI). After staining, the prepared coverslipped microscope slide is kept dry at room temperature overnight. After staining, a well-trained technician

places the prepared samples on a fluorescence microscope for cyst detection and enumeration manually. For this task, the technician screens the whole microscope slide for characteristic FITC-fluorescence, as well as the size and shape of the *Giardia* spp. cysts, confirming their identification through DAPI-stained nuclei and cytoplasm.

Over the last decade, developed countries have widely utilized Method 1623.1, but the effort and associated costs involved in various constituent procedures still make it difficult to deploy in low- and medium-income countries (Fava *et al.* 2021). When using fluorescence microscopy, the high costs and technical analytical complexity of *Giardia*-specific fluorescent dyes, laboratory infrastructure, and equipment are evident (Maciel & Sabogal-Paz 2016). The primary cost is the separation process involving immunomagnetic beads (Fava *et al.* 2021), followed by the two final steps of the analysis procedure involving staining and detection using fluorescence microscopy. From a practical perspective, microbiological methods used to detect parasites should ideally meet the criteria of being simple, easy-to-use, and rapid but reliable and low cost (Shrestha *et al.* 2020).

There is a wide range of alternative methods used for detecting *Giardia* spp. cysts in water, including flow cytometry, laser scanning cytometry, and molecular techniques (Efstratiou *et al.* 2017). However, even if these methods offer more sensitivity (SE) and specificity (SP) than the microscopic technique used in the U.S. EPA protocol, they also have the disadvantages of being time- and resource-consuming (El Mallahi *et al.* 2013).

Researchers have developed various other detection methods to partially mitigate some of these disadvantages. Mudanyali *et al.* (2010) investigated the use of a field-portable lens-free holographic microscope to record cell holograms within a custom-made cuvette and also automatically detected and enumerated *Giardia* spp. cysts at concentrations of ~190 cysts/mL. This is several orders of magnitude higher than those expected in raw water samples (Efstratiou *et al.* 2017). Besides, specimen fixation is required. Mobile platforms for *Giardia lamblia* cyst detection and quantification have also been introduced to replace bulky and costly optical/mechanical components. Koydemir *et al.* (2015) presented a smartphone-based fluorescence microscope to capture images of fluorescently labelled cysts over a disposable sample processing cassette, achieving a lower detection limit (~12 cysts/10 mL). Shrestha *et al.* (2020) developed a smartphone-based brightfield microscope for the simultaneous detection of *Cryptosporidium* and *Giardia* spp. (oo)cysts stained with Lugol's iodine. After loading a standard haemocytometer with suspensions, technicians manually analyse magnified images of (oo)cysts using a ball lens directly on the screen of a smartphone to provide estimates based on size, shape, and contrast at a detection limit of ~1 cyst/g (derived from reported 73 cysts/100 g). This is a subjective, tedious, and time-consuming task. As all these methods involve chemicals affecting the purity of the suspension, which is an expensive input, the analysed cysts cannot be used for other trials or microbiological studies. Furthermore, apart from the proposal by Koydemir *et al.* (2015), experts only analyse aliquots, meaning representative estimates can only be obtained by sampling, preparing, and analysing numerous samples. More recently, Nakarmi *et al.* (2023) reported the usage of deep learning-based object detection to automate (oo)cyst detection in images from the smartphone-based microscope developed by Shrestha *et al.* (2020). Although the ability to detect cysts using a smartphone can be valuable for field testing scenarios, the results showed that because of the curvature effect of the ball lens in the smartphone microscope on the imaged objects, the models could falsely predict *Giardia* as *Cryptosporidium* and *vice versa*. Besides, the field of view using this lens is too small, making scanning an equivalent area to that of the standard microscopy slides a tedious and time-consuming task. Automated microscopy approaches have also combined imaging flow cytometry and deep learning-based models to predict *Cryptosporidium* and *Giardia* oocysts in drinking water (Luo *et al.* 2021). Despite the ability of flow cytometry to rapidly acquire thousands of single-cell images to be analysed at rates of 346 frames per second, the cost of acquiring and maintaining this equipment makes it prohibitive for low-resource settings. Besides, the created machine learning models were not able to generalize the prediction on non-seen data in the training phase.

Only a few detection methods that allow for eliminating the labelling steps, which are time-consuming and expensive and simultaneously enable the analysis of larger volumes, have been presented. Göröcs *et al.* (2020) reported a field-portable holographic imaging flow cytometer to detect and count low levels of *Giardia* contamination (10 cysts/50 mL) at a volumetric throughput of 100 mL/h. However, a new set of disposable components is required for each measurement. More recently, our research group introduced *in situ* microscopy for *Giardia* spp. cyst detection and enumeration in agitated suspensions (Belini *et al.* 2022). Here, we capture high optical resolution brightfield microscopic images of cysts directly from a 1 mL sampling chamber and process them within 10 min to produce automated estimates at a detection limit of 30 cysts/mL, without requiring either mechanical/optical adjustments or disposable components. Despite providing estimates without involving labelling,

the capability of these methods for detecting other waterborne pathogens such as *Cryptosporidium* spp. oocysts still needs to be demonstrated.

Although improvements have been achieved in several aspects compared to standard detection methods, all the aforementioned methods rely on custom-designed devices with varying degrees of complexity. Therefore, implementing these methods still requires some level of skill, inputs, and costs. From these observations, it seems obvious that the usage of existing laboratory equipment and available personnel would be preferable.

The low contrast of (unstained) *Giardia* spp. cysts, especially under brightfield microscopy, makes them difficult to visualize without the help of staining procedures (e.g., Lugol's iodine in brightfield) or fluorescence-based dyes specific for this protozoan such as FITC. Recently, Belini *et al.* (2018) demonstrated the effectiveness of using inexpensive darkfield microscopy setups to acquire higher contrasted images of *Giardia* spp. cysts of unstained microscopy slides. By using this illumination technique, well-resolved cyst walls, specific intracellular structures, including nuclei, axostyles, and retracted cytoplasm in *Giardia* spp. cysts could be visualized, all exhibiting characteristic features comparable to those observed in fluorescence-based microscopy (FITC/DAPI) following Method 1623.1.

While sensitive detection methods are crucial, it is equally important to address the high costs associated with standard detection methods to make them viable for water quality assessment in low- and medium-income countries. Despite numerous alternative methods available for detecting this protozoan in water samples, there appears to be a lack of effort in detecting and enumerating *Giardia* spp. cysts in less complex suspensions, such as those from drinking or treated water samples, using existing standard laboratory settings. The ability to rapidly detect and enumerate *Giardia* spp. cysts (e.g., post-recovery *via* EPA 1623.1) in a label-free process would reduce both the detection time and level of human intervention required, as well as contribute to saving resources (McGrath *et al.* 2017).

The objective of this study was to develop and evaluate an automated method for detecting and enumerating *Giardia* spp. cysts in darkfield microscopic images of unstained suspensions, aiming to create a technique easily applicable in microbiologists' routine practice. The proposed solution differs from existing methods as it does not require custom-made imaging devices to perform label-free estimates. By using standard microscopes in microbiology laboratories, the proposed method has the potential to make cyst suspension assessments more cost-effective and accessible in regions with limited infrastructure as the method does not require expensive equipment, consumables for labelling, or specialists to conduct and interpret analyses. Specifically, we used standard image processing algorithms for detecting cyst *candidates* in darkfield images. After extracting characteristic object features, we trained machine learning-based models with *Giardia* cysts and non-giardia particle images. The purpose of this training was then to determine the number of *Giardia* cysts found in testing images.

The novelty of this study is the usage of darkfield images in conjunction with machine learning-based methods to enumerate *Giardia* cysts in microscopic images of unstained suspensions. Access to a more economical and accurate detection technique, as described here, can potentially reduce people's vulnerability to these parasites in developing countries, where infection is likely underdiagnosed and underreported due to limited detection resources (Shrestha *et al.* 2020).

2. MATERIALS AND METHODS

2.1. Sample preparation

To obtain the suspensions for the darkfield experiments, commercial suspensions of *G. duodenalis* cysts (Waterborne™, Inc., New Orleans, LA, USA) containing $\sim 5 \times 10^6$ cysts in 8 mL were homogenized for 2 min using a vortex and vigorously pipetted 20 times. Our research group adopted these suspensions in various experiments using them as stock suspensions in the present study. Afterwards, we serially diluted the suspensions with distilled water. For each resulting suspension, we prepared two aliquots of 50 µL with a Merifluor® (Meridian Bioscience, Inc.) kit according to the manufacturer's recommendations and used them to count (in two wells of a microscopy slide) the number of cysts on a fluorescence microscopy (BX51, Olympus®, $\lambda_{excitation} = 475$ nm, $\lambda_{emission} = 520$ nm). From this process, we obtained average concentrations of *Giardia* spp. cysts at 100, 250, 500, and 1,000 cysts/mL. We also prepared a negative control using cyst-free distilled water. For each established concentration of organisms, we loaded 25 µL replicates onto a standard glass 1.0 mm thickness microscope slide and covered the mounting with a coverslip (0.17 mm thickness), sealing it with enamel for darkfield imaging.

2.2. Darkfield image acquisition

Darkfield microscopy is an inexpensive imaging technique that enables the typical oval-elliptical shape of *Giardia* cysts to be seen as a much brighter structure than the background of the image, as confirmed by the inspection of images published by our research group (Belini *et al.* 2018). This seemingly trivial observation motivated us to choose darkfield imaging because this facilitates cyst detection using standard edge detection techniques (Gonzalez & Woods 2008). As also shown, specific intracellular structures such as nuclei, axostyles, and (retracted) cytoplasm, as assessed using fluorescence microscopy, are much easier to be observed in darkfield than in brightfield images (Belini *et al.* 2018). These structures also help discriminate cysts from morphologically similar non-cyst particles in the images.

Darkfield microscopy was used to acquire images of distilled water that had been contaminated with *Giardia* spp. cysts. The images were acquired using an Olympus BX60 microscope, which was equipped with an Olympus DP74 digital camera (1/1.2″ colour CMOS imaging sensor, $5.86 \times 5.86\ \mu m^2$ pixel size, 1,600 × 1,200 pixels, RGB, tiff, 12-bit). The images were acquired in darkfield by using a condenser for darkfield illumination and an Olympus UPlanFL objective lens (40× magnification and 0.75 numerical aperture). The field of view resulting from the imaging system corresponds to a slide region of $0.27 \times 0.20\ mm^2$.

The prepared microscope slides were placed onto the microscope stage for darkfield imaging. The microscope light intensity was set for proper imaging (~50%), as recommended by the microscope's manufacturer. The aperture of the diaphragm of the illumination source was set close to its minimum so that the image background is dark, and potential artefacts such as dust onto the slide are minimized. The darkfield condenser was positioned just beneath the mounted slide.

The exposure time of the camera was adjusted to 60 ms, so that the internal content of the cysts was not overexposed. Just as for the light intensity, the exposure time was kept constant during the whole experiment. We adjusted the focus manually, so that the cyst wall appeared well defined and brighter than the internal content of the cyst. When needed, we readjusted the focus within different fields of view of the same microscope slides.

The whole microscope slide was screened for *Giardia* spp. cysts. Darkfield images of unlabelled slides were acquired after they had been thoroughly scanned in a systematic scanning pattern following the haemocytometer counting principle (Mundt & Shanahan 2020). The aim is to avoid missing *Giardia* spp. cysts or counted cysts more than once. In this process, the cysts were visually detected based on their typical oval shape, size (8–18 μm long by 5–15 μm wide, U.S. EPA 2012), well-defined wall (under darkfield illumination), and either by the presence of retracted or non-retracted cytoplasm (Belini *et al.* 2018).

2.3. Detection, segmentation, and classification of cyst *candidates*

Figure 1 shows the flowchart of the proposed system. All algorithms were implemented using the MATLAB.

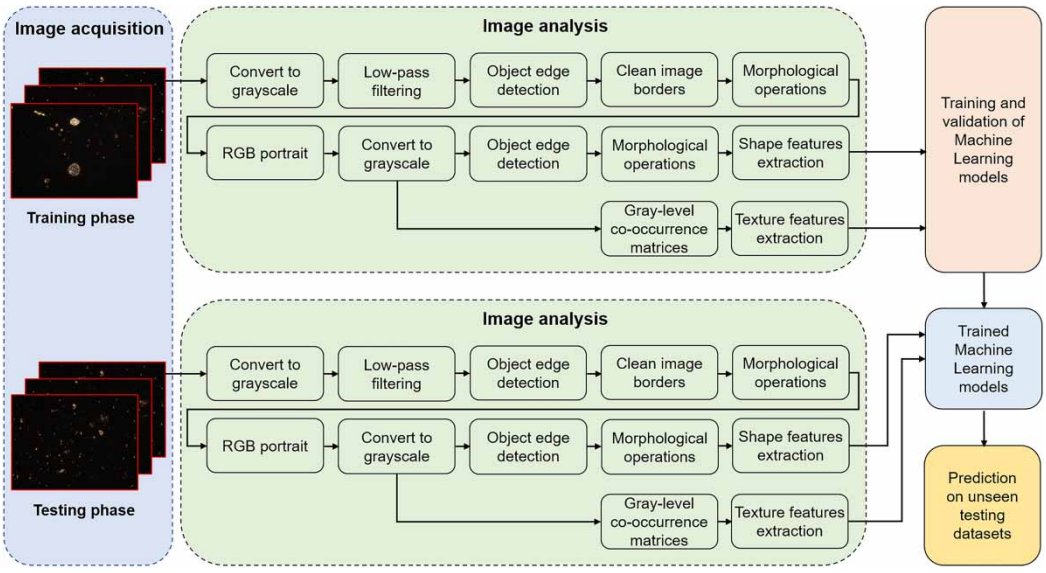


Figure 1 | Flowchart of the automated detection and quantification of *Giardia* spp. cysts in darkfield images using machine learning-based methods.

2.3.1. Object detection

The first step is to detect objects from the darkfield images to create portraits containing cyst candidates. Before detecting objects in the acquired images, we performed the following image preprocessing steps: first, a noise reduction step was carried out using a 7×7 median filter (Gonzalez & Woods 2008) to reduce high-frequency noises.

To find objects in the processed image, a suitable thresholding-based edge detection technique is applied to the gray values in the processed image to create a binary image, depicting detected structures as groups of white pixels against a black background. This detection step aims to identify and extract the edges of objects having a specific brightness level. However, the acquired images may contain cysts exhibiting brightness variances between different fields of view. To address this, the Canny edge detection algorithm (Canny 1986) was chosen because it can adapt to different image characteristics by using two thresholds. The Canny's threshold values were 0.07 and 0.18, respectively, chosen by visually examining the resulting images. Short gaps in the lines of pixels representing the object's edge in the binary image were closed by a morphological dilation operation (Gonzalez & Woods 2008) using 2-pixel line structuring elements followed by hole filling and border image cleaning. Afterwards, the shape of the object is smoothed by applying the morphological opening operation using 5-pixel disk structuring elements (Gonzalez & Woods 2008) in the resulting binary image. Size filtering discards detected objects in the resulting image that do not fall within the expected range of values for cysts (1,000–3,500 pixels, corresponding to 20–70 μm^2). The size of each object in the final image was determined by counting the pixels inside the borders of the objects detected. Each pixel in the acquired image corresponds to a real area given by pixel size over magnification squared:

$$\text{Pixel area} = \frac{5.86\ \mu m\ \times\ 5.86\ \mu m}{1,600} = 0.02\ \mu m^2 \tag{1}$$

2.3.2. Object segmentation

Complex visual data analysis over the entire original (large) image is a time-consuming task. To save computational time, the algorithm initially generates portraits containing cyst *candidates* for further analysis. This involves computing the centroid of the detected objects in the final binary image and then cropping a corresponding portrait image (sized 80×80 pixels) from the original darkfield RGB image, which is subsequently converted to a grayscale image. The corresponding lengths of the portrait are $12\ \mu m \times 12\ \mu m$ by counting the pixels forming the sides of the portrait. Each pixel corresponds to a real length given by the pixel length over optical magnification. Before detecting objects within the portrait, we performed a high-frequency noise reduction step using a 5×5 median filter (Gonzalez & Woods 2008).

Afterwards, the Canny operator (threshold values equal to 0.07 and 0.17, respectively) prepared the portraits for segmentation because the portraits may still contain more than one object. To generate portraits containing only one centralized object inside the portrait, a morphological closing operation (disk-shaped, 5 pixels) is performed, followed by hole filling and portrait border cleaning. After the morphological opening operation (disk-shaped, 7 pixels), only the largest object within the portrait is kept. Afterwards, the centroid of the segmented object is computed, and this information is used to centralize the selected object within the portrait.

2.3.3. Feature extraction

After the automated generation of cyst candidate portraits, a descriptive feature vector was computed for each portrait containing *Giardia* spp. cyst or non-cyst particles. In this respect, each portrait was represented by an 11-dimensional vector, combined of seven morphology-based features (size, perimeter, circularity, eccentricity, solidity, Feret's length, and Feret's width) and four texture features (contrast, correlation, energy, and homogeneity), computed as follows:

$$\text{Circularity} = \frac{4 \cdot \pi \cdot \text{Area}}{(\text{Perimeter})^2} \tag{2}$$

where Area is the size of the object determined as the number of pixels forming the object, and Perimeter is the number of pixels on the object's contour.

$$\text{Eccentricity} = \frac{\text{Distance between the foci of the ellipse}}{\text{Major axis length}} \tag{3}$$

The *solidity factor* (Soltys *et al.* 2005) of each object was computed as the scalar that specifies the proportion of pixels in the convex hull that are also inside the detected object.

The length and the width of the cyst *candidates* were determined as the maximum and minimum Feret's diameters (MathWorks™ 2023), respectively.

The texture-based features were derived from the gray level co-occurrence matrix (GLCM), as introduced by Haralick *et al.* (1973). A GLCM is a square matrix that has a dimension equal to the number of gray levels within portraits (i.e., 80 rows × 80 columns). Each element in the normalized GLC matrix represents the joint probability occurrence $p(i,j)$ of two adjacent pixels i and j located at a distance d apart, along a given direction θ. In this study, $d = 1$ pixel and $\theta = 0, 45, 90, 135°$. The texture features derived from GLC matrices are computed as follows (MathWorks™ 2023):

$$\text{Contrast} = \sum_{i=1}^{\text{columns}} \sum_{j=1}^{\text{rows}} (i-j)^2 . p(i,j) \tag{4}$$

$$\text{Correlation} = \sum_{i=1}^{\text{columns}} \sum_{j=1}^{\text{rows}} \frac{(i-\mu_i).(i-\mu_j).p(i,j)}{\sigma_i.\sigma_j} \tag{5}$$

$$\begin{aligned} \mu_i &= \sum_{i=1}^{\text{columns}} \sum_{j=1}^{\text{rows}} i.p(i,j), \qquad \mu_j = \sum_{i=1}^{\text{columns}} \sum_{j=1}^{\text{rows}} j.p(i,j), \\ \sigma_i &= \sum_{i=1}^{\text{columns}} \sum_{j=1}^{\text{rows}} (i-\mu_i)^2 . p(i,j), \qquad \sigma_j = \sum_{i=1}^{\text{columns}} \sum_{j=1}^{\text{rows}} (j-\mu_j)^2 . p(i,j) \ . \end{aligned} \tag{6}$$

$$\text{Energy} = \sum_{i=1}^{\text{columns}} \sum_{j=1}^{\text{rows}} [p(i,j)]^2 . \tag{7}$$

$$\text{Homogeneity} = \sum_{i=1}^{\text{columns}} \sum_{j=1}^{\text{rows}} \frac{p(i,j)}{1+|i-j|} \tag{8}$$

The final value for each texture feature was determined by averaging the four co-occurrence matrices, each one computed for each direction of adjacency θ.

2.4. Generation of training and validation datasets

We used example images of *Giardia* spp. cysts and non-cysts to train and validate the machine learning-based classifiers. Developing algorithms for automatic detection and segmentation facilitated the process of generating training sets of annotated images from the entire darkfield image. From this process, object portraits (80 × 80 pixels^2) were extracted from original RGB darkfield images. These images were manually annotated as either a *Giardia* spp. cyst or non-cyst.

Aiming at providing a more diverse training image dataset, accounting for, e.g., the variability of the cyst's optical appearance and different imaging conditions, we acquired hundreds of microscopic images of cysts from different microscope slides containing different cyst concentrations.

2.5. Machine learning-based classification

The final step is classification, whereby *Giardia* spp. cysts are identified and counted within a target darkfield image using a trained classifier. To achieve this, we developed a supervised machine learning algorithm to distinguish *Giardia* spp. cysts from non-cysts. The learning technique is based on *bagging* (an acronym from bootstrap aggregating), as introduced by Breiman (1996). This learning system manipulates the training data by randomly replacing the original training data with replicated training sets. These sets are sampled with replacement from the original training data in order to train different base classifiers. In this process, some features

describing the cyst *candidates* may not appear in the training sets, while others appear more than once. Finally, every base-classifier's output is then combined by the majority vote to form the final prediction.

To evaluate the ability of the model to generalize to similar but unseen data in the training/validation stage, 10-fold cross-validation was used for tuning the learning system parameters and also for assessing the overall prediction performance of the (*embedded*) classifier. In the validation method, the training dataset was randomly partitioned into 10 equal-sized subsets. Each subset, in turn, was then used as test data for the classifier from the remaining nine validation subsets. This procedure was repeated 10 times, whereby a different subset was used as the test set, and the performance was computed as the average of the 10 repetitions.

2.6. Performance evaluation

We visually inspected the classified objects within a target darkfield image to verify the correct classification by the trained machine learning algorithm. In this context, we manually classified the classifier outcome as true positive (TP), false positive (FP), or false negative (FN). The classification is based on the following criteria: complete cyst bodies, the largest fragment of over-segmented cysts (i.e., cysts divided into many parts), and one cyst for each case of under-segmentation (more than one cyst detected as one cyst) were classified as TP. FP was non-cyst particles predicted as cysts. FNs were missed cysts in cyst clusters.

To evaluate the classification performance, we computed the SE (i.e., the proportion of cysts that were recognized correctly), the SP (i.e., the proportion of detected objects that were cysts), and accuracy (ACC) as follows (Buggenthin *et al.* 2013):

$$\mathrm{SE} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{9}$$

$$\mathrm{SP} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{10}$$

$$\mathrm{ACC} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN} + \mathrm{FP}} \tag{11}$$

All algorithms developed in this study were run on a standard Desktop (Intel PC, Quad-CPU, 2.66 GHz, 4.0 GB RAM). All parameters involved in both image analysis procedures, and machine learning algorithms were kept fixed for the whole experiment.

2.7. Statistical analysis

We used the average *Giardia* spp. cyst counts obtained from both manual and proposed methods for statistical analyses. To assess the linearity between both methods, a correlation analysis was performed over a range of cyst concentrations. Agreement of the manual and the automated methods regarding cyst concentrations were also assessed using Bland–Altman plots (Bland & Altman 1999). Statistical analyses were carried out using Excel software (Microsoft, USA).

3. RESULTS AND DISCUSSION

3.1. Datasets

Microscopic images of unlabelled cyst suspensions were acquired under darkfield illumination to form image datasets. Figure 2 shows examples of darkfield images.

Microscopic images acquired under darkfield illumination show not only the target microorganism but also the whole background complexity, which appears as much brighter structures than the background itself (Figure 2). As seen in Figure 2(a), most objects in the image are non-cyst particles. Artefacts, such as dust particles on the cover glass (Figure 2(b)), might be highly refractive in darkfield microscopy and, since they are not usually in the focal plane of the target, they appear in the form of bright spotted halos (Ciurte *et al.* 2018). *Giardia* spp. cysts from different suspensions show different optical appearances and retracted cytoplasm (Belini *et al.* 2018). The cyst-free image in Figure 2(d) shows higher numbers of non-cyst particles.

Different densities of non-cyst particles may change the overall brightness of the image, as can be observed by comparing the darker background in Figure 2(c) to that in Figure 2(a) and 2(b). The slight interimage brightness variances and intraimage brightness inhomogeneities (Figure 2(b)) observed in the acquired images can influence the predictions made by the proposed system, since all chosen descriptors are intensity level-based. This could be

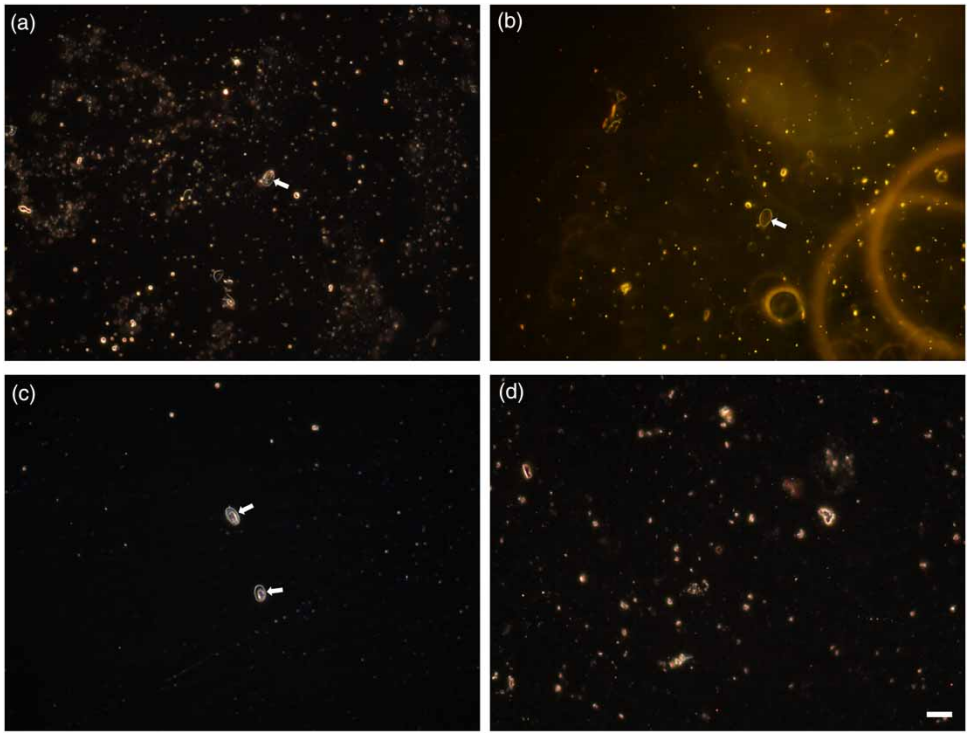


Figure 2 | Examples of original RGB darkfield images acquired from different microscopic slides containing unlabelled suspensions. (a) Image containing one *Giardia* spp. cyst and high density of impurities. (b) Image having *Giardia* spp. cyst and optical artefacts. (c) Image having *G.* spp. cysts and low numbers of impurities. (d) Image containing non-cyst particles only. The arrow points to *Giardia* spp. cyst. Scale bar = 20 μm.

avoided or reduced by using an enhanced image acquisition system and slide cleaning protocols but were omitted to present the proposed system to some degree of real-world interference factors.

3.2. Segmentation and feature extraction

Figure 3 shows an example of the output at each stage of the image analysis algorithm. The image analysis steps, as shown in Figure 1, were performed to generate both training and testing portraits for the machine learning algorithms.

As shown in Figure 3, the input RGB darkfield image is firstly converted into a grayscale image for further processing. After applying a smoothing median filter to reduce high-frequency noise, a Canny edge detector (Gonzalez & Woods 2008) is performed to highlight sharp intensity changes in foreground and background pixels (Figure 3(c)). Appropriate Canny threshold values (empirically determined as 0.07 and 0.18 by inspection in hundreds of images) yield a binary image having detected objects as white structures on a black background (Figure 3(d)). In this process, most pixels belonging to optical artefacts (as shown in Figures 2(b) and 3(a)) are not detected because their pixels have much lower intensities than those belonging to objects imaged in the focal plane.

After determining the centroid of each detected object, this information is used to crop 80×80 pixels^2 portraits of the corresponding object from the original RGB darkfield image. Figure 4 shows examples of segmented objects.

The portraits show images of *Giardia* spp. cysts exhibiting cells of different sizes, shapes, and distinct internal content. In common with most visually inspected portraits, there was the presence of oval-shaped cells and a well-defined bright contour corresponding to the cyst wall (Figure 4(a)). From the point of view of image analysis, the cyst border helps to detect them, while the internal content, characterized by texture features, is useful in helping to distinguish cysts from non-cyst particles that are morphologically similar to the target (Figure 4(b)).

3.3. Machine learning training and validation

After extracting the features for each cyst *candidate* within portraits, we used this information to train the machine learning algorithm to distinguish *Giardia* spp. cysts from non-cyst particles. From 820 acquired darkfield

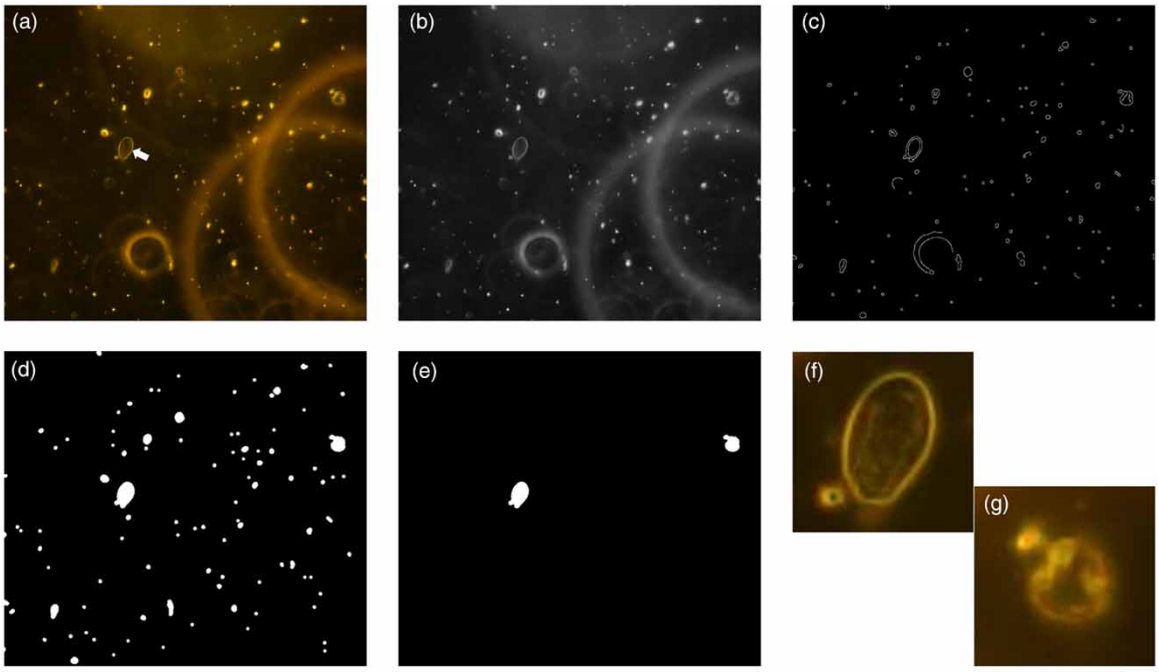


Figure 3 | Image preprocessing and object segmentation to generate portraits of cyst *candidates*. (a) Input RGB image (1,000 × 890 pixels). The arrow points to *Giardia* spp. cyst. (b) Input image converted into the grayscale image. (c) High-frequency noise attenuation followed by contour detection. (d) Binary image resulting from morphological operations, hole filling, and image border cleaning. (e) Resulting image after object size filtering. (f, g) Portraits of the corresponding detected objects (i.e., cyst *candidates*) cropped from the input RGB image.

images, a total number of 681 and 528 *Giardia* spp. cyst and non-cyst portraits, respectively, were generated by the automatic segmentation algorithm. The number of training examples used was balanced to avoid the model from becoming biased towards one class (i.e., *Giardia* spp. cyst or non-cyst). For each object within the portraits, 11 different features were computed and then used for training machine learning-based classifiers. Figure 4 shows examples of both classes of objects. The classifier-type model was based on the decision tree and the classification training used 10-fold cross-validation, achieving an accuracy rate of 88%. This prediction performance means that the classifier can properly identify most of the *Giardia* spp. cysts in the validation step.

3.4. Machine learning-based classifier applied to testing slides

Once we trained and validated the machine learning algorithm, we actively applied it to independent testing slides containing different cyst concentrations and compared its predictions to manual counts carried out in the same darkfield images.

Besides images of *Giardia* spp. cysts, we also acquired an average of 25 images from different fields of view within each microscopy slide (for 0, 100, 250, 500, and 1,000 cysts/mL), all containing a variety of non-cyst particles. The acquisition of cyst-free images aimed to capture the complexity of the background and also helped assess the robustness of the classification method.

For each acquired image, we visually inspected the objects classified as *Giardia* spp. cyst by the algorithm to evaluate the accuracy, SE, and SP of the proposed method. Table 1 shows the results in detail.

Table 1 shows that the proposed system can properly identify and count an average of 87% *Giardia* spp. cysts over a wide range of cyst concentrations. The percentages of correct identifications, despite the varying numbers of non-cyst particles exhibiting a diversity of shapes and sizes, are an indication of the robustness of the classification method. The presence of higher numbers of non-cysts resembling the target in the highest cyst concentrations showed no significant effect on the performance of the classification algorithm, where most *Giardia* spp. cysts were recognized correctly by the algorithm.

The two missed single cysts at the 500 cysts/mL concentration were due to impurities stuck to the detected cysts, making the classifier fail to identify them because this category of object was not considered in the proposal. At the highest cyst concentration, we also acquired images of two cysts touching each other and a four-cyst agglomerate.

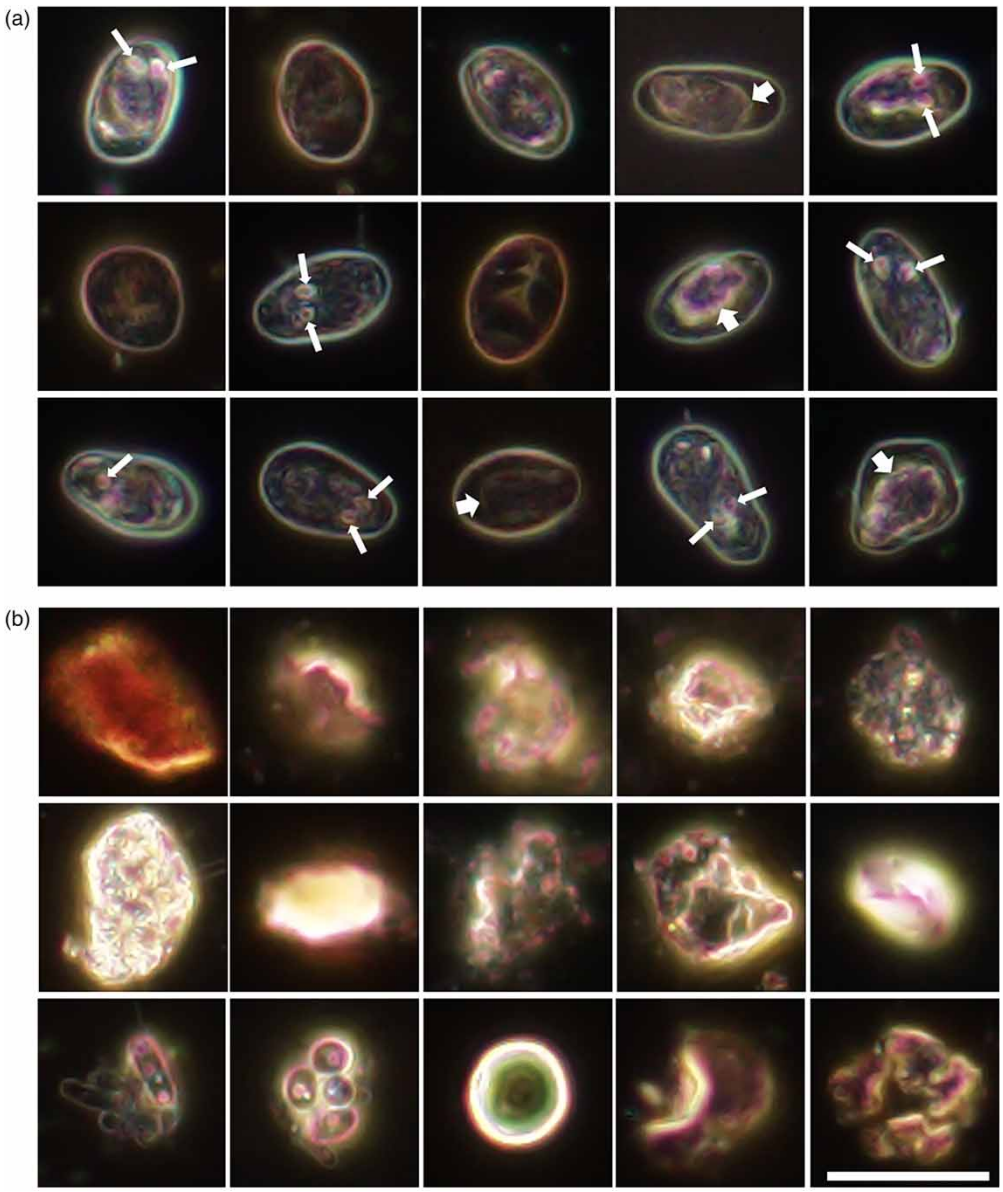


Figure 4 | Microscopy portraits cropped from darkfield images of unlabelled suspensions. (a) *Giardia* spp. cysts. Thin and thick arrows point to structures resembling nuclei and retracted cytoplasm, respectively (Belini *et al.* 2018). The outermost bright structure is the cyst wall. (b) Non-cyst structures morphologically similar to the target. *From left to right in the last row*: two portraits exhibiting clusters of other microorganisms (probably from contamination within different experiments) followed by an image of gas bubble. The micro-particles in the other portraits could not be identified. Scale bar = 10 μm.

In the former case, the algorithm was able to detect and identify two cysts but counted only two cysts in the agglomerate. This result was expected as the detection algorithm was not designed to cope with either cyst-cluster or cyst-agglomerates: the two identified cysts in the agglomerate were segmented correctly by the algorithm because they were just touching each other, thus facilitating their segmentation as two single cysts. On the other hand, the two missed cysts were partially overlapping each other and were detected as a much larger structure than that expected for single cysts and therefore were not classified by the algorithm as the target.

Table 2 shows the relative error observed between automatic and manual counting for the five cyst concentrations analysed. The average value of the error observed in the highest concentrations of cysts is within that acceptable by conventional methods (U.S. EPA 2012) and can be explained by the inability of the detection algorithm to segment clusters of cysts, thus increasing the number of lost cysts (not detected). It should be noted that Method 1623.1 considers recovery rates of *Giardia* spp. cysts from 8 to 100% as an acceptability criterion for modified detection methods, i.e., losses of 92% are considered acceptable.

Table 1 | Classification performance of the proposed system

	Concentration (cysts/mL)			
Parameter	**100 (8 ± 1.0)**	**250 (13 ± 2.5)**	**500 (28 ± 3.5)**	**1,000 (57 ± 4.5)**
Manual cyst count	3	5	14	26
Automated cyst count	3	5	12	21
TP	3	5	12	21
FP	0	0	0	0
FN	0	1	2	5
Sensitivity	100%	83%	86%	81%
Specificity	100%	100%	100%	100%
Accuracy	100%	83%	86%	81%

Manual inspection in the same darkfield images acquired from unseen testing suspensions forms the basis for SE and SP determination. The machine learning algorithm classifies the total number of portraits as the target (i.e., TP + FP). We analysed a total of 28, 29, 44, and 39 images using the proposed system and then manually examined them for the concentrations of 100, 250, 500, and 1,000 cysts/mL, respectively. The actual number of cysts, as determined in two 50 μL microscope wells on fluorescence microscopy, is presented within parentheses as the mean ± SD.

Table 2 | Relative error for cyst counts determined using the proposed method in five independent testing suspensions

Concentration (cysts/mL)	**Manual counting**	**Automated counting**	***n***	**Relative error**
1,000	26	21	39	19%
500	14	12	44	14%
250	5	5	29	0
100	3	3	28	0
0	0	0	27	–

Automated counts are shown as TP + FP, and *n* denotes the number of darkfield images evaluated by both manual and automatic methods.

The data analysis for the five cyst concentrations (0, 100, 250, 500, and 1,000 cysts/mL) resulted in a linear relationship between the automatic and manual counts, as shown in Figure 5.

The strong correlation ($R^2 = 0.99$) between cysts counted by both automated and manual methods demonstrates the efficacy of the proposed method as a tool to estimate numbers of unlabelled *Giardia* spp. cysts directly from darkfield images in the range from 0 to 1,000 cysts/mL.

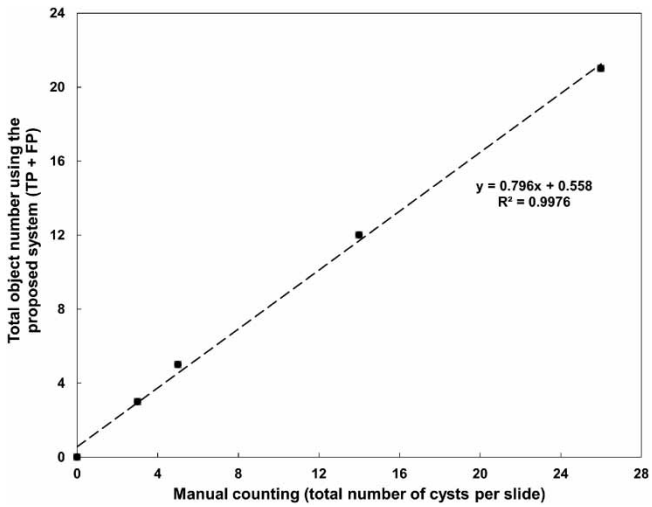


Figure 5 | Correlation between the total cyst count using the proposed system and the manual method in the same darkfield images of the testing suspensions. Each cyst concentration was documented with an average of 25 images. The solid dots represent the experimental data points.

We compared the manual and automated methods using Bland–Altman plots (Figure 6). Although the classification algorithm was designed to prevent missing cysts with minimal FPs to increase the overall slide-level SE (as in the case of 0, 100, and 250 cysts/mL), a low degree of non-agreement of ~1 cyst was observed at higher cyst concentrations. This bias means that the algorithm underestimated the counts when compared to the values from manual counts (as depicted by FNs in Table 2). Two independent factors influence the effectiveness of the proposed system: (1) the presence of impurities stuck to the body of *Giardia* spp. cysts, altering the values of shape- and texture-based descriptors to value ranges outside those expected for the target and (2) the occurrence of cyst clusters or agglomerates, leading to missed constituent cysts within these structures as the detection algorithm is based on standard thresholding techniques that were not designed to deal with more complex structures.

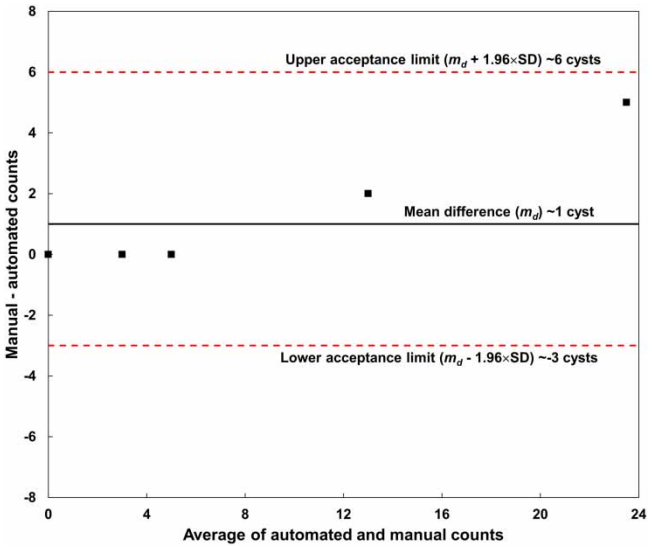


Figure 6 | Bland–Altman comparison between the mean manually determined cyst counts and the automated counts performed in the same darkfield images of the testing suspensions. Each square represents the mean cyst count at each cyst concentration. The black line shows the bias (mean difference) between the automated and manual counts. Dashed red lines show the 95% confidence intervals of manual counts versus automated counts.

The proposed system has some distinct advantages over approved methods using immunofluorescence microscopy that make it an attractive tool for automated detection and enumeration of (unlabelled) *Giardia* spp. cysts in resource-limited settings: (i) darkfield imaging is a low-cost illumination technique that can be easily implemented on standard laboratory microscopes, as demonstrated by Belini *et al.* (2018); (ii) no preparation of samples using either stains or fluorescent dyes is required, thus saving money and time.

The significant costs associated with fluorescence microscopy-based detection methods (e.g., U.S. EPA 2012) make them difficult to be performed in less-developed countries. Table 3 shows comparative costs of implementing both the proposed system and the 1623.1 method (U.S. EPA 2012). The cost associated with the purchase of a computer to register either fluorescence or darkfield images was not included as standard desktop computers can be used for this purpose, as well as to run image analysis and machine learning algorithms.

The estimated values do not include costs associated with electricity, maintenance, and also the basic staff qualification to prepare samples and operate the equipment. As a matter of fact, by eliminating the need for specific components for fluorescence microscopy such as optical filters, UV lamp, and its power supply, the operating complexity and costs associated with the maintenance of darkfield microscopes might even be reduced. Darkfield imaging-related costs can be further reduced by converting a standard laboratory microscope into a darkfield imaging system by using low-cost custom-made darkfield setups, as described in detail by Belini *et al.* (2018).

In addition to markedly reducing costs (Table 3), the proposed method quantifies suspensions without modification, as it does not require any chemicals. This enables the quantitative transfer of counted organisms to test volumes of water, as described in Ongerth & Saaed (2017).

Table 3 | Comparison costs accounting main physical components to implementing the method 1623.1 and the proposed system

Input	Standard method using fluorescence microscopy[a]	Proposed method using darkfield microscopy[a]
Microscope frame	$4,200	$1,100
Mercury lamp	$400	Not required
Exciter, dichroic, and emission filters	$5,100	Not required
Power supply for UV lamp	$1,300	Not required
Objective lens	$1,200	$1,200
CMOS camera	$8,700	$8,700
Total equipment cost	$20,900	$11,000

[a]ASTEL Laboratory Equipment and microscopes. https://astelbtu.com.br/. Brazilian quote in June 2023.

Compared with the standard method, the proposed system is less labour-intensive; it only requires preparing the microscopy slide with the suspension to be analysed. Once the prepared slide is placed onto the microscope stage, manual focusing, slide scanning, and image acquisition were performed by one of the authors, taking an average of 1.5 h per slide. Afterwards, a computer automatically analysed an average of 40 images within ~30 s to produce an estimate of the number of cysts for a single sample. In contrast, ~4 h are required for labelling the *Giardia* spp. cysts with fluorescent dyes (U.S. EPA 2012) followed by 30–45 min, so that an expert on fluorescence microscope can complete the counts (Keserue *et al.* 2012). However, at the current state of the proposed method, cysts cannot be distinguished from another genus of protozoan such as *Cryptosporidium parvum*. This would certainly be interesting to be explored in future studies. When needing this information, one still requires standard methods that involve antibody-specific fluorescent labelling.

The proposed method does have some limitations: if cysts physically overlap each other, our method could not separate them properly, producing underestimates. Although the presence of clustered cysts is not common in environmental samples, purified suspensions often contain agglomeration of cysts. In addition, although the method performed well on images exhibiting different complexities, such as optical artefacts and varying densities of non-cyst particles, we did not robustly validate it in more complex water matrices because the number and diversity of non-cyst particles are limited in the present conditions. As for the former limitation, a promising avenue is to extend the segmentation algorithm to include, for instance, methods based on the morphological watershed (Meyer & Beucher 1990) to separate touching cells in microscopy images as published elsewhere (e.g., Chiang *et al.* 2018; Belini *et al.* 2021). We prepared the suspensions to eliminate potential interferences from, e.g., organic matter, metals, and other microorganisms that naturally exist in environmental water matrices, aiming to understand the principles, potential, and limitations of the proposed system. The full potential of the method in assessing water samples from natural matrices will be explored in further studies by incorporating images of particulates found in water samples into the training phase of the machine learning models.

By using the standard method 1623.1 (U.S. EPA 2012), the estimated cost per sample due to fluorescent dyes was ~$120 (Brazilian quote in June 2023), whereas no cost was associated with the proposed method as no chemicals were required. This cost refers to the Merifluor® (Meridian Bioscience, Inc.) kit and Fluoroshield™ with the DAPI (Sigma-Aldrich®) solution, computed per three 50 µL aliquots (U.S. EPA 2012).

We believe that when using microscopes featured with image processing-based autofocusing systems (e.g., Zhang & Yao 2021) integrated into the process of motorized microscope stages, the analysis of a large routine batch of suspensions could benefit from a fully automated, cost-effective, fast, reliable, label-free, and expert-independent *Giardia* spp. cyst counting system.

4. CONCLUSIONS

In this study, we developed a machine learning-based approach and evaluated it to identify and count unlabelled *Giardia* spp. cysts in darkfield microscopic images. Approved standard methods for detecting

Giardia spp. cysts are expensive, laborious, and time-consuming procedures based on sample preparation that require specific fluorescent dyes and trained technicians to conduct and interpret analyses using fluorescence microscopy. The developed system demonstrated the general feasibility of label-free detection and enumeration for *Giardia* spp. cysts based on darkfield imaging, shape and textural features, and machine learning techniques. The results showed satisfactory performance exhibiting an average identification ratio of 88%, thus proving the applicability of the developed system to perform automatic counting. By using the proposed system, we simultaneously overcame the cost limitation and lack of trained personnel to analyse *Giardia* spp. cyst suspensions. We believe that the *Giardia* spp. cyst detection system proposed herein has the potential to be a promising alternative method for fluorescence microscopy in approved standard methods, thus allowing the routine analysis of suspensions in low-income countries as the proposed system is expert-independent and free of any kind of consumables. By eliminating the subjectivity inherent to manual counting using standard methods, the application of the proposed method could enable more accurate estimates in water treatment-related research involving, for instance, the assessment of filters efficiency for pathogen removal.

ACKNOWLEDGEMENTS

We acknowledge the Biological Processes Laboratory, Department of Hydraulics and Sanitation, São Carlos School of Engineering, University of São Paulo, for their support with the darkfield imaging presented herein.

AUTHOR CONTRIBUTION

V.L.B. conceptualized the study, developed the methodology, conducted the software investigation, performed validation, and contributed to the original draft and editing. N.M.N.F. contributed to the conceptualization, methodology, and reviewed and edited the writing. L.P.S.-P. obtained funding, provided resources, validated the study, and reviewed and edited the writing. All authors have read and approved the final manuscript.

FUNDING

This work was partially supported by The Royal Society (ICA/R1/201373 – International Collaboration Awards 2020) and by the National Council for Scientific and Technological Development (CNPq-Brazil, processes n° 308070/2021-6 and 442074/2023-9).

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

Belini, V. L., Freitas, B. L. S., Sabogal-Paz, L. P., Branco, N. & Bueno Franco, R. M. 2018 Label-free darkfield-based technique to assist in the detection of *Giardia* cysts. *Water Air Soil Pollut.* **229**, 195. https://doi.org/10.1007/s11270-018-3834-x.

Belini, V. L., Junior, O. M., Ceccato-Antonini, S. R., Suhr, H. & Wiedemann, P. 2021 Morphometric quantification of a pseudohyphae forming *Saccharomyces cerevisiae* strain using *in situ* microscopy and image analysis. *J. Microbiol. Methods* **190**, 106338. doi:10.1016/j.mimet.2021.106338.

Belini, V. L., de Melo Nasser, F. N., Garcia, L. A. T., da Cunha, M. J. R. & Sabogal-Paz, L. P. 2022 Label-free detection and enumeration of *Giardia* cysts in agitated suspensions using *in situ* microscopy. *J. Microbiol. Methods* **199**, 106509. doi:10.1016/j.mimet.2022.106509.

Bland, J. M. & Altman, D. G. 1999 Measuring agreement in method comparison studies. *Stat. Methods Med. Res.* **8**, 135–160. doi:10.1177/096228029900800204.

Breiman, L. 1996 Bagging predictors. *Mach. Learn.* **24**(2), 123–140. https://doi.org/10.1007/BF00058655.

Buggenthin, F., Marr, C., Schwarzfischer, M., Hoppe, P. S., Hilsenbeck, O., Schroeder, T. & Theis, F. J. 2013 An automatic method for robust and fast cell detection in bright field images from high-throughput microscopy. *BMC Bioinform.* **14**, 297. https://doi.org/10.1186/1471-2105-14-297.

Canny, J. A. 1986 Computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **8**(6), 679–698. doi:10.1109/TPAMI.1986.4767851.

Chiang, P. J., Wu, S. M., Tseng, M. J. & Huang, P. J. 2018 Automated bright field segmentation of cells and vacuoles using image processing technique. *Cytometry A* **93**(10), 1004–1018. doi:10.1002/cyto.a.23595.

Ciurte, A., Selicean, C., Soritau, O. & Buiga, R. 2018 Automatic detection of circulating tumor cells in darkfield microscopic images of unstained blood using boosting techniques. *PLoS One* **13**(12), e0208385. doi:10.1371/journal.pone.0208385.

Daniels, M. E., Smith, W. A. & Jenkins, M. W. 2018 Estimating *Cryptosporidium* and *Giardia* disease burdens for children drinking untreated groundwater in a rural population in India. *PLoS Negl. Trop. Dis.* **12**(1), e0006231. https://doi.org/10.1371/journal.pntd.0006231.

Efstratiou, A., Ongerth, J. & Karanis, P. 2017 Evolution of monitoring for *Giardia* and *Cryptosporidium* in water. *Water Res.* **123**, 96–112. https://doi.org/10.1016/j.watres.2017.06.042.

El Mallahi, A., Minetti, C. & Dubois, F. 2013 Automated three-dimensional detection and classification of living organisms using digital holographic microscopy with partial spatial coherent source: Application to the monitoring of drinking water resources. *Appl. Opt.* **52**, A68–A80. doi:10.1364/AO.52.000A68.

Fava, N. M. N., Silva, K. J. S., Snelling, W. J., Ternan, N. G., Dooley, J. S. G. & Sabogal-Paz, L. P. 2021 Does each bead count? A reduced-cost approach for recovering waterborne protozoa from challenge water using immunomagnetic separation. *J. Water Health* **19**(3), 436–447. https://doi.org/10.2166/wh.2021.005.

Feng, C., Xu, Z., Li, Y., Zhu, N. & Wang, Z. 2021 Research progress on the contamination status and control policy of *Giardia* and *Cryptosporidium* in drinking water. *J. Water Sanit. Hyg. Dev.* **11**(6), 867–886. https://doi.org/10.2166/washdev.2021.151.

Gonzalez, R. & Woods, R. 2008 *Digital Image Processing*, 3rd edn. Pearson Education, Upper Saddle River, NJ, USA.

Göröcs, Z., Baum, D., Song, F., de Haan, K., Koydemir, H. C., Yunzhe, Q., Cai, Z., Skandakumar, T., Peterman, S., Tamamitsu, M. & Ozcan, A. 2020 Label-free detection of *Giardia lamblia* cysts using a deep learning-enabled portable imaging flow cytometer. *Lab Chip* **20**, 4404–4412. https://doi.org/10.1039/D0LC00708K.

Haralick, R. M., Shanmugam, K. & Dinstein, I. 1973 Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* **3**(6), 610–621. doi:10.1109/TSMC.1973.4309314.

Keserue, H. A., Füchslin, H. P., Wittwer, M., Nguyen-Viet, H., Nguyen, T. T., Surinkul, N., Koottatep, T., Schürch, N. & Egli, T. 2012 Comparison of rapid methods for detection of *Giardia* spp. and *Cryptosporidium* spp. (oo)cysts using transportable instrumentation in a field deployment. *Environ. Sci. Technol.* **46**(16), 8952–8959. doi:10.1021/es301974m.

Koydemir, H. C., Gorocs, Z., Tseng, D., Cortazar, B., Feng, S., Chan, R. Y., Burbano, J., Mcleod, E. & Ozcan, A. 2015 Rapid imaging, detection and quantification of *Giardia lamblia* cysts using mobile-phone based fluorescent microscopy and machine learning. *Lab Chip* **15**(5), 1284–1293. https://doi.org/10.1039/C4LC01358A.

Luo, S., Nguyen, K. T., Nguyen, B. T. T., Feng, S., Shi, Y., Elsayed, A., Zhang, Y., Zhou, X., Wen, B., Chierchia, G., Talbot, H., Bourouina, T., Jiang, X. & Liu, A. Q. 2021 Deep learning-enabled imaging flow cytometry for high-speed *Cryptosporidium* and *Giardia* detection. *Cytometry A* **99**(11), 1123–1133. https://doi.org/10.1002/cyto.a.24321.

Maciel, P. M. F. & Sabogal-Paz, L. P. 2016 Removal of *Giardia* spp. and *Cryptosporidium* spp. from water supply with high turbidity: Analytical challenges and perspectives. *J. Water Health* **14**(3), 369–378. https://doi.org/10.2166/wh.2015.227.

MathWorks™. 2023 *Image Processing Toolbox (TM)*: *User's Guide* (R2021b). www.mathworks.com/help/images/ref/regionprops.html (accessed 15 February 2023).

McGrath, J. S., Honrado, C., Spencer, D., Horton, B., Bridle, H. L. & Morgan, H. 2017 Analysis of parasitic protozoa at the single-cell level using microfluidic impedance cytometry. *Sci. Rep.* **7**, 2601. https://doi.org/10.1038/s41598-017-02715-y.

Meyer, F. & Beucher, S. 1990 Morphological segmentation. *J. Vis. Commun. Image Represent.* **1**, 21–46. https://doi.org/10.1016/1047-3203(90)90014-M.

Mudanyali, O., Oztoprak, C., Tseng, D., Erlinger, A. & Ozcan, A. 2010 Detection of waterborne parasites using field-portable and cost-effective lens free microscopy. *Lab Chip* **10**, 2419–2423. https://doi.org/10.1039/C004829A.

Mundt, L. & Shanahan, K. 2020 *Graff's Textbook of Urinalysis and Body Fluids*. Jones & Bartlett Learning, Burlington, MA, USA.

Nakarmi, S., Pudasaini, S., Thapaliya, S., Upretee, P., Shrestha, R., Giri, B., Bhakta Neupane, B. & Khanal, B. 2023 Submitted deep-learning assisted detection and quantification of (oo)cysts of *Giardia* and *Cryptosporidium* on smartphone microscopy images. arXiv e-prints. doi:10.48550/arXiv.2304.05339.

Ongerth, J. E. & Saaed, F. 2017 *Cryptosporidium* oocysts and *Giardia* cysts – A practical and sensitive method for counting and manipulating small numbers. *Acta Trop* **169**, 119–121. doi:10.1016/j.actatropica.2017.02.008.

Shrestha, R., Duwal, R., Wagle, S., Pokhrel, S., Giri, B. & Neupane, B. B. 2020 A smartphone microscopic method for simultaneous detection of (oo)cysts of *Cryptosporidium* and *Giardia*. *PLoS Negl. Trop. Dis.* **14**(9), e0008560. https://doi.org/10.1371/journal.pntd.0008560.

Soltys, Z., Orzylowska-Sliwinska, O., Zaremba, M., Orlowski, D., Piechota, M., Fiedorowicz, A., Janeczko, K. & Oderfeld-Nowak, B. 2005 Quantitative morphological study of microglial cells in the ischemic rat brain using principal component analysis. *J. Neurosci. Methods* **146**, 50–60. doi:10.1016/j.jneumeth.2005.01.009.

United States Environmental Protection Agency – USEPA 2012 Method 1623.1 *Cryptosporidium* and *Giardia* in Water by Filtration/IMS/FA. EPA 816 – R-12-001. Office of Water (MS-140). Environmental Protection Agency, Washington, DC.

Zhang, H. & Yao, J. 2021 Automatic focusing method of microscopes based on image processing. *Math. Probl. Eng.* 8243072. https://doi.org/10.1155/2021/8243072.

First received 22 January 2024; accepted in revised form 28 March 2024. Available online 9 April 2024