

An Affordable Semi-Automated Optical Cell and Colony Counting Device*

Emily Wu, Jared Scott, and Trevor Murphy

Abstract—This work details the development of an affordable system to identify cell colonies in petri dishes with high fidelity and limited user interaction to facilitate the research of proton beam cancer treatment. Methods: The proposed system uses consumer camera equipment combined with semi-automated staging and open source software for image acquisition and processing. To assess system performance, system-generated colony and cell counts of 3T3 mouse fibroblast samples similar to those used in proton beam therapy research are conducted and compared against human results. The generated results are lower in variability than manual counting, although generated counts are lower than human estimates for colonies and higher than estimates for individual cells. Further analysis and tuning can be conducted to assess and improve the accuracy of the counter.

Clinical relevance— By developing a framework to make colony counting more robust to human error, more efficient, and more accessible, this work aims to facilitate proton beam therapy research, and thus the development of safer and more effective cancer treatment.

I. INTRODUCTION

Proton beam therapy precisely targets tumors to treat cancer. A proton beam carries lethal radiation to a desired cell to ultimately kill it in a more precise, localized way than other radiation therapy techniques [1]. Despite its proposed benefits in cancer treatment, proton beam therapy requires more research to determine its usage and effectiveness.

Colony-forming cell (CFC) assays assess the effectiveness of proton beam therapy. The general technique is described in [2] and reviewed in brief here. A set of cells are seeded in a dish and irradiated with a proton beam, which disrupts the cell growth potential. The cells are then grown up to six or seven generations, since successful disruption of cell growth potential takes multiple generations to observe. After growth there will be multiple clumps (groups of collocated cells generated from the same progenitor cell) of different cell counts. Clumps with more than 50 cells (a midpoint between successful replication over five and six generations) are considered colonies generated by a cell that survived the therapy. By counting the number of colonies and comparing to a control dish that is not irradiated, the effectiveness of the proton beam therapy is assessed. The CFC assay often relies on manually counting of colonies for its main metric.

Manual colony counting is unreliable. A proton beam therapy experiment can generate around 6000 colonies that need to be counted. Additionally, the colony and cell sizes are small and require switching between a magnifying glass and microscope to view. Manual counting takes a large amount

of time, focus, and repetitive motion, resulting in a count that is not consistently repeatable and has large operator to operator variability.

There are a variety of assistive tools and devices to count colonies available on the market. They range in cost from \$30 to \$300,000 and vary in complexity, degree of user interaction, and expected sample format [3][4][5][6][7]. Many of these devices focus on identifying bacterial colonies, and thus do not offer sufficient magnification or resolution to count cells. On the other hand, the devices capable of automatically counting both cells and colonies are expensive, limiting their accessibility.

The work proposes a system using consumer camera equipment combined with semi-automated staging and free software for image acquisition and processing. With this system, the aim is to accelerate proton beam therapy research, and thus the development of more effective and safe cancer treatment.

II. MATERIALS AND METHODS

This work describes the design of an affordable system to count 60 mm diameter petri dishes stained with Giemsa containing around 100 clumps of 20 um wide cells. To achieve colony counting, the system must be able to identify cell clumps within a petri dish, and determine which are colonies. To do so, the system must be able to count cells, since clumps vary in width and density. To this end, an optical imaging device with several modules is developed. The modules are as follows: visualization of clumps and cells, staging, image acquisition, and image analysis. The overall approach is a multi-step process illustrated in Fig. 1, and in more detail in Fig. 2. To visualize the clumps and cells, the system scans and images the entire petri dish, then stitches the images together. The fully stitched image is analyzed with a two-step categorization process. In the first step, the system applies sizing thresholds to categorize the clumps as being too small to be colonies, large enough to surely be colonies, or in the threshold size range. In the second step, the system conducts cell counts for the threshold clumps to determine if they are colonies.

A. Selection of optical and imaging components to affordably and efficiently visualize cells and clumps

The goal of the visualization module is to efficiently and affordably generate images conducive for image analysis to extract colony and cell counts. To visualize cells, the optics must have sufficient contrast, spatial resolution, magnification, and pixel sampling.

*This work was supported by Sumitomo Heavy Industries

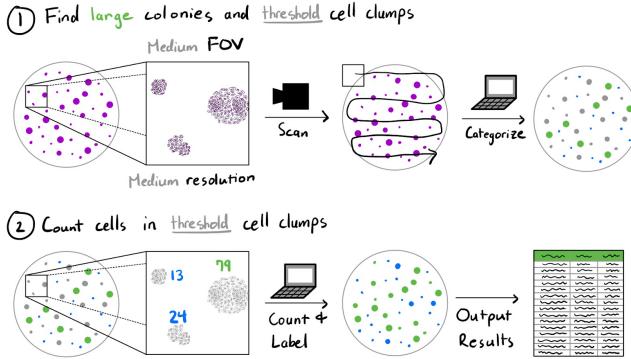


Fig. 1. The system has implemented a two step process for imaging and identifying cell colonies in a petri dish. In the first step, the dish is imaged via scanning and cell clumps are categorized by size. For the clumps that are around the threshold size range and cannot definitively be categorized by area, cell counting is conducted to determine if they are colonies or not.

Four-phase workflow

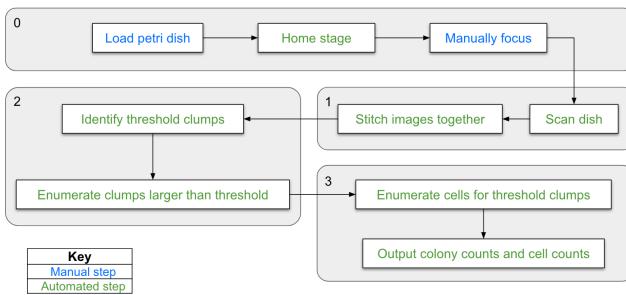


Fig. 2. Implementing the two-step counting process actually requires four phases. In the current proof of concept device, most of the steps are automated, but the user is still required to execute transitions between phases as well as load the dish and focus the camera in the preparation phase.

Giems stain absorbs red light with peak absorbance at around 645 – 655nm and provides intracellular color contrast, since it strongly stains the nucleus. To facilitate the initial intended strategy of counting cells by identifying nuclei, a four LED array designed with peak emission around 655nm (key components in BOM Table link) used to transilluminate the sample. To ensure uniform illumination, the light is diffused with a sheet of ground glass.

To ensure sufficient detail for cell counting, the spatial resolution is selected so that the nuclei spanned 6–10 pixels. Mammalian nuclei range from 5–10 μm , resulting in a target resolution of approximately 1 $\mu\text{m}/\text{pixel}$ [8]. In photography, the resolution is typically limited by the magnification and the pixel size and can be estimated by the following equation where P represents the pixel width in μm and M represents the magnification.

$$R = \frac{P}{M} \quad (1)$$

The target resolution guided the camera selection. Selecting the camera also requires balancing time cost and monetary cost. To efficiently image the whole petri dish, the camera resolution or field of view (FOV) should be

maximized so that each image captures more of the dish. However, there is a tradeoff between larger FOVs and lower imaging time, and higher monetary cost. To find a camera that balanced this tradeoff, the number of pictures required to image a 60 mm diameter petri dish versus cost is assessed. It is found that the Canon EOS R5 camera combined with the Canon MP-E 65 mm macro lens offers an affordable and simple image acquisition system with high camera resolution (45 megapixels) and magnification (1-5X). Since the EOS R5 has a pixel size of 4.4 μm , a magnification of 4X achieves the target resolution, resulting in nuclei spanning 5 – 9 pixels, and cells spanning about 18 pixels.

B. Selection of imaging settings

To select imaging settings (f-stop, ISO, exposure time, illumination strength), images are taken of cells transilluminated with red or white light. Petri dishes of 3T3 mouse fibroblast cells are prepared for testing of the system. Cells are plated and grown for 3-7 days, then fixed and stained with Giems. To select adequate image settings, images are acquired with varied settings and qualitatively assessed for the best contrast and image sharpness across a single dish.

C. Stage design

XY-staging moves the petri dish sample relative to the camera so that the whole dish can be imaged. The camera used in this work has an FOV that suggested linear staging is the best option. The project ended up using two lead screw positioner linear stages [9] fixed to each other.

Because the optical setup does not have autofocus functionality, a z-stage is needed to manually focus the system before image acquisition. Manual staging is acceptable since a benchtop test of imaging different regions of the dish confirmed that once focused, the whole dish stayed in focus. Manual focusing would be necessary just once for each new dish. A manual linear Z stage is selected with sufficient focusing step resolution (based on the depth of field of the camera) and sufficient total range to conveniently adjust the camera during testing and assembly.

The structure of the staging platform is created from commercial off the shelf components to provide a stiff platform that would not significantly eat into focusing step resolution. The entire device is shown in Fig. 3.

D. Image Acquisition and Stitching

A Python program is developed to control the XY-stage and camera during the image acquisition process. The program alternates between sending commands to the stepper motor controllers and sending image capture commands to the EOS R5 using a Python wrapper for the Canon EDSDK [13] in order to move the petri dish through a grid pattern and sequentially collect images of a rectangular region of interest. The captured images are then computationally stitched into an aggregate image of the region of interest using the Microscopy Image Stitching Tool provided by the National Institute of Standards and Technology [14][15].

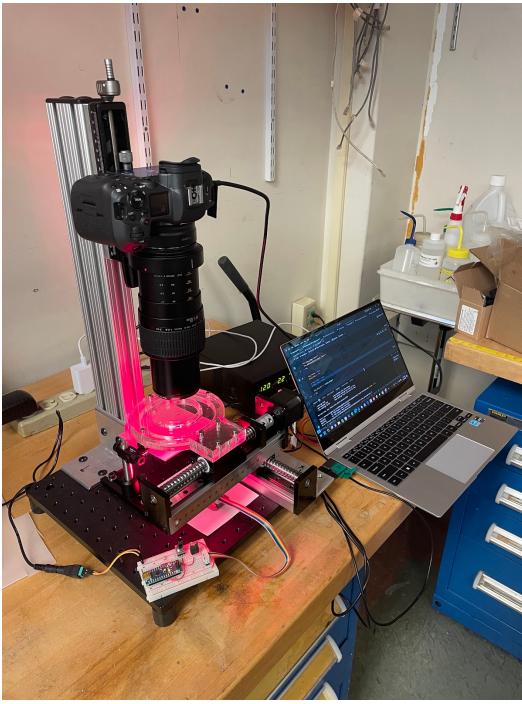


Fig. 3. The proof of concept device is benchtop and requires an external power supply for stage control, as well as a computer to execute the stage control and image analysis.

E. Image Processing

To extract meaningful information about colonies and cells from the acquired images, a Python program interfaces with the Fiji open-source software developed by the National Institutes of Health [16]. To identify colonies in an imported image, the script first runs a rolling ball background subtraction function with a radius of 10 pixels on the image to correct any aberration or inhomogeneous illumination, and to remove debris and unwanted dish artifacts before counting. An automatic color threshold is then applied to isolate illuminated Giemsa-stained cells within the image. The image is then converted to a binary mask, and a Gaussian blur ($\sigma = 30$) is applied to consolidate cells into clumps before reconverting the image to binary. After the binary image of clumps is derived from the original, the script runs a particle analysis function to identify and enumerate large particles (colonies) within a particular size threshold (currently tuned at 5×10^5 to 1×10^7 square pixels for fibroblast cells imaged in trials). This threshold is set by the user to exclude clumps that are too small to be considered colonies, and force the partitioning of detected image regions that are too large to be single colonies (e.g. conjoined clumps). The program saves the areas and indexes of identified colonies in CSV format. In order to identify individual cells within a clump, the Python script takes an imported image of cells and runs a background subtraction function, then employs the use of Fiji's Find Maxima function to identify the pixels in the image with locally maximal intensity values that exceed a prominence threshold (currently tuned at 17 for red-illuminated fibroblast cell images); these pixels correspond to the nuclei of cells



Fig. 4. On the left, a region of a petri dish containing fibroblast colonies illuminated using red light; on the right, an image of the results of automated colony counting using this image. The program identifies the approximate locations and sizes of three major colonies, highlighted in green; additional parameter tuning may be desired to improve area identification for sparse clumps such as those in the lower left quadrant of the image.

in the image. The indices and coordinates of the identified maxima are saved in CSV format.

F. Initial Data Collection by Humans

For preliminary evaluation of the image analysis, human data is collected for comparison. An expert, as well as 4 different non-experts are given 10 images to count clumps and 9 images to count cells. All the images are manually taken with the device. Two sets of images of the same dish regions are captured: one for manual counting, illuminated using white light, and one for automatic counting, using red light.

For the clump images, the humans are instructed to identify all cell clumps they see in the image and label them as a colony, not colony, or unclear based on first pass evaluation with no cell counting. The unclear label implies that upon initial viewing, the human believes that an actual count of the cells should be performed to determine if it is a colony or not. The total number of colonies and unclear clumps among all the images are added for comparison with the number generated by the device's initial thresholding image processing step.

For the cell images, the humans are instructed to count all the cells in a cropped image of a small clump. The total number of cells among the images are added for comparison with the number generated by the device's cell counting step.

The non-expert values are combined into an average value and standard deviation to get a sense of human variation and accuracy compared to an expert.

III. RESULTS

Images of Giemsa-stained mouse fibroblast cell clumps captured by the device, in addition to the results of automated counts are displayed in Figures 4 and 5.

Upon completion of the device, its performance is compared to the functional requirements generated by the sponsoring company at the beginning of the design in Table I. Specifically, the system ended up having a footprint of $305 \times 343 \times 540 \text{ mm}^3$, an image acquisition time of 8 minutes and 29s, an estimated aggregate image processing time on the order of 4 minutes, and a total hardware cost around \$5.6K for the system.

Table II and Table III detail the initial human data collection compared to the robot performance.

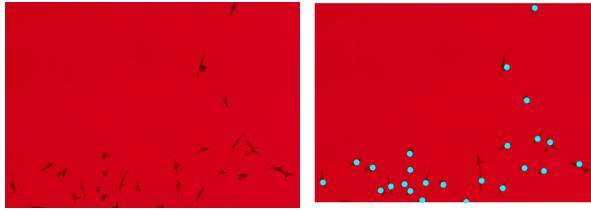


Fig. 5. On the left, a region of a petri dish containing fibroblast cells illuminated using red light; on the right, an image of the results of automated cell counting using this image. The cyan dots are the computed locations of cell nuclei in the image.

TABLE I
OVERALL SATISFACTION OF FUNCTIONAL REQUIREMENTS

Goal	Target	Realized
affordable	< \$50k per device	$\approx \$5.6k$ per device
table top	$300 \times 300 \times 300 \text{ mm}^3$	$305 \times 343 \times 540 \text{ mm}^3$
petri-dish based	yes	yes
fast	< 3 min per dish	> 8 min 29s per dish Successful clump identification Successful cell nuclei counting Threshold optimization required
distinguishes colony from non-colony clumps	± 1 colony	

IV. DISCUSSION

A. Human Variability

As expected, human variability is large, as seen in the large range of values. It should be noted that the human data set is small and not of a sufficient size for most statistical analyses. More data will be collected in a systematic manner to create a database for possible machine learning image analysis and better comparison to the device output. Despite the small amount of data, the variation in human choice is expected to be larger compared to the variation in the device output.

For comparisons with automatic devices, assessing human variability to establish a ground truth is a general area of future necessary work. Previous works make this assessment based on an arbitrarily chosen comparison of automatic and manual counting values. Specifically the automatic processes are acceptable if within $0.5 \log_{10}$ (the manual count) from the manual count [10][11]. This arbitrary datum gets larger for more objects to be counted presumably to account for higher variability of manual counting of larger numbers. If applied to the collected data here, the acceptable deviation of the system from the expert's value would be ± 0.7 clumps and ± 1.2 cells. Given the complex nature of counting the cells and the high human variability, these seem unreasonable. The question would then be what is reasonable? Further questions to be explored include: how sensitive is the cell assay process in general, to cell density variation, plating efficiency? What accuracies are needed from the cell assay testing? What are practitioners okay with in terms of accuracy and why?

To explore what is reasonable with regard to comparison,

TABLE II
TOTAL COUNTS OF CLUMPS THAT ARE DEFINITELY COLONIES OR REQUIRING CELL COUNTING AMONG 10 TEST IMAGES

Expert	Non-expert	System
56	21-77	22

TABLE III
TOTAL COUNTS OF CELLS IN 9 TEST IMAGES OF CLUMPS

Expert	Non-expert	System
274	214-312	350

future work is proposed to use this device to test the sensitivities of the CFC assay which requires assessment of the device's own sensitivities and repeatability.

B. System Performance Evaluation

Table I suggests that the performance only achieved two functional requirements of the sponsor pending future work to assess the accuracy and repeatability. However, since most of the functional requirements are estimated guidelines from the sponsor, the results are acceptable. The size and speed satisfy the need behind the functional requirements by still fitting on a table and freeing up the human operator to do other activities while it runs. The device meets the needs articulated through the functional requirements. However, future work is still needed to properly assess accuracy in a way meaningful for the sponsor and the general bioengineering field.

As the repeatability of the device's counting process is expected to be much better than the human variability, the process could be used to gather data relevant to determining the factors that affect CFC assay. While the system currently undercounts clumps and overcounts cells as shown in Tables II and III, the system removes a great degree of human subjectivity from the CFC counting process by making the assay a function of individual parameters such as the σ value of the Gaussian blur function and threshold values. In the future, these parameters can be varied independently by human users or machine learning algorithms to empirically discover settings for optimal enumeration accuracy when compared against expert manual counts. Future work would also be needed to assess the repeatability and sensitivity of the process to plate, cell, and hardware operation variation. With better understanding of the system and assay sensitivity, this system could be used to justify limits for simpler and faster automatic counting tests like area measurement [12].

C. Other Future Work

Updates to the staging can be developed to improve automation of the process and automation with prior and post processes with the dish. Examples of improvements could include automated z-stage motion for focusing and automated dish loading and unloading. Additionally, the optics could be adapted for fluorescent microscopy for possibly better cell counting accuracy.

V. CONCLUSIONS

The developed system affordably automates the enumeration of cell colonies in CFC assays to more precisely assess the effectiveness of proton beam therapy on treating cancer. The system removes the human variability and time required during manual counting and offers an accessible platform to better generate assessment of accuracies and sensitivities in the CFC assay process. Additionally, since the overall approach is not restricted to the analysis of cancer cells, it could be applied to other research that requires colony counting or 5 micron resolution imaging.

ACKNOWLEDGEMENTS

The team thanks Dr. Nevan Hanumara, Dr. Jay Connor, and Anthony Pennes for their mentorship, LED array creation, and design assistance. The team thanks Shinji Nomura for his colony counting expertise. The team also thanks Dr. Maxine Jones for preparing cell samples, and Steve Wasserman for optics guidance. Lastly, the team thanks Maxwell Hensley and Canon for aid in camera selection.

REFERENCES

- [1] "Proton Beam Therapy Program - Overview - Mayo Clinic." [Online]. Available: <https://www.mayoclinic.org/departments-centers/proton-beam-therapy-program/sections/overview/ovc-20185491>. [Accessed: 04-Mar-2023].
- [2] N. A. P. Franken, H. M. Rodermond, J. Stap, J. Haveman, and C. van Bree, "Clonogenic assay of cells in vitro," *Nat. Protoc.*, vol. 1, no. 5, pp. 2315–2319, 2006.
- [3] "Fisherbrand™ Counter-Pen/Colony-Counter." [Online]. Available: <https://www.fishersci.com/shop/products/fisherbrand-counter-pen/1464977>. [Accessed: 23-Feb-2023].
- [4] "SphereFlash® – Automatic Colony Counter." [Online]. Available: <https://iul-instruments.com/product/sphereflash-automatic-colony-counter/>. [Accessed: 23-Feb-2023].
- [5] "Protos 3 - Complete automated colony counting and chromogenic identification system." [Online]. Available: <https://www.synbiosis.com/product/protos-3-colony-counter/>. [Accessed: 23-Feb-2023].
- [6] "GelCount™ Mammalian-cell colony, spheroid and organoid counter." [Online]. Available: <https://www.oxford-optronix.com/gelcount-cell-colony-counter>. [Accessed: 23-Feb-2023].
- [7] "CloneSelect Imager (CSI and CSI FL)." [Online]. Available: <https://www.moleculardevices.com/products/clone-screening/mammalian-screening/cloneselect-imager>. [Accessed: 23-Feb-2023].
- [8] J. Lammerding. (2011, Apr). Mechanics of the Nucleus. *Compr Physiol.* [Online]. 1(2), 783–807. Available: <https://onlinelibrary.wiley.com/doi/10.1002/cphy.c100038>
- [9] "Yeebyee 100mm Effective Stroke Length Linear Rail Guide with Square Linear Rails Ballscrew SFU1605 with Nema 17 Stepper Motor for DIY CNC Router Parts X Y Z Linear Stage Actuator (Size:100mm): Amazon.com: Industrial & Scientific." [Online]. Available: <https://www.amazon.com/Yeebyee-Ballscrew-SFU1605-Stepper-Actuator/dp/B0BLVND5Z?th=1>. [Accessed: 27-Apr-2023].
- [10] A. K. Mahapatra, D. L. Harris, C. N. Nguyen, and G. Kannan, "Evaluation of an IUL Flash & Go automated colony counter," *Agric. Eng. Int.*, vol. 11, pp. 1–6, 2009.
- [11] E. GARRY, G. OUATTARA, P. WILLIAMS, and M. PESTA, "ENUMERATING CHROMOGENIC AGAR PLATES USING THE COLOR QCOUNT AUTOMATED COLONY COUNTER," *J. Rapid Methods Autom. Microbiol.*, vol. 17, no. 1, pp. 46–54, Mar. 2009.
- [12] C. Guzmán, M. Bagga, A. Kaur, J. Westermarck, D. Abankwa. (2014, Mar.). ColonyArea: an ImageJ plugin to automatically quantify colony formation in clonogenic assays. *PloS One.* [Online]. 9(3), e92444. Available: <https://pubmed.ncbi.nlm.nih.gov/24647355/>
- [13] Jiloc, "Jiloc/EDSDK-Python: Canon Edsdk Wrapper for python," GitHub, <https://github.com/Jiloc/edsdk-python> (accessed May 19, 2023).
- [14] J. Chalfoun et al., "Mist: Accurate and scalable microscopy image stitching tool with stage modeling and error minimization," *Scientific Reports*, vol. 7, no. 1, 2017. doi:10.1038/s41598-017-04567-y
- [15] T. Blattner et al., "A hybrid CPU-GPU system for stitching large scale optical microscopy images," 2014 43rd International Conference on Parallel Processing, 2014. doi:10.1109/icpp.2014.9
- [16] J. Schindelin et al., "Fiji: An open-source platform for biological-image analysis," *Nature Methods*, vol. 9, no. 7, pp. 676–682, 2012. doi:10.1038/nmeth.2019