Automated Cell Survival Colony Counting for Clonogenic Assays

Wanrong Mona Wang[1,2] WANRONG.WANG@MAIL.MCGILL.CA
[1] *Department of Bioengineering, Faculty of Engineering, McGill University, Montréal, Québec, Canada*
[2] *Medical Physics Unit, Gerald Bronfman Department of Oncology, McGill University, Montréal, Québec, Canada*

Joanna Li[*2,3] JOANNA.LI@MAIL.MCGILL.CA
Laya Rafiee Sevyeri[2,3] LAYA.RAFIEESEVYERI@MAIL.MCGILL.CA
Shirin A. Enger[*2,3] SHIRIN.ENGER@MCGILL.CA
[3] *Lady Davis Research Institute for Medical Research, Jewish General Hospital, Montréal, Québec, Canada*

Editors: Under Review for MIDL 2025

Abstract

Assessment of cancer cell radiosensitivity is essential for understanding the effectiveness of radiotherapy. The clonogenic assay is the gold standard for quantifying radiosensitivity by enumerating survived cell colonies *in vitro* post-radiation exposure. However, the technique is time-consuming and subject to observer variability. In this study, we present a computer vision pipeline to automate colony counting, tested on 787 HCT116 assay images. Our method addresses challenges: (i) lack of annotated data, (ii) variability in data collection, and (iii) unavailability of microscopic images of individual colonies. Using our augmented marker-based watershed algorithm, the pipeline achieved an average of 78.5% accuracy compared to the ground-truth colony count. Ongoing work focuses on improving the robustness of the pipeline and validating the approach across more cell lines.

Keywords: Computer vision, segmentation, cell counting, clonogenic assay

1. Introduction

The radiosensitivity of cancer cells is the response of the tumour to irradiation that can be measured by the extent of regression (Yashar, 2018). Radiosensitivity is influenced by various biological factors such as cell cycle position and the degree of metastasis, which differ significantly between cancer types and even individuals (Yashar, 2018). Assessing the radiosensitivity of given cells is crucial for understanding the radiotherapy efficacy, which eventually can lead to a better treatment plan for the patient. In radiobiology, the clonogenic assay is regarded as the gold standard for quantifying radiosensitivity *in vitro* by evaluating the cell reproductive abilities post-radiation exposure (Backer et al., 2011; Franken et al., 2006; Barber et al., 2001). Clusters containing more than 50 cells are considered colonies (Backer et al., 2011; Franken et al., 2006) and are counted to generate survival fraction curves, which describe the relationship between radiation dosage and the proportion of surviving cells.

Currently, colony counting in the clonogenic assay is done manually by observing individual colonies through a stereomicroscope (Barber et al., 2001; Lavitt et al., 2021; Zhang, 2022). Hence, this process is laborious, time-consuming and prone to interobserver variability.

In this study, we present a computer vision-based pipeline to automate the clonogenic assay counting process. We compiled 787 HCT116 cell line images, each with the ground truth colony count, radiation qualities[1] and the replicate information. The pipeline addresses the following challenges:(i) the lack of annotated data due to the tedious nature of annotating the wells; (ii) variability in data collection, leading to differences in images that require preprocessing; and (iii) the unavailability of microscopic-resolution images of cell clusters, which makes cell counting difficult, thereby necessitating reliance on morphological features.

2. Methodology

In preprocessing, a common challenge in automated cell counting is the edge effect, where cells near the well boundary are challenging to distinguish from the edge due to similar color intensity (Zhang, 2022; Dahle et al., 2004). To address this, we used OpenCV's Hough Circles[2] algorithm (Bradski, 2000) to detect and isolate wells in each image. A binary mask was applied to mask out non-assay regions; then, the images were cropped to the bounding box of the assays. Images were then resized to the medium dimension of the dataset for consistency (see Figure 1). This pipeline standardizes inputs for downstream processing.

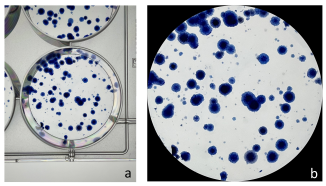

Figure 1: Clonogenic assay well image, a. before preprocessing and b. after preprocessing

We developed a colony enumeration pipeline, which begins with assay background masking to isolate cell clusters. Thresholds for masking were manually selected at colony edges, taking advantage of the high contrast between colonies and the background. Then, the masks were refined using OpenCV morphological closing[3] with a 3×3 kernel to smooth edges and suppress noise caused by thresholding (Bradski, 2000). Non-colony cell clusters were removed from the mask with a manually tuned area threshold of 400 pixels using `cv2.connectedComponentsWithStats`. An "overlay" mask was then generated by overlaying the masks on top of the original image using `cv2.bitwise_and`.

Colony aggregation, where bordering colonies appear as a single blob in a binary mask, was addressed with an augmented marker-based watershed algorithm shown in Figure 2. Foreground and background extraction are required for the algorithm. The background was generated via binary mask dilation (Figure 2 D). For the foreground generation, to segment the colonies, we leveraged the fact that the bigger HCT116 cell colonies are characterized by a darker inner core and a lighter outer ring (Figure 2 B). We applied a threshold to isolate the core, followed by morphological opening and `cv2.connectedComponents` for segmentation. From the original binary mask, small colonies were isolated by filtering out

1. Radiation quality is the radiation source used during irradiation experiment.
2. Hough Circles algorithm detects circles in an image within a range of radii; this range is manually defined.
3. Morphological closing is a process where the mask undergoes dilation followed by erosion (Bradski, 2000).

the large blobs; the resulting mask was concatenated with segmented large colonies to form the foreground (Figure 2 G). The foreground and the background were inputted into the watershed, producing the segmented binary masks (Figure 2 H).

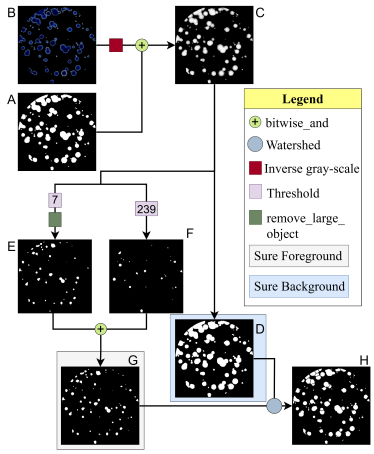

Figure 2: Flow diagram of the CV-based segmentation pipeline. (A) Raw binary mask. (B) Original image overlaid with a binary mask. (C) inverted gray-scale image of B. (D). Sure background from C for the watershed algorithm. (E) Sure foreground 1 from C. (F) Sure foreground 2 from C. (G) Sure foreground from E and F for the watershed algorithm. (H) Segmented mask.

3. Results

To evaluate the accuracy of the segmented masks, we used the OpenCV function `cv2.moments`[4] to count the number of segmented colonies. The observed count was then compared with the ground truth, and the accuracy was calculated with the following equation:

$$\%accuracy = |\frac{count_{observed} - count_{true}}{count_{true}}| \times 100\%$$

The accuracy of the dataset, categorized with radiation qualities, was tabulated in Table 1.

Results					
Radiation Qualities	Replica	Accuracy (%)	Radiation Qualities	Replica	Accuracy (%)
180 kVp	1	77.07	225 kVp	1	77.59
	2	80.76		2	86.87
	3	77.58		3	81.88
	4	70.07		4	92.46
	5	70.96		5	79.33
	6	77.96		6	79.42
80 kVp	1	73.71		1	83.71
	2	79.94		2	78.69
	3	88.55		3	82.1
	4	82.57		4	75.33
	5	76.57		5	80.34
	6	78.83	Ir-192	4	83.04
50 kVp	1	76.33		5	86.39
	2	76.44	6 MV	1	86.69
	3	63.59		2	83.33
	4	25.14		3	85.34
	1	77.75		4	84.9
	2	88.18		5	82.98
	3	76.3			
	4	75.38		**Overall:**	78.53

Table 1: Generated masks accuracy categorized with radiation qualities

Overall, the accuracy of the pipeline on the dataset is 78.5%. From Table 1, we can observe that the algorithm performed better on some experiment replicas than others; however, there is no correlation between radiation qualities and accuracy. Our ongoing work will focus on improving the robustness of the pipeline and validating the approach across more cancer cell lines.

4. cv2.moments detect the center of a contoured region, but can be used to count the number of centers.

References

M. V. Backer, J. M. Backer, and P. Chinnaiyan. Chapter three - targeting the unfolded protein response in cancer therapy. In P. Michael Conn, editor, *The Unfolded Protein Response and Cellular Stress, Part C*, volume 491 of *Methods in Enzymology*, pages 37–56. Academic Press, 2011. doi: https://doi.org/10.1016/B978-0-12-385928-0.00003-1. URL https://www.sciencedirect.com/science/article/pii/B9780123859280000031.

Paul R Barber, Borivoj Vojnovic, Jane Kelly, Catherine R Mayes, Peter Boulton, Michael Woodcock, and Michael C Joiner. Automated counting of mammalian cell colonies. *Physics in Medicine Biology*, 46(1):63, jan 2001. doi: 10.1088/0031-9155/46/1/305. URL https://dx.doi.org/10.1088/0031-9155/46/1/305.

G. Bradski. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000.

Jostein Dahle, Manish Kakar, Harald B. Steen, and Olav Kaalhus. Automated counting of mammalian cell colonies by means of a flat bed scanner and image processing. *Cytometry Part A*, 60A(2):182–188, 2004. doi: https://doi.org/10.1002/cyto.a.20038. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/cyto.a.20038.

N. A. Franken, H. Rodermond, J. Stap, J. Haveman, C. V. Bree, and H. Zheng. Clonogenic assay of cells in vitro. *Nature Protocols*, 1:2315–2319, 2006. doi: https://doi.org/10.1038/nprot.2006.339. URL https://www.nature.com/articles/nprot.2006.339.

Falko Lavitt, Demi J. Rijlaarsdam, Dennet van der Linden, Ewelina Weglarz-Tomczak, and Jakub M. Tomczak. Deep learning and transfer learning for automatic cell counting in microscope images of human cancer cell lines. *Applied Sciences*, 11(11), 2021. ISSN 2076-3417. URL https://www.mdpi.com/2076-3417/11/11/4912.

Catheryn M. Yashar. 23 - basic principles in gynecologic radiotherapy. In Philip J. DiSaia, William T. Creasman, Robert S. Mannel, D. Scott McMeekin, and David G. Mutch, editors, *Clinical Gynecologic Oncology (Ninth Edition)*, pages 586–605.e3. Elsevier, ninth edition edition, 2018. ISBN 978-0-323-40067-1. doi: https://doi.org/10.1016/B978-0-323-40067-1.00023-1. URL https://www.sciencedirect.com/science/article/pii/B9780323400671000231.

L. Zhang. Machine learning for enumeration of cell colony forming units. *Visual Computing for Industry, Biomedicine, and Art*, 5, 2022. ISSN 2076-3417. URL https://www.mdpi.com/2076-3417/11/11/4912.