2021

Facilitating Analysis of *Toxoplasma gondii* Bradyzoite Metabolic Activity via Image Processing and Multivariate Logistic Regression for High Throughput Classification of Mitochondrial Morphologies

Brooke Place

University of Kentucky, place.brooke.c@gmail.com

Author ORCID Identifier:

🆔 https://orcid.org/0000-0001-6247-3661

Digital Object Identifier: https://doi.org/10.13023/etd.2022.220

Right click to open a feedback form in a new tab to let us know how this document benefits you.

Recommended Citation

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Brooke Place, Student

Dr. Abhijit Patwardhan, Major Professor

Dr. Sridhar Sunderam, Director of Graduate Studies

FACILITATING ANALYSIS OF *TOXOPLASMA GONDII* BRADYZOITE
METABOLIC ACTIVITY VIA IMAGE PROCESSING AND MULTIVARIATE
LOGISTIC REGRESSION FOR HIGH THROUGHPUT CLASSIFICATION OF
MITOCHONDRIAL MORPHOLOGIES

_____

THESIS

_____

A thesis submitted in partial fulfillment of the
requirements for the degree of Master of Science in Biomedical Engineering in the
College of Engineering
at the University of Kentucky

By

Brooke Place

Lexington, Kentucky

Director: Dr. Abhijit Patwardhan, Professor of Biomedical Engineering

Lexington, Kentucky

2021

ABSTRACT OF THESIS

FACILITATING ANLAYSIS OF *TOXOPLASMA GONDII* BRADYZOITE
METABOLIC ACTIVITY VIA IMAGE PROCESSING AND MULTIVARIATE
LOGISTIC REGRESSION FOR HIGH THROUGHPUT CLASSIFICATION OF
MITOCHONDRIAL MORPHOLOGIES

Approximately a third of the population worldwide is chronically infected by the parasite *Toxoplasma gondii*. During chronic infection the parasite resides within tissue cysts as the poorly understood bradyzoite form. Observation of these bradyzoites via microscopic imaging within tissue cyst purified from infected mouse brains has shown metabolic activity with heterogeneous replication potential. With fluorescence microscopy imaging targeting the parasite's actively respiring mitochondria, the parasite's metabolic state can be further investigated as the morphology of mitochondria can be associated with specific physiologic states. However, manually classifying mitochondrial morphologies from these images can be tedious and prone to error as the bradyzoites within cysts can number into thousands. Towards this end, computer based tools were developed to facilitate and automate the identification and classification of the parasite's mitochondrial morphologies to aid the study of bradyzoite biology. The developed image processing based program assists the manual classification of mitochondrial morphologies by trained operators while collecting features and statistics from the manual classification of shapes. The manual classifications, obtained from a subset of images was used as the gold standard to train a multivariate logistic regression algorithm. Results from the machine learning based automatic classification, trained on a total of 1,138 discrete mitochondrial objects from 5 images of individual cysts, and tested on 5 different images of individual cysts showed that from a total of 927 discrete mitochondrial objects that were identified, an average overall accuracy of 82% was obtained in predicting the mitochondrial morphology as belonging to one of the five predefined classes of Blobs, Tadpoles, Donuts, Arcs, or Other. The Lasso morphology, typically associated with actively growing parasites was not observed with significant frequency. The majority of detected objects, 55%, of the objects from the 5 images used in the training set were identified as belonging to the Blob morphology which is presumptively associated with low bradyzoite metabolic activity. There was a lower incidence of Tadpole and Arc morphologies which are associated with more active parasites. The performance of the trained machine learning algorithm resulted in a high degree of confidence in the prediction of Blobs with an average F score of 0.91 and Arcs with an average F score of 0.73 which added together make up a majority of morphological classes present in most bradyzoites, 85% of the objects detected in the training set. These results indicate that the developed approach can advance the investigation of bradyzoite biology by allowing for a higher throughput of images analyzed and thus potentially assist in the evaluation of the efficacy of potential drug treatments.

KEYWORDS: image processing, machine learning classification, bradyzoite biology

Brooke Place

11/01/2021

FACILITATING ANALYSIS OF *TOXOPLASMA GONDII* BRADYZOITE
METABOLIC ACTIVITY VIA IMAGE PROCESSING AND MULTIVARIATE
LOGISTIC REGRESSION FOR HIGH THROUGHPUT CLASSIFICATION OF
MITOCHONDRIAL MORPHOLOGIES

By
Brooke Place

Abhijit Patwardhan
Director of Thesis

Sridhar Sunderam
Director of Graduate Studies

11/01/2021
Date

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

Technology advancements continue to aid in the expansion of our knowledge in many fields due to their broad application, one such area is parasitology allowing previously inaccessible aspects to be observed in greater detail. The work described in this thesis focused on the integration of fluorescence microscopy imaging methods along with the automation of image processing and machine learning to better facilitate analysis of the biology of *Toxoplasma gondii*, specifically bradyzoite mitochondrial morphology.

Toxoplasmosis is an infectious disease caused by the protozoan parasite *Toxoplasma gondii* (*T. gondii*) which chronically infects a third of individuals globally [1]. In most healthy individuals the infection is asymptomatic and readily controlled by a competent immune system[1]. However, the immune system does not rid the individual of the parasite instead the parasite enters a second life cycle stage defined by the tissue cyst composed of slow growing bradyzoites[1]. Tissue cysts develop primarily in neuronal and muscle tissues, wherein a life-long infection is established [1]. These tissue cysts retain the potential to reactivate either stochastically or in response to yet to be defined triggers initiating a round of acute infection. This event in the immunocompromised, especially in the case of HIV-AIDS, results in symptomatic toxoplasmosis manifesting primarily as toxoplasmic encephalitis that is lethal if not treated [1]. There is significant motive for further research on *T. gondii* due to the relatively high prevalence of the parasite, the potential lethality of toxoplasmosis, and the outstanding questions on how the parasites biology enables it to persist within humans and other intermediate hosts.

Little is known about the physiology of bradyzoites, which were long considered to be non-replicative and metabolically quiescent [2]. Prior work demonstrated that the bradyzoites within tissue cysts in vivo exhibited a surprising level of metabolic activity including mitochondrial activity and the ability to replicate [3]. This finding exposed the need for further studies on the physiology of bradyzoites, which has necessitated the development of new tools including the physiology based imaging approaches that are the focus of this work.

The parasite's development cycles and its growth during chronic infection are key indicators of its capability to persist. Inherent to parasite viability and the potential for growth is a requirement for energy and metabolites derived from the activity of the single parasite mitochondrion. Ongoing studies revealed a surprising level of heterogeneity in the level of activity and morphology of mitochondria and mitochondrial elements within tissue cysts purified from infected mouse brains [3]. It can be reasoned that these differences reflect underlying physiological states that can inform on the metabolism of individual bradyzoites and the populations that are housed within tissue cysts. As has been observed in tachyzoites, the varying mitochondrial morphologies are plastic and reflect on the parasites physiological state [4, 5]. The morphologies resembling that of a lasso (ring) and arcs (linear and curvilinear forms) can be associated with more active mitochondria compared to that of a tadpole (also termed sperm-like) and blobs (also termed collapsed and puncta) morphological forms [4, 5]. Under conditions of nutrient stress and drug pressure the mitochondrion can be fragmented into multiple punctate forms [4, 6]. This work aims to assist in the quantification and classification of mitochondrial morphologies imaged following accumulation of the

mitochondrion targeting reagent MitoTracker, as the observed morphologies can inform about the parasite's physiological state [4, 5]. MitoTracker accumulates within actively respiring mitochondria driven by the electrochemical membrane potential [7]. Once in the mitochondrion, this fluorescent dye is captured and can be detected using fluorescence microscopy[7]. Fluorescence microscopy imaging allows us to capture targeted entities, which are the respiring/active mitochondria marked by MitoTracker dye and associated nuclei using DNA dye 4',6-Diamidino-2-phenylindole DAPI (Invitrogen), illuminating the mitochondria and nuclei respectively within an optical section of tissue cyst following deconvolution. The profiles of labeled mitochondria viewed in the context of the nuclear staining using DAPI or other UV excitable DNA-specific dyes (Hoescht 33342), provides a spatial relationship of individual bradyzoites within labeled cysts which was noted to be potentially helpful in classifying the mitochondrial morphology by the experts involved in this study. The reason being that the mitochondrion of actively growing tachyzoites tend to surround the nucleus with a lasso shape [5]. The MitoTracker dye reveals the level of metabolic activity within the encysted bradyzoites as it accumulates selectively in active mitochondria and appear as objects with distinct morphologies within imaged encysted bradyzoites. Mitochondrial morphology has been observed to transition during different lifecycles, such as observed during tachyzoite mitochondrial autophagy (mitophagy) [4], as well as in the course of the active growth cycle [5].   Notably, the dynamics of mitochondrial morphology have not been investigated in bradyzoites. Observing bradyzoite mitochondria in greater detail can help to narrow the current gap in knowledge on the parasites 'clinically important life cycle stage', as the majority of active infections in immune compromised individuals are due

to the reactivation of a previously acquired (can be decades earlier) chronic infection [8]. This is particularly relevant as the only drugs exhibiting any efficacy in this hard to treat form are believed to target the parasite mitochondrion [9, 10]. A true longitudinal in vivo study is not feasible as the images are taken of extracted and purified tissued cyst therefore capturing specific time points throughout the progression of chronic toxoplasmosis can help in the derivation of the parasite's growth and development cycles. Once established the observed effect of drug treatment will provide unprecedented mechanistic insight into the variable efficacy of existing drugs against encysted parasites.

The observation of the subcellular entities has not been performed in detail throughout the parasite's lifecycle. This is partially due to the high demand of time and trained specialist it would require in performing this manual process of observing individually sampled images. Technology advances have helped to continue to overpass the technical challenges that have halted even some of the 'most basic studies on bradyzoites and tissue cysts in vivo' [8]. With the application of automated image processing and machine learning this manual process has the potential to be automated. In this study the input of trained specialist was used to help guide the development of a new tool to aid the analysis of the mitochondria within the tissue cysts. Quantitative analysis of the mitochondrial morphological diversity has direct implications on the assignment of physiological activity and thus druggability.

CHAPTER 2. BACKGROUND

Previous work from our lab developed image processing based tools to quantify bradyzoite nuclei, via DAPI (Invitrogen) stained microscopy images, to derive the number of parasites within tissue cysts (the packing density) at the chronic stage of toxoplasmosis infection [3]. The study illustrated that the bradyzoite burden within cyst over time displayed a high degree of heterogeneity challenging the concept of bradyzoites being predominantly dormant entities [3]. The wide range in bradyzoite burdens observed in that study could be greater than that of the effects of drugs that exhibit partial efficacy[8]. This previously unrecognized heterogeneity in organization and function of tissue cysts leads to further questions on what contributes to this complexity in bradyzoite physiology.

This prior work found that the packing density defined by varying bradyzoite counts within equal sized tissue cysts could differ by over two fold [8]. This work also derived that the level of metabolic function, associated with bradyzoites in vivo was much greater than what was long thought to be the case. This included the capacity to replicate within the tissue cyst, highlighting the need to reexamine the prevailing dogma's surrounding bradyzoite biology [3]. With the number of bradyzoites ranging into the hundreds for a single cyst there is the opportunity to observe and gather a large amount of useful information on bradyzoites to better understand their biology. At the same time, the large number of the bradyzoites also highlight the need for and the utility of computer based approaches in these studies.

The mitochondrion of *T. gondii* appears highly sensitive to nutrient limitation by exhibiting progressive changes in morphology in studies using cell culture based

tachyzoite infections [4]. The mitochondrial morphology of *T. gondii* has shown to reflect its physiological state [4, 5]. The pattern of morphologies progresses from loop to a more linear shape with a bulbous end (tadpole), to an arc or dumbbell shape, to a punctate shape [5]. The normal shape for active tachyzoites has been described as ring or loop shaped [4, 5], and a tachyzoite will possess a single large mitochondrion [11]. In addition, under conditions of stress and drug pressure, the mitochondrion can be fragmented into multiple puncta [4, 6]. In contrast, little is known about the morphology of mitochondria within encysted bradyzoites, suggesting that the distribution of morphologies could inform on the specific physiological state of the individual bradyzoites.

Image analysis programs have been developed to aid the analysis of pathogen recognition and host response via protein recruitment based on machine learning and deep learning [12] and the analysis of the tissue cyst morphology along with bradyzoite count within cyst was evaluated by Bauman et al. by looking at diameter, circularity, particle (bradyzoite count), fractal dimensions, lacunarity, and packing density of tissue cysts themselves [13]. Few studies have looked at the chronic phase of in vivo derived tissue cyst from the context of individual bradyzoites, and yet analysis at the level of parasites organelles specifically mitochondria have not been performed in the manner approached here.

With the use of added imaging techniques and the use of multiple stains in ongoing animal studies of exposure to *T. gondii* in our collaborator's (Dr. Sinai) laboratory, the diversity of the irregularly shaped parasite structures can be enumerated and explored in greater detail. As mentioned above, MitoTracker (Invitrogen) can be

used to illuminate active mitochondria. Using mitochondria as a measurable substitute for bradyzoite metabolic status we can better understand its growth in terms of expansion and replication.

This study uses image processing based methods to detect and segment fluorescently labelled mitochondrial entities and nuclei within microscopy images to localize distinct objects within the images and then uses a supervised machine learning approach to identify these objects morphologies. The images were obtained from mice infected with *T. gondii* with varying parasite burdens. These animal studies were conducted in Dr. Sinai's laboratory and approved by the Institutional Animal Care and Use Committee at the University of Kentucky. With hundreds of mitochondria illuminated in a single fluorescent microscopy image of a tissue cyst the end goal was to be able to identify each individual mitochondrion and classify them as belonging to one of the predefined morphologies. Image processing was used to extract features of localized objects and manual classification, performed by expert users, was used to obtain gold standard classification which was used to train the supervised machine learning approach.

There are currently a number of supervised classification machine learning approaches that could be implemented to best predict which morphological class the set of extracted features are best associated with based on the user labeled data set. The classification models include simple thresholding, regressions, support vector machines, and neural networks [14]. The approach chosen to implement in this thesis was logistic regression as it is a widely used technique that is relatively simple to implement yet robust.

CHAPTER 3. Methods

The following section describes how the developed computer program performs image processing for object detection, determines features of the detected objects and collects user input on the object's morphological classification. The collected information is then used in training a machine learning algorithm automating the time-consuming process that a trained user would perform of manually identifying and classifying individual objects and their shapes within fluorescence based images of tissue cysts. MATLAB's App Designer, Image Processing, Computer Vision, and Optimization toolboxes were used in the creation of a custom designed code that incorporates the image processing workflow and a graphical user interface (GUI) for a customizable and user-friendly data collection method.

3.1   Images

Tissue cysts of the Type II ME49 strain of chronically infected CBA/J mice were used to evaluate the chronic stage of infection. CBA/J mice show a high volume of cyst development within the brain [15]. The cysts were extracted from the central nervous system post infection and purified using a Percoll gradient as described by Watts et al [16]. The samples were then labeled for DNA and actively respiring mitochondria with DAPI (Invitrogen) and MitoTracker red (Invitrogen) dyes. Prior to being imaged the cysts were fixed in 3% paraformaldehyde, deposited on glass slides and cover-slipped. A Zeiss Axioplan microscope with 100x/1.4 numerical aperture objective was used with a grayscale AxioCam MRM digital camera. A single z- plain image was selected from each cyst to be analyzed based on having the widest diameter, hence having the potential to represent the maximum number of bradyzoites of all the z-plain images for that cyst.

Each image represents a slice of a three-dimensional section of the cyst, because of the z-plain slicing, thus there was the potential of having overlapping objects. The images were deconvolved using an iterative algorithm (AxioVision Deconvolution Suite; Zeiss). The deconvolved images were used for object detection and classification as described below. The optical magnification used provided images with a 15.5 pixel per micron scale. The resulting images, one labeled DAPI and one labeled Mito, were 8-bit grayscale image files in TIF format. An example of the DAPI and Mito image pair, obtained from the same cyst, is shown in Figure 1.
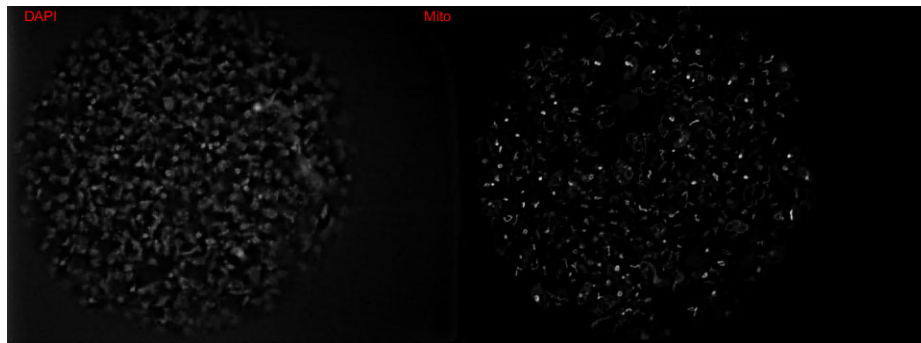
Figure 1. A representative DAPI (left) and Mito (right) image.

3.2   Image Processing

The 8-bit grayscale image undergoes linear intensity scaling, adjusting the pixel intensity range to that of the full range of resolution, i.e. 0 to 255. Enhancing the contrast between the background pixels and the target objects' pixels was then performed by utilizing top hat and bottom hat filters [17]. A disc shaped structuring element was used in implementing the erosion and dilation processes of the filters, the radius of the structuring element used in the top hat filter was 10 pixels while the radius used for the bottom hat filter was 3 pixels.

The two filters were used in conjunction as described in Equation 1 and an example of this filtering is depicted in Figure 2. The foreground was emphasized by the

9

application of the top hat filter to the original image, I, which made the bright spots brighter. The objects' perimeter was enhanced by the subtraction of the scaled bottom hat filtered image resulting in the contract enhanced filtered image, J.

$$[J] = [Tophat(I)] + [I] - 10 * [Bottomhat(I)] \ Equation \ 1. \ Contrast \ Enhancement$$
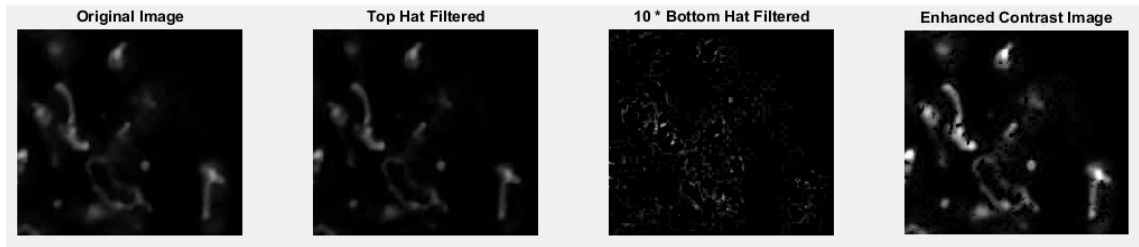
Figure 2. Depiction of the filters used, starting with a zoomed in section of the original Mito image [left], the top hat filtered image [left center], the bottom hat filtered image [right center], and the combination of images as described by Eq. 1 [right].

3.3   Object isolation and capturing user input

The image is composed of dark background pixels, bright and intermediate gray pixels belonging to the objects of interest within the cyst. The objects of interest are the organelles targeted by the fluorescent dye, in the case of the MitoTracker images it is the active mitochondria while in the DAPI images it is the nuclei of the individual bradyzoites residing within the cyst.

To analyze individual objects of interest the objects first needed to be automatically detected, or separated from the background, and segmented into individual entities from within the whole imaged cyst. A threshold value to separate background pixels from pixels representing the target objects was implemented, this method segmented the target objects sufficiently as well. In this approach only the pixels above the threshold value were retained with a value of 1 while the remaining pixels were set to 0, creating a binary image defined solely of having the target object pixels or background pixels. A key factor in choosing the appropriate threshold for this method of

segmentation was finding a balance between over segmentation and under segmentation. The former results in fragmented target objects while the later results in combined objects composed of multiple target objects. In increasing the threshold for binarization there may occur over segmentation while decreasing the threshold typically results in under segmentation.

To allow computational analysis with minimal user input, an initial i.e. starting threshold value based on intensity distribution was calculated to best separate the objects of interest from the rest of the image. The Otsu method determines a threshold which minimizes the interclass variance and maximizes the between class variance of foreground and background pixels [18]. Therefore, the Otsu method was used to find an initial threshold which was set as the default value. After the evaluation of a few images, it was determined that 40% of the Otsu threshold provided a more acceptable starting point to default to for capturing the objects of interest within the MitoTracker images. An example of the binarized Mito images using the unmodified and modified Otsu threshold is displayed in Figure 3. The decision to modify the Otsu threshold was based on wanting to be more conservative and limit over segmentation because the key attribute of the objects that was of interest was their morphologies, i.e. their shapes. The DAPI images required a higher percentage of the Otsu threshold which resulted in the exclusion of a greater number of lower value pixels. This exclusion prevented under segmentation in the DAPI images where the individual target objects were less defined. A slider was added to the GUI to allow the user to adjust these default thresholds and the resulting changes to the binary image could be observed immediately by the user.
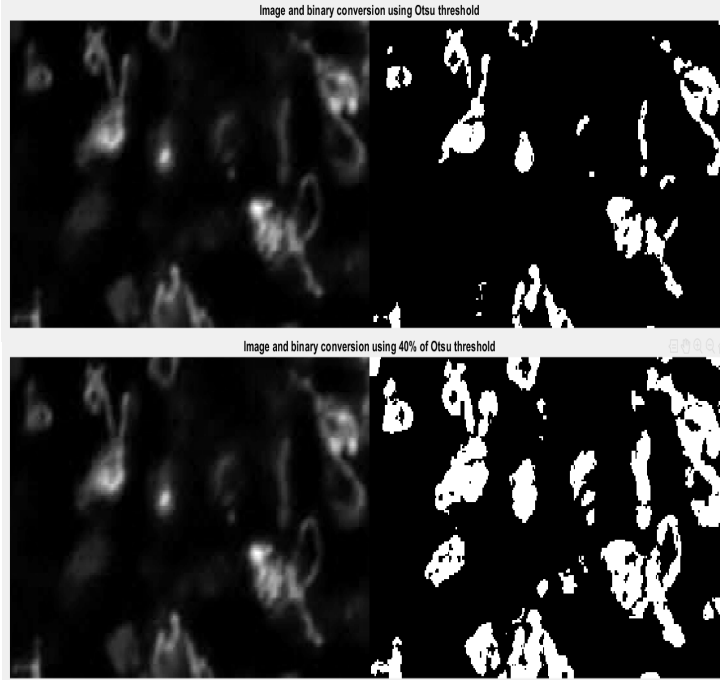
Figure 3. The left side of the top and bottom image shows the same zoomed in section of a Mito image, while the top right side shows its corresponding binary image using the Otsu threshold and the bottom right side shows the binary image using a modification of using 40% of the Otsu threshold.

Following thresholding, the DAPI images underwent watershed segmentation [19] for more accurate segmentation of the target objects, especially the nuclei within proximity of neighboring nuclei. Since the watershed method can be prone to over segmentation a height transform [20] was applied to suppress smaller peaks, of one standard deviation of the image pixel values or less, before applying watershed segmentation.

The boundary of each object of interest within the binary image was then obtained using the Moore-Neighbor tracing algorithm with Jacob's stopping criterion [17]. This step resulted in a label matrix enabling quick access to the total number of objects along with their location within the image. With the label matrix and by superimposing the binary image on the original image, geometric and pixel value information of each object

12

within the tissue cyst image were extracted. After this stage, a total of 22 features, which are defined in the next section within Table 2, were extracted for each object.

To exclude objects that would typically be ignored during manual evaluation the identified objects were filtered based on two criteria to exclude objects smaller than a certain size and dimmer in intensity than a certain threshold (but with an intensity greater than the binarization threshold). The features of minimum axis length and maximum intensity of each object were compared to user selected size and intensity thresholds. A 'low size' exclusion criterion defaulted to half a micron, i.e., seven pixels to exclude objects that had a minor axis, or width, less than this value in microns. A second exclusion criterion, 'low intensity', did not initially exclude any objects unless updated by the user as it defaulted to the binarization threshold value used. This criterion excluded objects that are dimmer and may have occurred due to over segmentation. The user could update these exclusion criteria as needed via the GUI in the appropriate fields to observe their effect.
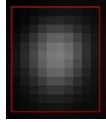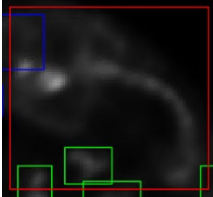
3.4 Mitochondrial morphology classes and features extracted
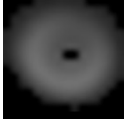
Upon initial analysis the same mitochondrial profiles observed in the tachyzoites served as the main morphologies to classify the objects identified in the Mito images. These morphologies included a lasso (ring), arcs (linear and curvilinear), tadpoles (sperm-like) and blobs (collapsed) forms [4, 5].

It was important to then have an additional 'Other' class for those objects that the user could not place in any one of the four specified classes since the morphological diversity of encysted bradyzoites mitochondrial profiles has not been explored before. After initial analysis it was observed that the Lasso class was rarely present and there

were a number of objects that appeared to have a similar shape but were markedly smaller. We referred to this new observed shape as "donut", therefore a Donut class was added, with the Lasso class left unused a total of five classes was maintained for further analysis. Table 1 below shows a representative image and description of each class mentioned here.

Table 1. Mitochondrial morphological class description

Class		Description
Lasso		Ring or loop shaped with a visible hole in the middle, typically larger than 1 micron (2-3 micron). Normal morphology of a functioning mitochondria. Will most likely surround a nucleus.
Tadpole		Tadpole or sperm like, having a bulbous bright end with a tail or linear string like protrusion that is often thinner and fainter. The tail is typically equal to or greater than the diameter of the head.
Arc		Dumbbell shaped, fairly uniform arms with a thin connection, often connected in the shape of an arc. This may also include linear shapes.
Blob		Punctate, circular in shape, ranging in brightness, typically smaller (between 0.5 microns and 1 micron in diameter). This is an abnormal mitochondrial shape, possible due to fragmentation from autophagy.
Other		Unidentifiable, typically a combination of more than one due to proximity and/or failure of thresholding to segment properly. A typical parasite is around 7 x 3 microns in size, so if it approaches this size it is going to be combination of more than one mitochondrion.

| Donut | | Comparable to the Lasso class but does not surround a nucleus, typically smaller in size compared to the Lasso class. |

Features describing shape and intensity attributes of the objects were extracted to best capture the geometric and pixel intensity descriptions outlined in Table 1. Many of the features were extracted using Matlab's 'regionprops' function [21], these are denoted with an asterisk in Table 2.

Table 2. Description of object features used

1	Area*	Total number of pixels composing the object.
2	Perimeter*	Total number of adjoining pixels around the border of the object.
3	Max Intensity*	Maximum pixel intensity within the object.
4	Min Intensity*	Minimum pixel intensity within the object.
5	Avg Intensity*	Mean pixel intensity of the object.
6	Intensity Variance	Variance of the pixel intensities composing the object.
7	Intensity Std Dev	Standard deviation of the pixel intensities composing the object, or the square root of the variance (feature 6).
8	Intensity MSE	The mean squared error between original object pixel intensities and the enhanced image object pixel intensities, both extracted with regionprops*, this intensity change is caused by the linear equalization and contrast enhancement performed.
9	Extent*	To calculate the extent, the area of the object is divided by the area of the bounding box, or the smallest box that will contain the object. It is the ratio of pixels within the object to the pixels in the bounding box.

10	Circularity*	Circularity describes the 'degree of roundness' or how close an object is to a circle. A result of 1 corresponds to a perfect circle and 0 corresponds to highly irregular/non-circular shapes; (4 X pi X Area) ./ Perimeter.^2.
11	Eccentricity*	This is calculated by using the second moments to find the ratio of foci distance to major axis length, assigning a circle a value of 0 and a line a value of 1, commonly defined as the deviation from circularity.
12	Major axis*	The axis length is determined by finding the corresponding major axis of an ellipse with the same normalized second central moments.
13	Minor axis*	The axis length is determined by finding the corresponding minor axis of an ellipse with the same normalized second central moments.
14	Aspect ratio	This is calculated by dividing the major axis by the minor axis.
15	Peak count	The number of peaks found in the object after suppressing peaks one standard deviation (of total object bounding box) high using a height transform (Matlab's 'imhmin') and then applying the watershed algorithm (Matlab's 'watershed'). This was useful for identifying object that may be a combination of multiple objects as having multiple peaks.
16	Hole count	The number of holes found in the object, Matlab's 'bwboundaries' was used to help calculate this feature.
17	Hole size	The total number of pixels that are part of a hole in the object, or the addition of the areas of all holes found in the object, Matlab's 'bwboundaries' was used to help calculate this feature.
18	DAPI proximity	The minimum distance found between the mitochondrial object extrema* to any surrounding DAPI object centroids*.
19	DAPI distance	The minimum distance found between the mitochondrial object centroid* to any surrounding DAPI object extrema*.
20	Mito proximity	The minimum distance found between the mitochondrial object extrema* to any surrounding mitochondrial centroids*.

| 21 | HOG peak count | The histogram of oriented gradients was found using Matlab's 'extractHOGFeatures' from the Computer Vision toolbox. This function evaluated the objects image as single block and found the magnitudes of the gradients that fell into 8 orientations (0-45,45-90, 90-135, 135-180, 180-225, 225-270, 270-315, 315-360). To get one value from these distributions the number of dominant orientations present was found. First the magnitudes were sorted in descending order, this helped to remove complications due to varying object orientations. Then the location in the array where the difference between the neighboring value was greater than the standard deviation was returned. If no value was returned, as in the case of a circular objects, then a value of zero was returned. |
| 22 | Extrema Intensity ratio | The objects' centroid* and orientation* were determined to then translate and rotate the objects' image to be centered around its centroid and align the major axis with the x-axis. The objects image is then split in half along its major axis* and the sum of the intensities of each half are determined to then calculate the ratio of the brighter half's pixel intensities to that of the other half's pixel intensities. |

3.5   Machine learning for morphology classification

To address the morphology classification problem a supervised logistic regression algorithm was chosen due to its ease of interpretation and implementation. Logistic regression determines the probability that a given example or object's discrete features are associated with a class or not by implementing the sigmoid function (eq.2) typically with a threshold of 0.5 to make Boolean predictions from the linear regression hypothesis (eq.3) [14]. The sigmoid function is derived from the log of the odds ratio, which is the ratio between the probability the event, i.e. class, occurs to the probability

17

that it does not occur. With linear regression the predictive output is on a linear and continuous scale that can have a large range, in applying the sigmoid function the predictive output gets scaled between 0 and 1, thus can be considered to be the probability of the class with a range of 0 to 1. Since we sought to not just predict the probability of one class, each class was evaluated individually to determine the probability for each class and then compared to select the class with the maximum likelihood.

$$g(z) = \frac{1}{1+e^{-z}} \quad \textit{Equation 2. Sigmoid function}$$
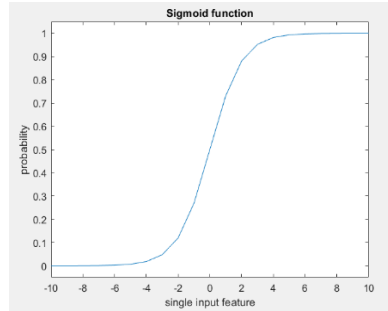
Figure 4. Sigmoid function plotted in MATLAB

$$h(x^M) = \theta_0 x_0^M + \theta_1 x_1^M + \cdots + \theta_N x_N^M \quad \textit{Equation 3. Linear regression hypothesis}$$
superscript M for the Mth object and subscript N for the Nth feature
implicit is that $x_0^M = 1$ and θ_N parameters set the contribution from each N feature

To easily implement the concepts of logistic regression in code the features (variable X), actual classifications (variable Y), and predicted classifications (or h(X) in the following equations) for each object were stored in a vectorized form which allowed for the linear regression and logistic regressions equations to be simplified to eq 4 and 5 below [14].

18

$$h(X) = X\theta^T \quad \text{\textit{Equation 4. Matrix implementation of linear regression hypothesis}}$$

$$h(X) = g(X\theta^T) \quad \text{\textit{Equation 5. Logistic regression hypothesis}}$$

X is a matrix composed of M rows, one for each object, and N+1 columns of features with an implicit feature with the value of 1 added to allow for the simplification of eq 3 to eq 4.

h(X) and Y are column vectors with M rows, each row representing the predicted and actual (known) class for each M object.

θ is a parameter row vector with the first column, θ_0, being constant with a value of 1, and the following columns representing the coefficiants defining the contribution of each N feature.

(For example if there were 200 objects, each with 22 features, dimensions of X would be 200 by 23. While h(X) and Y would be 200 by 1 row vectors, and θ would be 1 by 23 row vector)

When developing a machine learning algorithm an optimum solution for the $\boldsymbol{\theta}$ parameter can be determined more efficiently with features that have been scaled to values of the same order via eq.6 [14]. For example, the range for the feature 'Area' may range from 7 to 700 pixels while the 'Extent' feature may range from 0.13 to 0.96 therefore, the features were first normalized due to the varying ranges of values, this is also referred to as feature scaling.

$$X_N^M = \frac{X_N^M - mean(X_N)}{std(X_N)} \quad \text{\textit{Equation 6. Feature scaling}}$$

$$mean(X_N) = \frac{1}{M}\sum_{i=1}^{M} X_N^I \quad \text{\textit{Equation 6.1 Feature means}}$$

$$std(X_N) = \sqrt{\frac{1}{M-1}\sum_{i=1}^{M}\left|X_N^i - mean(X_N^i)\right|^2} \quad \textit{Equation 6.2 Feature standard}$$

deviations

The logistic regression algorithm is trained by finding a parameter vector, represented by $\boldsymbol{\theta}$, that minimizes the cost function, or error. The cost function is a measure of how the hypothesis, i.e. predicted class from eq. 5, compares to the ground truth, or in this example the gold standard, (Y) as described in eq.7 [14]. The function returns a lower cost when the hypothesis is closer to the ground truth, or gold standard.

$$J(\theta) = \frac{1}{M}\sum_{i=1}^{M}\left[-Y^{(i)} log\left(h(X^{(i)})\right) - (1 - Y^{(i)})log\left(1 - h(X^{(i)})\right)\right] + \frac{\lambda}{2M}\sum_{j=1}^{N}\theta_j^2$$

Equation 7. Regularized cost function; superscript (i) represents the ith row or object

This cost equation includes a regularization parameter, $\boldsymbol{\lambda}$, to prevent overfitting the model to the provided training objects. Regularization helps to generalize the model for future predictions by adding a cost for the parameter vector used and limiting the algorithms overall fit to the training set used.

To find the parameters which minimize the cost and in return find the most accurate hypothesis, the 'fminunc' function from MATLABs Optimization Toolbox was used. This function uses the Quasi-Newton algorithm with BFGS Quasi-Newton method and a cubic line search procedure [21]. The optimum parameters are found for each class, essentially creating 5 models. To do this the ground truth vector, Y, is modified when each model is trained to consist of 1's for the class of interest and 0's for all other classes. Next eq.5 is used to find the hypothesis with each model for the provided objects; this results in 5 hypotheses for each M object. Finally, a one-vs-all approach is

used as described in eq 8 where the class associated with the model resulting in the greatest hypothesis value is then set to be the predicted class for that object [14, 22].

$$prediction^M = max\,[\,h_{blob}(X^M)\,,h_{tadpole}(X^M),h_{donut}(X^M),h_{arc}(X^M),h_{other}(X^M)\,]$$

Equation 8. Multi-class logistic regression prediction; superscript M represents the Mth object

The performance of the algorithm is then determined by comparing the prediction vector, $prediction^M$, with the ground truth vector Y.

3.6   Process for capturing manual classification and implementation of machine learning

The design of the user interface of the developed program is depicted in Figure 5 and a flowchart of how the input and output of the program are handled is portrayed in Figure 6. The first step of the program requires a Mito image input, then the corresponding DAPI image is loaded, and image processing of both images is performed automatically. The user reviews the DAPI image first and its default settings to update them as needed to outline the nuclei within the image, this step is performed on the first tab of the program which is setup similar to what is shown in Figure 5. The user then moves to review the Mito image and its default settings, on the second tab of the program GUI, updating the default settings to identify each mitochondrial object as accurately as possible within the program. The user is warned if the number of mitochondrial objects found is greater than the number of nuclei found in the DAPI image. This is done as the parasite is a eukaryote with a single nucleus and mitochondrion per organism indicating the number mitochondria is generally expected to be similar to the number of nuclei. The presence of "inactive" mitochondria, which are

21

not visualized by the MitoTracker dye, can result in the number of mitochondrial objects recorded being lower than that for DAPI-stained nuclei. In addition, mitochondrial fragments would typically be excluded using the size criteria, although extensive mitochondrial fragmentation [4] could result in a higher relative mitochondrial count. Once the user is satisfied with the settings, a file is saved containing all the necessary variables to replicate the users' results. This first step in the process was performed by a designated lead user with an extensive background in reviewing *T. gondii* microscopy imaging.

The second step was performed by the lead user along with two other trained users varying in experience and background in the investigation of mitochondrial function and *T. gondii*. The lead user had 25+ years of experience while the other two had less than 4 and 2 years. To train and help calibrate the group of users Table 1 was reviewed by all and edited to best define the mitochondrial morphological classifications.

Each trained user then performed the classification step by inputting the settings file created by the lead user in step one. This step automatically processes the images to display the individual objects ensuring that the same objects were seen by all users. The objects were reviewed, and an appropriate morphology classification was selected independently by each user. Within the GUI, each object started with a green bounding box and once the object had been reviewed the bounding box for it turned from green to blue, while the current object being reviewed was red to allow for easy tracking of the current object being reviewed along with the user's overall progress. When complete a file was saved with the user's classifications.
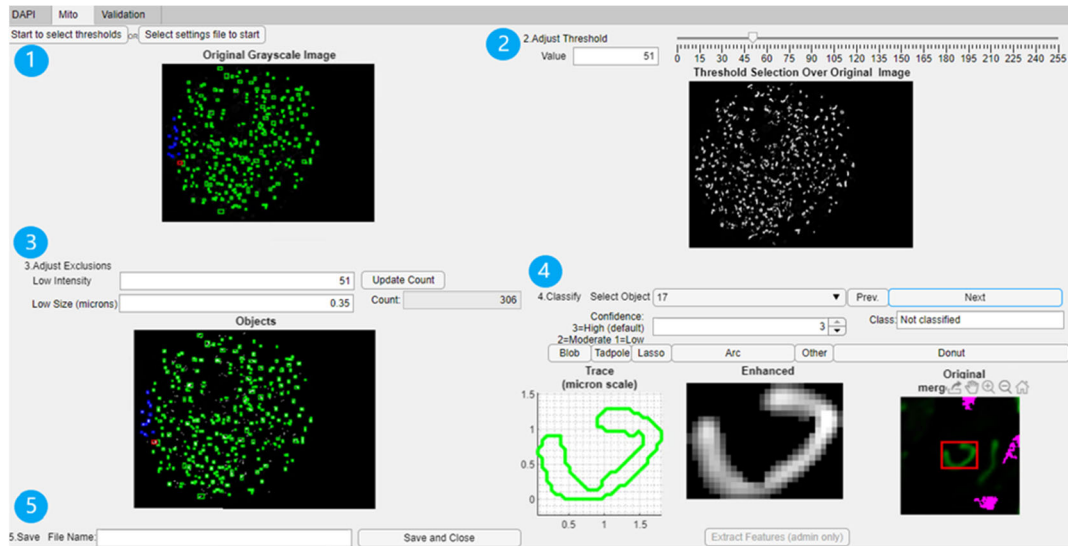
Figure 5. Mito tab, the second tab of the GUI with numbers representing the sequential steps to be taken: 1) selecting the button to begin settings selection or the button to load a settings file if a settings file is used step 2 and 3 are skipped, 2) adjusting the threshold setting, 3) adjusting the exclusion criteria, 4) classifying the object morphology, and 5) saving the output files.
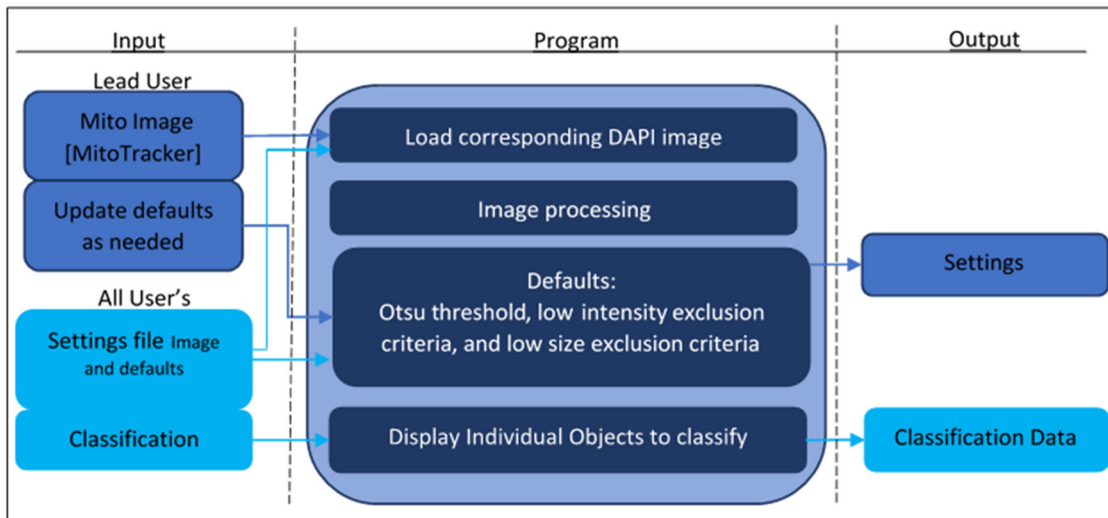
Figure 6. General program design for input and output. The classification data file is a M x 1 vector, M number of objects. The settings output file stores the image file name, DAPI threshold, DAPI low intensity criteria, DAPI low size criteria, DAPI object count, DAPI object count after watershed, Mito threshold, Mito low intensity criteria, Mito low size criteria, Mito object count, and an indicator to extract features or not.

3.7   Congruence in manual classification and validation of automatic classification

A total of 12 images were classified individually by all three users. Of these 12, 5 images they had previously reviewed once before, 2 images were repeated from the first set of 5 but with different settings, the remaining 5 were new images. The 5 images previously reviewed and the 5 new images were randomly selected from a set of 20 imaged tissue cysts. A review of the congruence among manual classifiers, i.e. the degree to which different users placed the same objects within the same class was performed to determine whether certain morphologies were readily classifiable. This review was previously performed on the first 5 images after classifications were made with the 5 classes of Blob, Tadpole, Lasso, Arc, and Other [23], referred to as round 1 in Figure 7. These findings were then compared to the congruency determined from this study of the same 5 images, referred to as round 2, to evaluate if there was an improvement in congruency amongst these 5 images.
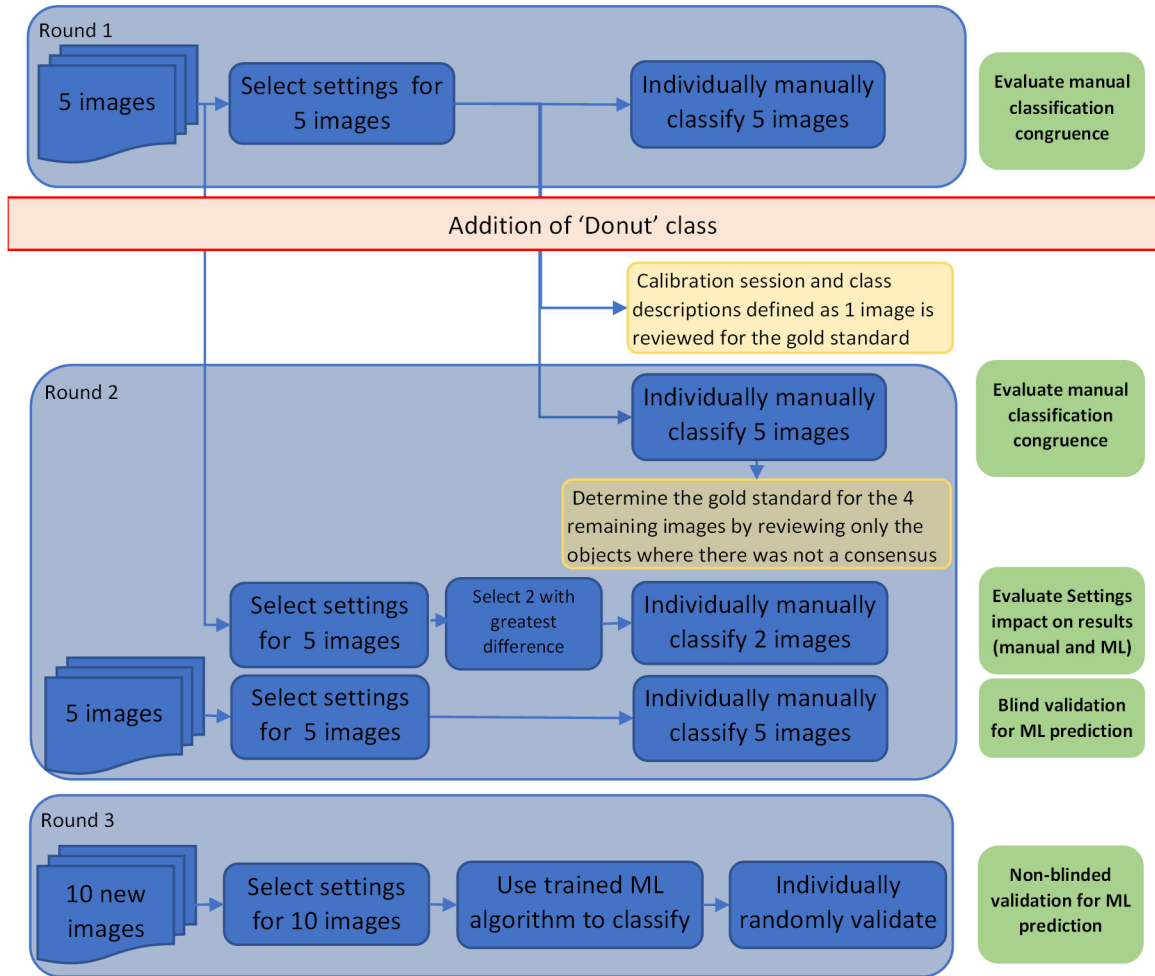
Figure 7. Flowchart for analysis

The settings for the first 5 images were made previously in one sitting by the lead user, then the lead user was tasked with selecting the settings for the first 5 images again along with 5 new images at a later sitting. The change in settings for the first 5 images was evaluated, of which 2 were selected that had the largest differences. These 2 images were classified twice, once with the first settings and again with the new settings to evaluate the impact the change in settings had on the class distributions within.

The gold standard, which is used as the ground truth in this application, was used for training the machine learning algorithm and was determined in two sittings, each

with all users present to assist in determining the most accurate class to be used for each object. In the first sitting all the objects from one of the initial 5 images, from round 1, was reviewed as a calibration session before the users individually classified all 12 images. In the second sitting the remaining 4 of the initial 5 images were reviewed to determine a gold standard for just the objects they did not already have a consensus on from their individual classifications.

The trained algorithm then predicted the classification of objects in images that were not used in training the algorithm. The performance of the algorithm is reported for both the training set and test set, of which there is one test for each user. The performance is reported for each user's classification to evaluate if it performs well for all users. The performance is based on several calculations as follows: The first is the overall accuracy, the percent of all objects that were correctly classified. Then the sensitivity, specificity, precision, and F score are calculated per class as defined in eq 9,10,11, and 12.

$$Sensitivity = \frac{True\ Positives}{True\ Positives + False\ Negatives} \qquad Equation\ 9.\ Sensitivity$$

$$Specificity = \frac{True\ Negatives}{True\ Negatives + False\ Positives} \qquad Equation\ 10.\ Specificity$$

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \qquad Equation\ 11.\ Precision$$

$$F\ score = \frac{2 * Sensitivity * Precision}{Sensitivity + Precision} \qquad Equation\ 12.\ F\ score$$

Lastly a non-blind validation was performed as outlined in Figure 7, denoted 'Round 3'. This allowed the users to observe the predictions made by the machine learning algorithm and provide further feedback. In this step the settings for 10 new

images, remaining from the initial set of 20, were made by the lead user. The objects identified in these images were input into the trained machine learning algorithm to predict their class. A validation input file was created containg the settings used and the predicted class. Each user was then randomly selected several objects to verify the predictions machine learning algorithm by using the 'Validation' tab, displayed in Figure 8.
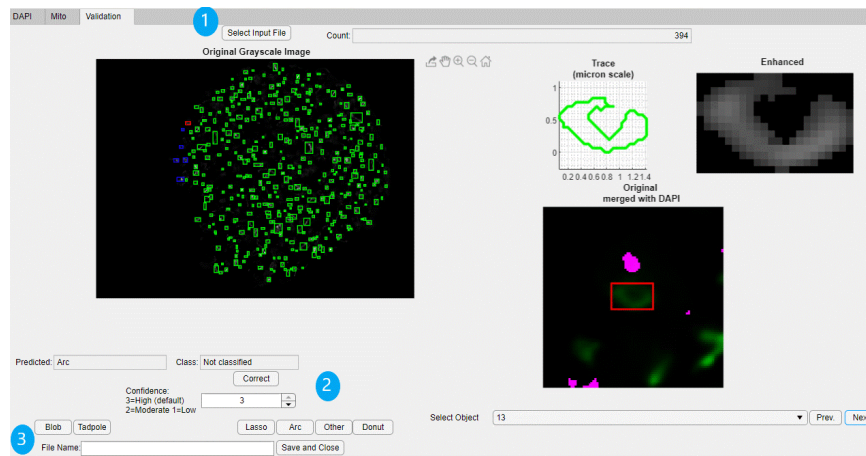
Figure 8 Validation tab, the third tab of the GUI with numbers representing the sequential steps to be taken: 1) selecting the button to load a validation input file if a settings file, 2) observing the images on the right and the predicted class displayed then selecting 'Correct' if the predicted class is accurate and if not selecting the appropriate class, 3) saving the output files.

4.1   Settings and manual classification

The first image analyzed by all users together to obtain manually made classifications that could be used as gold standard classifications, prior to individual classifications of the remaining images resulted in the following class breakdown for the 334 objects that were detected.
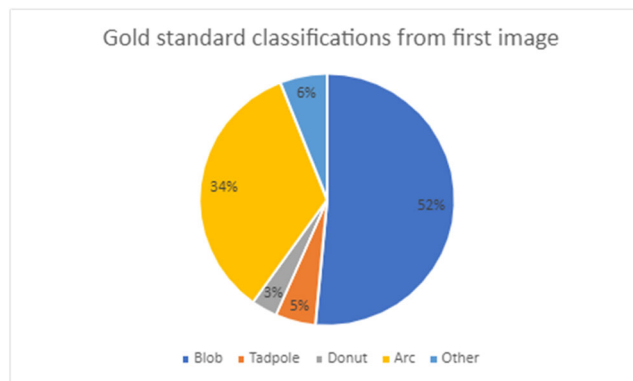
Figure 9. Class distribution for first 334 objects used for gold standard from one image

The congruency, i.e. agreement, between users from round 1 classifications, when using the Lasso class for this image, was at 52.7% for all three users, 95.5% for at least two users, and 4.5% where none of the users agreed. As described in the methods section, the addition of the Donut class was made after round 1 classification's due to having no objects that the users unanimously agreed were Lassos and the observation of multiple Donut shaped objects. With the addition of the Donut class a total of 11,3% objects were classified as donuts when determining the gold standard. Therefore, in evaluating each users' round 1 classifications to the gold standard it is expected to be slightly lower due to not having the donut class when these classifications were performed, this data resulted in the following percent match: 73.1% for user 1, 82.6% for user 2, and 72.8% for user 3.

The congruency between users' classifications for this first image, after the gold standard was determined, in round 2 was 71.3% for all three users, 99.1% for at least 2 users, and 0.9% where none of the users agreed. For 6 of the objects (which was 1.8% of all objects), the users all agreed on a class in their individual classifications that differed from the gold standard i.e. what they had classified as the class for these objects was different than what the consensus class ended up being agreed for the objects. There were 9 objects (2.7% of all objects), where the users did not have a consensus classification and none of the users matched the gold standard class. Evaluating each user to the gold standard with the round 2 individual classifications resulted in the following percent match: 78.4% for user 1, 86.8% for user 2, and 88% for user 3.

The congruency between users when using the Lasso class for the first 5 images in round 1 produced the results in Figure 10(A) with 57% of the objects being classified similarly by all users and 96% classified the same by 2 of the 3 users. In round 2 of classifications for these images with the addition of the Donut class, 71.8% of the objects were classified similarly by all users and 98.8% were classified the same by 2 of the 3 users. Figure 10(B) shows a higher congruency in Blob recognition occurred in round 2 as compared to the first time the objects of these images were classified in round 1, while both graphs in Figure 10 display that Blobs were the most frequent class to occur as well as the most likely to be uniformly recognized by all operators. There was an increase in instances were all three users agreed for the Arc and Other classes except Tadpoles which remained constant.

Figure 10. Round 1 (A) and round 2 (B) percentage of all objects classified separated by class and the percent agreement between users for the first 5 images

For all 5 images from which a gold standard classification was determined, there were a total of 1,138 objects and the class breakdown within these remained comparable to that of the first image, with the most objects belonging to the Blob class followed by Arcs, Tadpoles, Others, then Donuts as displayed in Figure 11. Comparing these consensus (i.e. gold standard) distributions to that of each user's round 2 individual classification the following percent matches were determined: 83.4 for user 1, 86.7 for user 2, and 87.1 for user 3. Objects for each class where there was 100% agreement between all users and which matched the gold standard are shown in Figure 12.
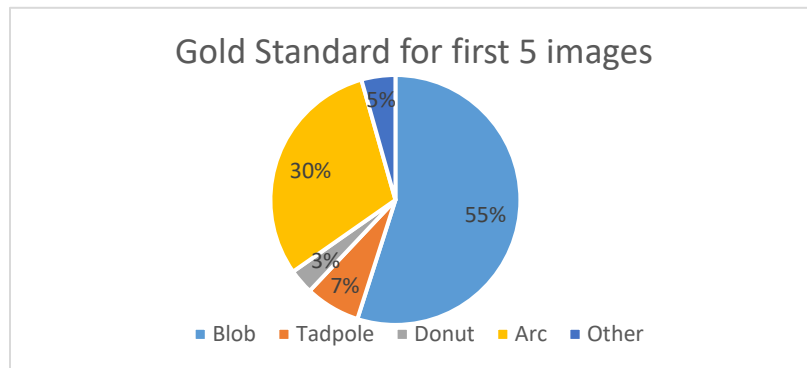
Figure 11. Class distribution for the 5 gold standard images
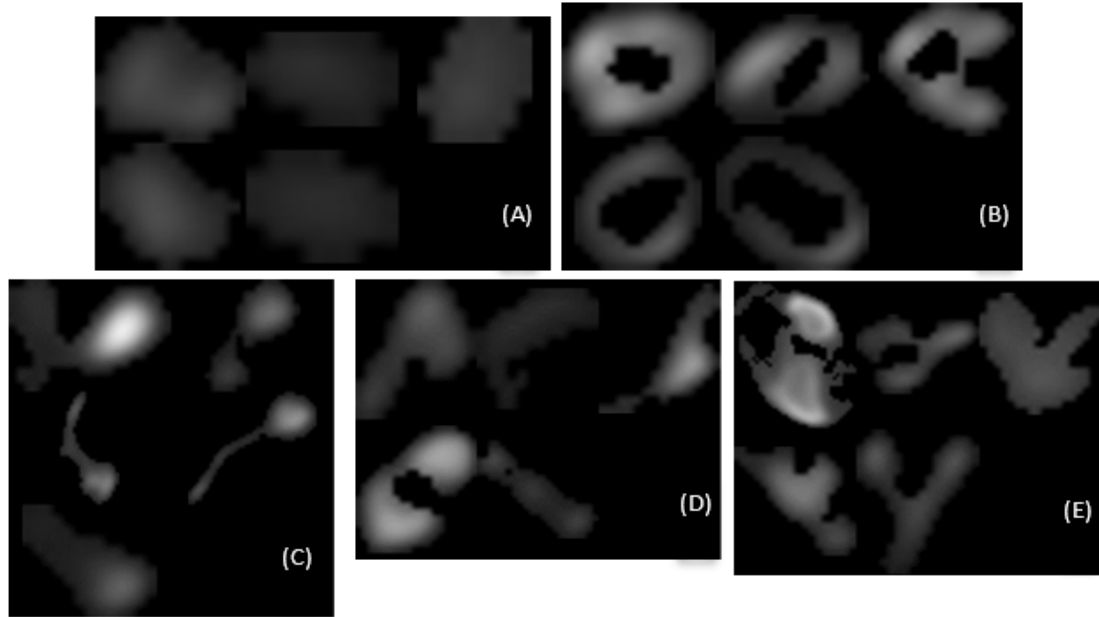
Figure 12. Objects classified congruently with all users and the gold standard for Blobs (A), Donuts (B), Tadpoles (C), Arcs (D), and Others (E)

From the previous classifications in round 1 the users did not unanimously determine any object to be a Lasso in these images, user 1 classified a total of 3 Lassos (0.3%), user 2 and 3 classified 2 and 0 objects as Lassos, respectively. Since the Donut class made up only 3% (37) of all objects in the gold standard classifications, evaluating the degree of agreement between each users' round 1 classifications and the gold standard is expected to be slightly lower due to not having the donut class when these classifications were performed. This data resulted in the following percent match: 74.9 for user 1, 82.2 for user 2, and 75.6 for user 3.

To visualize the normalized features distribution for each class as defined by the gold standard from the first 5 images Figure 13 displays a boxplot for each class and each of the 22 features where the top of the box marks the 75th percentile and the bottom the 25th percentile with a red line indicating the median, the red '+' markers indicate the outliers which are greater than one and a half times the interquartile range. These results

31

show the trends for each feature for each class, for example the circularity and extent are highest for the Blob class followed by the Donut class while the eccentricity and aspect ratio are highest for the Tadpole and Arc class. A telling feature of the Donut class is the presence of holes which correlates with the higher hole size while a telling feature of the Tadpole class is to have a bright head with dim tail which correlates to the higher extrema intensity ratio shown.
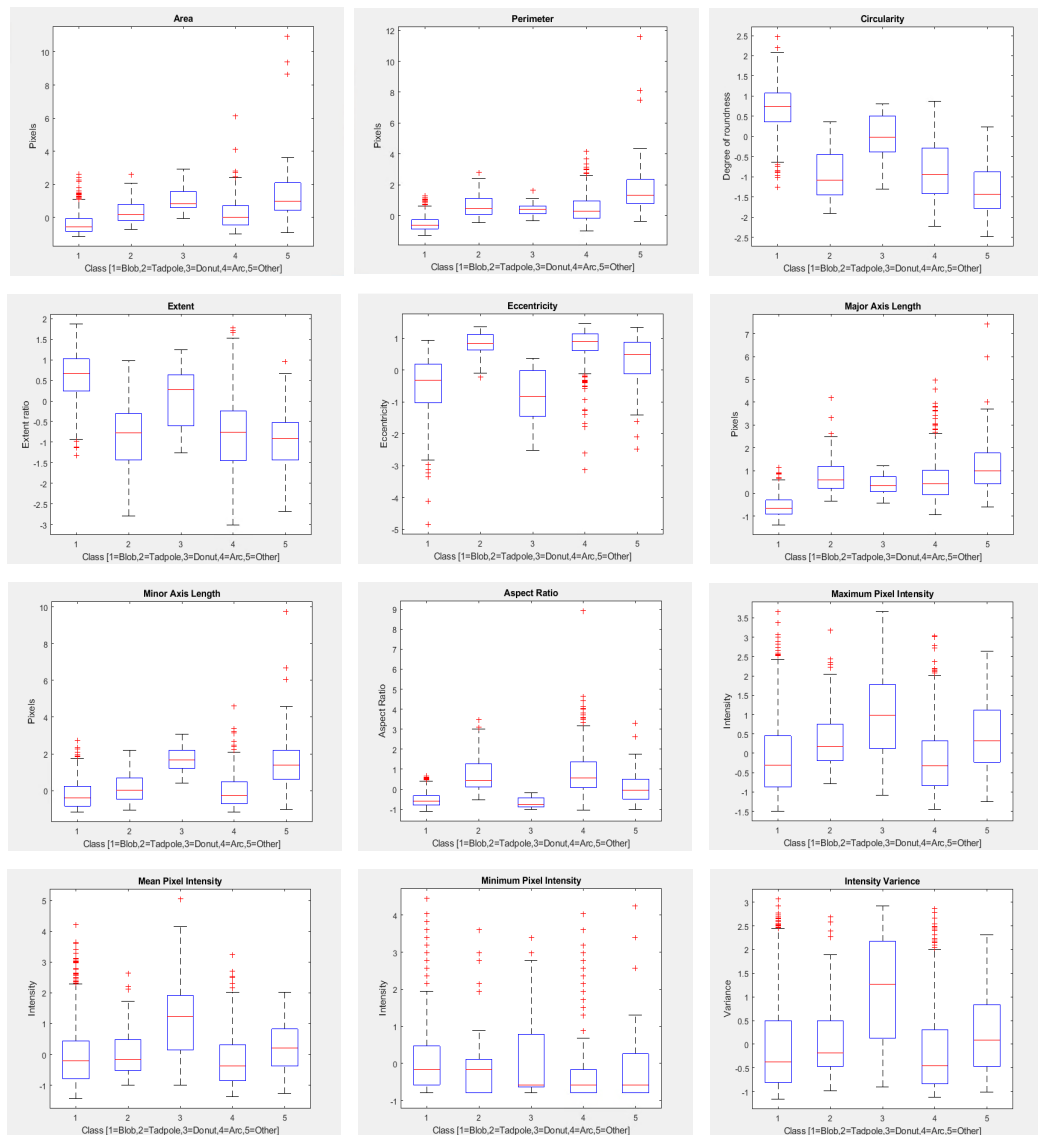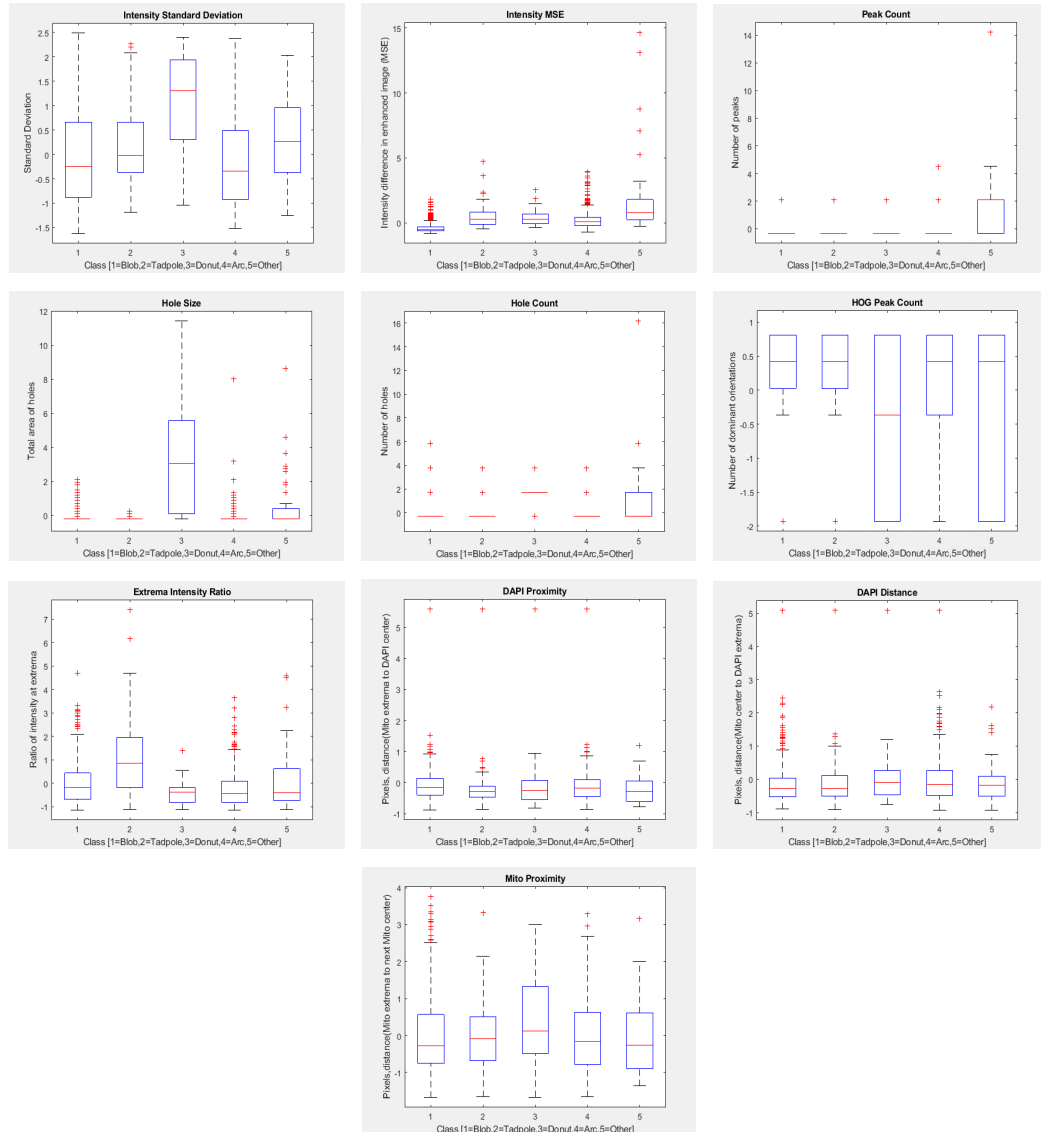
Figure 13. Depiction of the 22 normalized features and their distribution for each class from the 1138 gold standard objects

There was a total of 927 objects that were classified from the new set of 5 images in round 2, that had not previously been analyzed. Of the 927 objects, 697 were placed in the same class by all 3 users (75.2% of the objects) i.e. they were classified similarly by all users. A total of 918 objects (99%) were classified as belonging to the same class by 2 out of the 3 three users. These results, displayed in Figure 14 consistently show Blobs as the most frequent class to occur and a high congruency amongst users recognizing this

class. The agreement for and frequency of occurrence for the Arc, Donut, and Other classes were comparable to the results from the first five images with a slight increase in user congruency for Tadpoles.

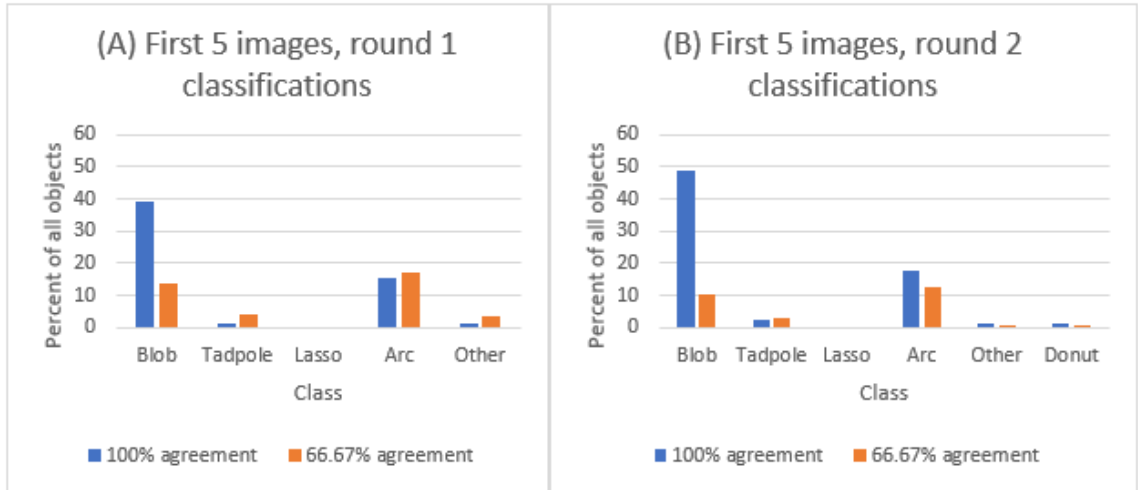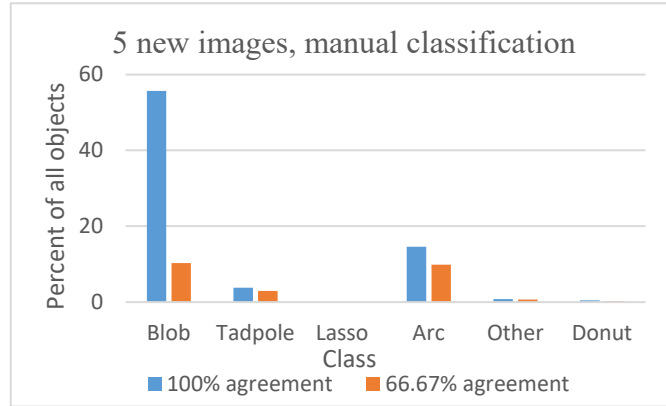*Figure 14. Percentage of all objects classified separated by class and the percent agreement between users for the 5 new images*

The settings determined in a second sitting were compared to the settings originally chosen in Table 3 and Table 4 for the DAPI settings and Mito settings, respectively. The object counts for the DAPI images increased for all images with the new settings, the percent increase ranged from 18.3% to 83.3% (averaging to a 42.4% increase). This increase in count resulted from a lower threshold being selected. The exclusion criteria were not adjusted from the defaults for any of the images.

Table 3. DAPI settings comparison

Threshold	(original) Threshold	Otsu%	(original) Otsu%	Watershed Count	(original) Watershed Count
69	97	54.337	76.378	558	375
50	62	39.3701	48.8189	614	335
81	112	63.7795	88.189	442	326
82	95	64.5669	74.8031	420	355
81	104	63.7795	81.8898	274	218

For three of the Mito images the user selected threshold values were increased and for two of the images the threshold values were decreased, the percent decreases were greater than that of the percent increases. The image with the greatest percent difference (decrease) in threshold value also had a decrease in the low size exclusion criteria of 0.15 from 0.5 microns resulting in the largest percent increase in the number of objects counted, a 91.2% increase in objects from the original (third in Table 4). The second largest percent difference in object counts was an increase by 25.2% which was caused by the 22.5% decrease in the threshold value. The low intensity exclusion criterion was not adjusted from the default for any of the images. These two images, highlighted in Table 4, were chosen to repeat with these new settings due to the greatest difference in resulting objects detected, and were included in round 2 for manual classification.

Table 4. Mito settings comparison

Threshold	(Original) Threshold	Otsu%	(Original) Otsu%	Low Size	(Original) Low Size	Count	(Original) Count
55	51	43.3071	40.1575	0.35	0.35	208	218
51	45	40.1575	35.4331	0.35	0.35	306	334
42	56	33.0709	44.0945	0.35	0.5	216	113
51	45	40.1575	35.4331	0.35	0.35	324	358
54	70	42.5197	55.1181	0.35	0.35	144	115

A summary of the impact of the settings change on class distribution and congruency is shown in Figure 15. The congruency was comparable with a slight decrease for the images with the new settings. Comparing the 2 images that were repeated with new settings, which had a total of 360 objects, 73.9% of the objects were

classified similarly by all users, and 98.1% of the objects were classified the same by 2 of the 3 users. When the original settings were used, a total of 228 objects were detected, 75% of the objects were classified similarly by all users, and 98.7% of the objects were classified the same by 2 of the 3 users. The distribution of classes identified by each user for the images remained relatively consistent with a slight decrease in Blobs and increase in Arcs in the images with new settings.

Figure 15. Percentage of all objects classified separated by classification and user for the 3 repeat images with original settings (A) and new settings (C). Percentage of all objects classified separated by classification and the percent agreement between users for the 2 repeat images with original settings(B) and new settings(D)

4.2   Machine learning performance

The first gold standard image, with 334 objects, was initially used to train the machine learning algorithm. Once the θ parameter arrays were determined (one for each

of the five classes), which is called Model 1, the training accuracy was found to be 86.8%, this was the fraction of the total number of objects that were classified correctly regardless of class that the trained algorithm achieved when predicting the class for the same 334 training objects. The breakdown per class, shown in Table 5 showed a high degree of performance for the Blob, Donut, and Arc classes.

Table 5. Training performance with first gold standard image

(Model 1)	Blob	Tadpole	Donut	Arc	Other
Sensitivity	0.97093	0.41176	0.90909	0.85088	0.45
Specificity	0.91358	0.98738	1	0.90455	0.98408
Precision	0.92265	0.63636	1	0.82203	0.64286
F score	0.946176	0.5	0.952381	0.836207	0.529412

Tables 5,7-9,11 and 13-15 break down the performance per class showing the sensitivity, specificity, precision, and F score as calculated by equations 9-10 in the methods section. The sensitivity, or recall, is the percentage of actual positives predicted accurately or the true positive rate while specificity is the percentage of actual negatives predicted accurately or the true negative rate. Precision also called positive predictive value is the number of true positives divided by the number of predicted positives. Lastly the F score uses the calculated sensitivity and precision values to provide a single metric to compare the performance for each class [14].

Testing this trained algorithm on the remaining 11 images that were classified individually produced an overall accuracy or 73.8%, 82.8%, and 82.6% for user 1,2, and 3 respectively. The overall accuracy was computed as the percent of object where the predicted class matched that of the user's classification. A summary of the objects that were identified by each user in the test set is shown in Table 6 and graphically depicted in bar chart form in Figure 16. The performance breakdown is displayed in Tables 7-9 by

user. The performance for Model 1 on the Blob class remained high and there was overall high specificity for all classes.
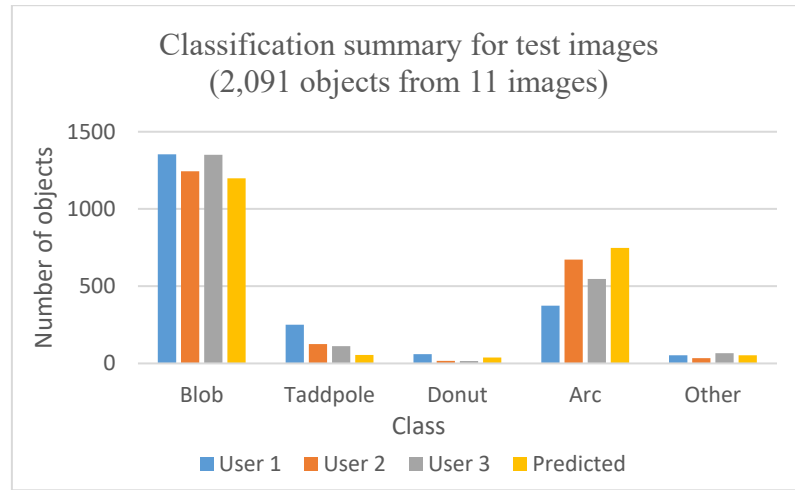
Figure 16. Summary of the number of objects each user classified as one of the five classes for the objects in the 11 test images

Table 6. Number of objects each user classified as one of the five classes for the 2,091 objects in the 11 test images

	Blob	Tadpole	Donut	Arc	Other
User 1	1354	250	60	374	53
User 2	1244	125	16	672	34
User 3	1351	112	15	547	66
Predicted	1198	55	38	748	52

Table 7. Model 1 testing performance with user 1 classifications

	Blob	Tadpole	Donut	Arc	Other
Sensitivity	0.84343	0.124	0.33333	0.90642	0.20755
Specificity	0.92402	0.98696	0.99114	0.76179	0.97988
Precision	0.95326	0.56364	0.52632	0.45321	0.21154
F score	0.894984	0.203279	0.408163	0.604278	0.209524

Table 8. Model 1 testing performance with user 2 classifications

	Blob	Tadpole	Donut	Arc	Other
Sensitivity	0.89309	0.272	1	0.83333	0.29412
Specificity	0.89728	0.98932	0.9894	0.86751	0.97958
Precision	0.92738	0.61818	0.42105	0.74866	0.19231
F score	0.90991	0.377778	0.592593	0.788732	0.232558

Table 9. Model Testing performance with user 3 classifications

	Blob	Tadpole	Donut	Arc	Other
Sensitivity	0.86454	0.24107	1	0.91773	0.24242
Specificity	0.95946	0.98585	0.98892	0.84067	0.98222
Precision	0.97496	0.49091	0.39474	0.67112	0.30769
F score	0.916438	0.323353	0.566038	0.77529	0.271186

When training the algorithm with all 1,138 gold standard object classifications from the first 5 images, the training accuracy was 84.8% with the following performance breakdown per class in Table 11. The training performance remained similar to that when training with one image. The parameters determined for the five classes is displayed in Table 10.

Table 10. Theta parameters

(Model 2)	Blob	Tadpole	Donut	Arc	Other
Constant (θ_0)	-1.11254	-6.49823473	-7.84	-1.9699	-5.11838
Area	0.34055	0.01724091	0.262	0.974739	1.771135
Perimeter	0.17384	-2.030202224	-1.67	-2.83466	1.055137
Circularity	3.00977	-1.520282926	-0.14	-2.42467	-2.7186
Extent	-0.1671	0.319468028	0.562	0.015405	0.382758
Eccentricity	1.68559	4.932588178	0.921	1.315383	-0.61986
Major Axis Length	-1.24554	-0.553633524	0.528	-0.11369	-1.98938
Minor Axis Length	-1.13004	1.835126685	3.016	1.607829	-1.30582
Aspect Ratio	-6.31523	-0.327253727	-0.42	1.123867	-0.26114
Maximum Pixel Intensity	1.04769	3.537202161	-0.87	-2.35301	-0.19481
Mean Pixel Intensity	-1.92191	-2.715971158	1.333	0.609366	-0.41831
Minimum Pixel Intensity	-0.00376	0.218155991	0.463	-0.09563	-0.13372
Intensity Variance	0.11547	-2.956779969	0.271	1.057729	-0.05514
Intensity Standard Deviation	0.71496	2.500957553	-0.41	0.257916	0.906855
Intensity MSE	1.28341	0.119710932	-2.5	-0.31131	0.039224
Peak Count	-0.52124	-0.387752903	-1.38	0.117529	0.343296
Hole Size	-1.26692	-1.838875404	2.036	-0.46652	0.046713
Hole Count	0.48507	0.013877879	-0.17	-0.07419	0.533047
HOG Peak Count	0.11276	0.281818941	-0.34	-0.07014	-0.15992
Extrema Intensity Ratio	-0.41361	0.756114891	-0.1	-0.49678	0.209065
DAPI Proximity	0.25784	-0.220159517	0.039	0.038435	-1.14201
DAPI Distance	-0.08573	0.139730564	-0.82	-0.03937	0.471918
Mito Proximity	-0.06414	0.210152542	0.03	-0.01026	0.168185

Table 11. Training performance with all gold standard objects

(Model 2)	Blob	Tadpole	Donut	Arc	Other
Sensitivity	0.9664	0.35802	0.67568	0.83768	0.36
Specificity	0.88694	0.9877	0.99637	0.88903	0.99081
Precision	0.91239	0.69048	0.86207	0.76658	0.64286
F score	0.938617	0.471545	0.757576	0.800554	0.461538

Testing this second trained algorithm on the second set of 5 new images that were classified individually produced an accuracy of 76.9%, 83.1%, and 85% for user 1,2, and 3 respectively. A summary of the number of the total objects of each class that were identified by each user in this test set is shown in Table 12 and graphically depicted in Figure 17. The performance breakdown is displayed in Tables 13-15 by user. The overall accuracy for each user increased and the performance for the Blob class remained high along with the overall high specificity for all classes.

Table 12. Number of objects each user classified as one of the five classes for the 927 objects in the 5 test images

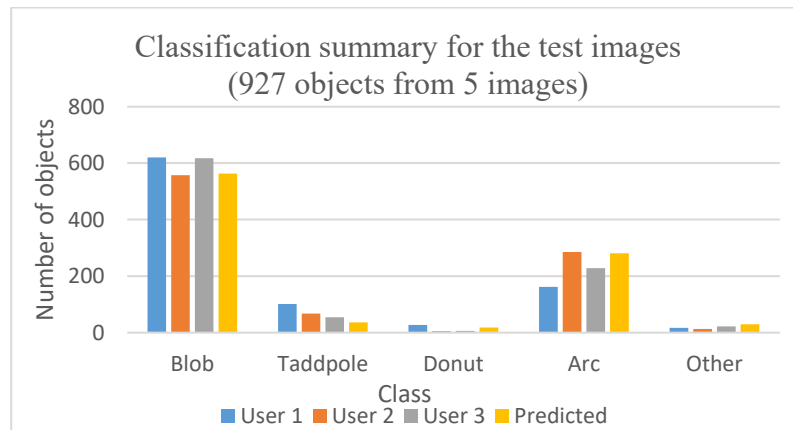| Number of Objects	Blob	Tadpole	Donut	Arc	Other
User 1	620	101	27	162	17
User 2	557	67	5	285	13
User 3	617	54	6	228	22
Predicted	563	36	18	281	29

Figure 17. Summary of the number of objects each user classified as one of the five classes for the objects in the 5 test images

Table 13. Model 2 testing performance with user 1 classifications

	Blob	Tadpole	Donut	Arc	Other
Sensitivity	0.85968	0.26733	0.37037	0.85185	0.29412
Specificity	0.90228	0.9891	0.99111	0.81307	0.97363
Precision	0.94671	0.75	0.55556	0.4911	0.17241
F score	0.901099	0.394161	0.444444	0.623025	0.217391

Table 14. Model 2 testing performance with user 2 classifications

	Blob	Tadpole	Donut	Arc	Other
Sensitivity	0.91741	0.43284	0.8	0.77544	0.38462
Specificity	0.85946	0.99186	0.98482	0.90654	0.97374
Precision	0.90764	0.80556	0.22222	0.78648	0.17241
F score	0.9125	0.563107	0.347826	0.780919	0.238095

Table 15. Model 2 testing performance with user 3 classifications

	Blob	Tadpole	Donut	Arc	Other
Sensitivity	0.88817	0.51852	0.83333	0.87719	0.31818
Specificity	0.95161	0.99084	0.98588	0.88412	0.97569
Precision	0.97336	0.77778	0.27778	0.71174	0.24138
F score	0.928814	0.622222	0.416667	0.785855	0.27451

In the final validation step, i.e. round 3, a total of 438 objects were collectively validated by the three users (each user validated a subset of these 438 objects). Of these 438 objects, the same 24 were randomly selected by two users for validation. These results are summarized in Table 16 below, which shows 87% of the validated objects were accurately predicted. A breakdown per class of the validated objects' predicted class to manual class is presented in Table 17. Looking at each user individually: user 1 validated 173 objects indicating 89% were correct, user 2 validated 122 objects indicating 88% were correct, and user 3 validated 167 objects indicating 86% were correct.

Table 16 Validation Summary

	#validated	#correct	#incorrect	%correct	%incorrect
All (excluding duplicates)	438	381	57	86.9863	13.0137
All users (including duplicates)	462	404	58	87.44589	12.55411

Table 17 Validation classification breakdown

%error (total of 13%)	Predicted class	Manual classification (excluding duplicates)				
		Blob	Tadpole	Donut	Arc	Other
5	Blob	[209]	5	11	3	3
1.6	Tadpole	1	[19]	0	6	0
0	Donut	0	0	[14]	0	0
4.3	Arc	5	9	0	[125]	5
2.1	Other	0	2	3	4	[14]

CHAPTER 5. DISCUSSION

T. gondii bradyzoites within tissue cysts have been found to be metabolically active with heterogeneous replication [3]. To better understand the metabolic activity and replication potential of these encysted bradyzoites investigation of the mitochondrial function can provide valuable information [8]. The classification of mitochondrial morphologies that are present within the bradyzoites can inform about the metabolic state of the bradyzoite as a functional state of any given cell is likely to be associated with the particular morphology of the mitochondrion [4, 5]. Image processing and classification algorithms can facilitate the analyses of larger numbers of images than what would be possible by the current process of manually classifying the mitochondria from the thousands of possible bradyzoites within a cyst as it is tedious and time consuming. This increased throughput will allow the gathering of information about the biology of this parasite at a scale that is not possible using current approaches.

The developed program helped to facilitate the manual expert classification of mitochondrial shapes and these classifications were used in the evaluation of a machine learning approach for the automation of the morphology classifications. A common machine learning method used in object detection for multi-class classification is multinomial logistic regression approach with a one vs all methodology [14], which was used in this case to predict the detected objects morphological class. With the extracted features and user labeled data collected via the developed program this approach proved to be successful for the mitochondria morphology classification problem presented here, with being able to achieve up to 85% overall accuracy, i.e. the fraction of correctly classified object regardless of class.

Consistent with previous findings [23] the Blob morphology (condensed or punctate mitochondrial profile) was the simplest to classify with a readily apparent appearance which likely resulted in the high level of congruency between users and the higher accuracy that was seen for the trained machine learning algorithm. Most of the objects were identified as belonging to the Blob morphology, with lower incidences of Tadpole and Arc morphologies which is not surprising as the Tadpoles and Arcs are associated with more active parasites and bradyzoites metabolic activity is still expected to be low [4]. When training the algorithm with gold standard objects of one image compared with that of five images there was a slight decrease in training accuracy but the testing accuracies slightly increased; with the larger training set the model was better generalized to fit the variance in the new test data but these differences in accuracy were minimal, suggesting one image may be sufficient for training but that more may be helpful. The trained machine learning performance resulted in a high degree of confidence in the prediction of Blobs and Arcs which make up a majority of morphological classes present in most bradyzoites. The results also showed that the variation in settings chosen for each image resulted in relatively large differences in the number of mitochondrial objects found within, but the class distribution remained fairly constant.

In summary, results of our study provide a robust program that can successfully automate the classification of the majority of mitochondria morphologies within encysted bradyzoites which can minimize the burden of evaluating these images and present opportunities for hybrid workflows involving partial automation in classifying the majority of objects present. The classes which attained lower accuracy and were of

low incidence could be presented to the user for manual classification, while the objects that were predicted to belong to a class with higher accuracy and where there was overall high congruency such as seen with the Blobs would not need to be reviewed, thereby allowing the majority of the objects within the cyst to be automatically classified which would dramatically reduce the time required for manual classification.

REFERENCES

1 Tenter, A.M., Heckeroth, A.R., and Weiss, L.M.: 'Toxoplasma gondii: from animals to humans', Int J Parasitol, 2000, 30, (12-13), pp. 1217-1258

2 Weiss, L.M., and Kim, K.: 'The development and biology of bradyzoites of Toxoplasma gondii', Front Biosci, 2000, 5, pp. D391-405

3 Watts, E., Zhao, Y., Dhara, A., Eller, B., Patwardhan, A., and Sinai, A.P.: 'Novel Approaches Reveal that Toxoplasma gondii Bradyzoites within Tissue Cysts Are Dynamic and Replicating Entities In Vivo', mBio, 2015, 6, (5), pp. e01155-01115

4 Ghosh, D., Walton, J.L., Roepe, P.D., and Sinai, A.P.: 'Autophagy is a cell death mechanism in Toxoplasma gondii', Cell Microbiol, 2012, 14, (4), pp. 589-607

5 Ovciarikova, J., Lemgruber, L., Stilger, K.L., Sullivan, W.J., and Sheiner, L.: 'Mitochondrial behaviour throughout the lytic cycle of Toxoplasma gondii', Scientific reports, 2017, 7, (1), pp. 42746

6 Lavine, M.D., and Arrizabalaga, G.: 'Analysis of monensin sensitivity in Toxoplasma gondii reveals autophagy as a mechanism for drug induced death', PLoS One, 2012, 7, (7), pp. e42107

7 Perry, S.W., Norman, J.P., Barbieri, J., Brown, E.B., and Gelbard, H.A.: 'Mitochondrial membrane potential probes and the proton gradient: a practical usage guide', Biotechniques, 2011, 50, (2), pp. 98-115

8 Sinai, A.P., Watts, E.A., Dhara, A., Murphy, R.D., Gentry, M.S., and Patwardhan, A.: 'Reexamining Chronic Toxoplasma gondii Infection: Surprising Activity for a "Dormant" Parasite', Current Clinical Microbiology Reports, 2016, 3, (4), pp. 175-185

9 Doggett, J.S., Nilsen, A., Forquer, I., Wegmann, K.W., Jones-Brando, L., Yolken, R.H., Bordon, C., Charman, S.A., Katneni, K., Schultz, T., Burrows, J.N., Hinrichs, D.J., Meunier, B., Carruthers, V.B., and Riscoe, M.K.: 'Endochin-like quinolones are highly efficacious against acute and latent experimental toxoplasmosis', Proc Natl Acad Sci U S A, 2012, 109, (39), pp. 15936-15941

10 Araujo, F.G., Lin, T., and Remington, J.S.: 'The activity of atovaquone (566C80) in murine toxoplasmosis is markedly augmented when used in combination with pyrimethamine or sulfadiazine', J Infect Dis, 1993, 167, (2), pp. 494-497

11 Seeber, F., Ferguson, D.J., and Gross, U.: 'Toxoplasma gondii: a paraformaldehyde-insensitive diaphorase activity acts as a specific histochemical marker for the single mitochondrion', Exp Parasitol, 1998, 89, (1), pp. 137-139

12 Fisch, D., Yakimovich, A., Clough, B., Wright, J., Bunyan, M., Howell, M., Mercer, J., and Frickel, E.: 'Defining host-pathogen interactions employing an artificial intelligence workflow', Elife, 2019, 8

13 Bauman, N., Ilic, A., Lijeskic, O., Uzelac, A., Klun, I., Srbljanovic, J., Cirkovic, V., Bobic, B., Stajner, T., and Djurkovic-Djakovic, O.: 'Computational image analysis reveals the structural complexity of Toxoplasma gondii tissue cysts', PLoS One, 2020, 15, (8), pp. e0234169

14 Rebala, G., Ravi, A., and Churiwala, S.: 'An Introduction to Machine Learning' (Cham: Springer International Publishing AG, 2019. 2019)

15 Watson, G.F., and Davis, P.H.: 'Systematic review and meta-analysis of variation in Toxoplasma gondii cyst burden in the murine model', Exp Parasitol, 2019, 196, pp. 55-62

16 Watts, E.A., Dhara, A., and Sinai, A.P.: 'Purification Toxoplasma gondii Tissue Cysts Using Percoll Gradients', Curr Protoc Microbiol, 2017, 45, pp. 20C 22 21-20C 22 19

17 Gonzalez, R.C., Woods, R.E., and Eddins, S.L.: 'Digital Image processing using MATLAB' (Pearson Prentice Hall, 2004. 2004)

18 Otsu, N.: 'A Threshold Selection Method from Gray-Level Histograms', IEEE transactions on systems, man, and cybernetics, 1979, 9, (1), pp. 62-66

19 Meyer, F.: 'Topographic distance and watershed lines', Signal Processing, 1994, 38, (1), pp. 113-125

20 'Morphological Image Analysis: Principles and Applications', Sensor review, 2000, 20, (3)

21 MathWorks, I.: 'MATLAB', 2018, (version 9.4.0 (R2018a))

22 Barber, D.: 'Bayesian reasoning and machine learning' (Cambridge University Press, 2012. 2012)

23 Place, B.C., Troublefield, C., Murphy, R.D., Sinai, A.P., and Patwardhan, A.: 'Computer Aided Image Processing to Facilitate Determination of Congruence in Manual Classification of Mitochondrial Morphologies in Toxoplasma gondii Tissue Cysts*'. Proc. Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2021. (Accepted)

VITA

Brooke Place

Education:

- University of Kentucky (Lexington, KY), 2018-2021
  - Master of Science in Biomedical Engineering
  - Engineering in Healthcare Graduate Certificate
- University of Dayton (Dayton, OH), 2010-2015
  - Bachelor of Science in Electrical Engineering

Work Experience:

- Research Assistant - University of Kentucky (Lexington, KY), March 2019 – Current
- Technical Service Engineer - Epic Systems (Madison, WI), September 2015 - May 2018
- Co-op (3 rotations) at Defense Engineering Corporation (Dayton, OH), Spring 2013

Conference Presentations:

- Brooke Place, Cortni Troublefield, Robert D. Murphy, Anthony Sinai, and Abhijit Patwardhan: "Computer Aided Image Processing to Facilitate Determination of Congruence in Manual Classification of Mitochondrial Morphologies in Toxoplasma gondii Tissue Cysts*". Annual International Conference of the IEEE Engineering in Medicine and Biology Society, October 31-November 4, 2021 (accepted)
- Brooke Place, Anthony P. Sinai, Abhijit Patwardhan. "Automated Image Processing Algorithms to Aid Understanding of Toxoplasma gondii Bradyzoite Biology". BMES Annual Meeting abstract presented October 14-17, 2020.
- Mohammad Javad Mollakazemi, Dibyajyoti Biswal, Brooke Place, and Abhijit Patwardhan. "Effects of respiratory pathway on processing of emotions, autonomic and brain response". BMES Annual Meeting abstract presented October 14-17, 2020.
- Dibyajyoti Biswal, Mohammad Javad Mollakazemi, Brooke Place and Abhijit Patwardhan. "Heart Rate and Breathing Rate calculated from Cheeks and Lips Using Green and Derived Colors from Video". IEEE iSES paper presented December 14-16, 2020.

Publications:

- Thomas Raffay, Mandy Brasher, Brooke Place, Abhijit Patwardhan, Peter Giannone, Henrietta Bada, Philip Westgate, and Elie Abu Jawdeh. "Response to First Dose of Inhaled Albuterol in Mechanically Ventilated Preterm Infants". Journal of Perinatology. 2021.

Awards:

- Outstanding Master's Student 2021, University of Kentucky Department of Biomedical Engineering
- 3rd place as a team in Von Allmen Center for Entrepreneurship's COVID-19 Hack-A-Thon, September 2020