Accurate and Accessible Organoid Analysis with OrganoSeg2

A Technical Paper submitted to the Department of Biomedical Engineering

Presented to the Faculty of the School of Engineering and Applied Sciences

University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

University of Virginia, Department of Biomedical Engineering

Spring 2025

Cameron Wells

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

Advisors: Najwa Labban, Róża Przanowska, PhD

Faculty Advisor: Kevin Janes, PhD

Accurate and Accessible Organoid Analysis with OrganoSeg2

Cameron Wells[a], Najwa Labban[a], Róża K. Przanowska[a], Kevin A. Janes[a,b,1]

[a] Department of Biomedical Engineering, University of Virginia, Charlottesville, Virginia 22908, United States.

[b] Department of Biochemistry and Molecular Genetics, University of Virginia, Charlottesville, Virginia 22908, United States[b]

[1] Correspondence: kjanes@virginia.edu

Abstract

Organoids have great power to model diseases and treatment response while capturing patient-to-patient heterogeneities. Characterizing the behavior of organoids requires automated and unbiased segmentation methods to quantify their size and shape. Organoid segmentations also provide a means to identify biological states, such as cell death, at the individual organoid level. These features evaluate the effectiveness of treatments that act by different mechanisms. It is important that methods for organoid analysis are built with human factors in mind, as there is no value in a perfectly accurate program that is impossible to use. Previously, our lab created OrganoSeg as a user-friendly platform to segment organoids. OrganoSeg is a popular tool, but is limited to population-level morphological analysis and struggles with certain formats of organoid images. Alternative options for organoid analysis exhibit deficiencies in usability, accuracy, or functional capacity. Here, we introduce OrganoSeg2 to accurately collect organoid data related to biomedical research settings. OrganoSeg2 provides a graphical user interface that segments, tracks, and quantifies fluorescence of individual organoids, without requiring training data or prior coding knowledge. We find that OrganoSeg2 significantly improves segmentation accuracy from its predecessor, and is highly generalizable to different organoid types in comparison to alternative segmentation platforms. We applied OrganoSeg2 to collect individual-organoid fluorescence data and identify variable patterns of organoid death in response to radiotherapy, within and between cases of patient-derived organoids of breast cancer. OrganoSeg2 successfully addresses deficiencies in organoid segmentation technology and presents a versatile tool for studying disease with organoids.

Keywords: organoids, image segmentation, graphical user interface, breast cancer, radiotherapy

Introduction

Organoids are 3-dimensional structures derived from stem cells or adult tissue that self-organize and mimic the behavior of biological tissues[1]. Organoids capture biological complexity and diversity without the constraints of in vivo experimentation[2]. More specifically, patient-derived organoids capture patient heterogeneities and recapitulate patient phenotypes in terms of molecular composition[3,4] and response to treatment[5]. One way to assess these phenotypes is by quantifying organoid morphology. For example, area is an indicator of disease progression for tumor organoids, while morphological features such as circularity or intensity are indicators of distinct biological states[6,7]. Interpreting these morphological readouts is valuable for planning effective treatment or screening for molecular targets across diseases[6,8].

The large-scale collection of morphological data from organoid cultures offers insight into diseases and better treatment regimens for patients. Collecting such data requires organoid segmentation - identifying and outlining the region of organoids within an image. A single culture may contain several hundred organoids, and manually segmenting these images would introduce time burdens and low replicability that would interfere with research or therapeutic potential. Automated segmentation enables efficient and unbiased collection of data[9]. Thus, it is important to have a platform for biomedical researchers to

segment organoids accurately and efficiently and collect morphological data.

Several platforms for biological image analysis existed before their use in organoid research, serving to segment cells or other 2-dimensional image features[9,10]. However, organoids grow in a 3-dimensional hydrogel with different focal planes, presenting barriers to segmentation of organoid data. Our lab previously developed OrganoSeg[11] to address this problem. OrganoSeg uses conventional image processing techniques geared towards brightfield organoid images. Upon publication, OrganoSeg demonstrated better organoid detection and segmentation than image analysis platforms ImageJ, CellProfiler, and MorphoLibJ[11]. Those results indicate the benefit of an organoid specific segmentation platform, arising from the multi-windowed adaptive thresholding technique employed by OrganoSeg. OrganoSeg remains a popular tool, having been cited over 100 times in the past 5 years and used as a part of studies of treatment response[12] and disease modeling[13].

Alternative platforms for organoid segmentation have since developed, eliciting comparisons of OrganoSeg with said alternatives[14–16]. Many of the alternatives use machine learning techniques[17], claiming higher accuracy and generalizability. They also suggest that conventional image processing techniques are not suitable as they require parameter tuning for each image or set of images[18,19]. However, machine learning platforms present multiple barriers. Firstly, machine learning requires large amounts of training data to be successful. For certain scales of experiments, it is not practical or feasible to manually segment sufficient images for training data. Secondly, machine learning platforms are difficult to use for researchers with minimal coding experience. Tuning training algorithms is challenging given their black-box nature. Additionally, most of these programs require use of a command-line interface. OrganoSeg requires no training data and is in the format of a graphical user interface (GUI), which users simply load images into for segmentation. Although machine learning presents advantages, it is crucial to maintain an accessible means for organoid segmentation.

To address deficiencies of the original OrganoSeg, our lab developed OrganoSeg2, an updated GUI using the same base algorithm and interface (Supplementary Figure 1). OrganoSeg2 has increasing GUI and workflow flexibility for users, including options to resize image panels, perform in-app brightness adjustments, and work on multiple sets of images at once. OrganoSeg2 improves efficiency by lowering the runtime for segmentation, data exporting, and image display, and by introducing options for post-segmentation tuning that attempt to gather accurate organoid data in a more streamlined manner. In addition to improvements to usability, OrganoSeg2 has an updated segmentation algorithm that improves on edge recognition in organoid images with uneven lighting or focus. Finally, certain constant parameters in were made customizable to increase user control over segmentation for heterogenous image sets. Here, we show that the improved GUI is comparable to alternative segmentation programs and capable of collecting organoid morphological data in a wide array of research settings.

In addition to changes to the segmentation process, OrganoSeg2 includes new options for downstream analysis using individual organoid tracking and fluorescence quantification. In its original state, individual organoid data is not connected between images, and therefore any longitudinal data collected can only be assessed at the population level. Individual organoid tracking increases the statistical power of results and highlights sub-population trends. Additionally, OrganoSeg was not able to report any molecular information on organoids, but with fluorescence quantification, morphological and molecular information can be linked. These extended capabilities had not previously been tested in a research setting, but we hypothesized that they would be informative in monitoring breast cancer organoids throughout exposure to irradiation. Our lab previously used OrganoSeg to quantify breast cancer organoid growth in response to treatment with drugs such as tamoxifen[20]. Tamoxifen exhibits cytostatic mechanisms, making organoid size an important readout[21]. However, radiation treatment is primarily cytotoxic, meaning that cell death is more relevant[22]. Using fluorescent markers for caspase activation and thus cell death, OrganoSeg2 collects this information longitudinally.

Results

Validation of Segmentation Accuracy

Edge Correction Improves Perimeter Recognition in OrganoSeg2

Since the changes to the OrganoSeg algorithm focused on edge detection, we compared OrganoSeg and OrganoSeg2 in their recognition of spheroid perimeters. We used spheroid images for which OrganoSeg was previously validated to have high call rate and area segmentation

accuracy compared to traditional cell segmentation platforms, but exhibited some failures in edge detection (Figure 1a). These failures occur when glare causes the organoid surface to be brighter than the background. OrganoSeg2's algorithm accounts for this glare using background-adjusted gradient thresholding, so that changes at the organoid edge are detected whether they are bright or dark. Using edge correction, OrganoSeg2 showed a significant improvement in the colocalization of the segmented perimeter with the manually traced perimeters (Figure 1b). By improving perimeter recognition, OrganoSeg2 enables greater representation of morphological features such as area, circularity, eccentricity, and solidity, which convey descriptive phenotypic information about organoids/spheroids.

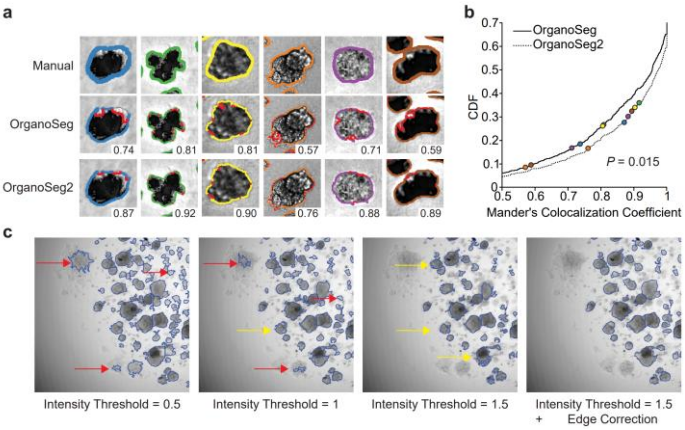


Figure 1. OrganoSeg2 improves edge recognition. a) Perimeter segmentation from OrganoSeg and OrganoSeg2 were compared to manual segmentation. Red regions of segmentation lie outside of manually segmented target region. Number is Mander's Colocalization Coefficient (MCC) **b)** CDF plot of distribution of MCCs from OrganoSeg and OrganoSeg2 perimeters. Colored dots correspond to representative spheroids in **a. c)** Representative image of poor segmentation with changing intensity threshold alone, and improvement with edge correction. Red arrows show over segmentation and yellow arrows show under segmentation

Edge correction also facilitated the segmentation of more complicated breast cancer organoid cultures. Wells containing hundreds of organoids are highly informative, but also present complications due to crowded/overlapping organoids, different focal planes, and varying brightness levels. Using OrganoSeg's original options, the primary parameter controlling sensitivity is the intensity threshold. Varying this parameter alone, it is difficult or impossible to find a threshold which segments organoid edges accurately while avoiding oversegmentation (i.e., non-organoid artifacts or out-of-focus organoids which are morphologically inaccurate). Applying a higher intensity threshold in combination with edge correction alleviates this problem (Figure 1c). The high intensity threshold improves specificity to eliminate incorrect organoid identifications, and edge correction then increases sensitivity of edge detection, only for the organoids which pass the intensity threshold. Edge correction successfully reduced the amount of post-segmentation fine-tuning needed to analyze breast cancer organoid cultures.

OrganoSeg2 outperforms alternative segmentation platforms across organoid types

We compared the segmentation accuracy of OrganoSeg2 to OrgaExtractor[23], OrganoID[18], and OrganoLabeler[16] (Figure 2). Comparisons were made using the datasets which these platforms were originally tested on, including colon, lung, pancreas, brain, and embryoid organoids. This diverse set of organoid images was used to test generalizability and to identify the strengths and weaknesses of OrganoSeg2 relative to the other platforms. OrganoSeg2 and OrganoLabeler required parameter tuning for each image set, as well as for subsets of the pancreatic ductal adenocarcinoma (PDAC) organoids, representing one notable concern for platforms which use conventional image processing techniques. However, the default models for machine learning platforms OrgaExtractor and OrganoID performed poorly in most cases. Whether or not machine learning was used, the dissimilar set of images introduced pre-processing burdens for all platforms.

To quantify segmentation quality, we first calculated the intersection-over-union for whole-image segmentations. This metric describes how well the segmenter recognizes pixels as belonging to an organoid, and is m inimally affected by differences in organoid splitting. OrganoSeg2 only had significantly lower segmentation quality in 1/22 comparisons for whole-images and was significantly better than other segmenters in 15/22 comparisons, respectively (Figure 2a). The only comparison in which OrganoSeg2 was outperformed was when OrgaExtractor segmented colon organoids, its internal dataset. In the colon organoid images, OrganoSeg2 appears to identify many regions which are not part of the manual segmentation. However, manual inspection suggests that many of these regions are in fact organoids and are recognized by multiple segmenters (Figure 2b). These initial results suggest that OrganoSeg2 is not only an appealing option due to its usability, but its strong organoid recognition abilities relative to other platforms.

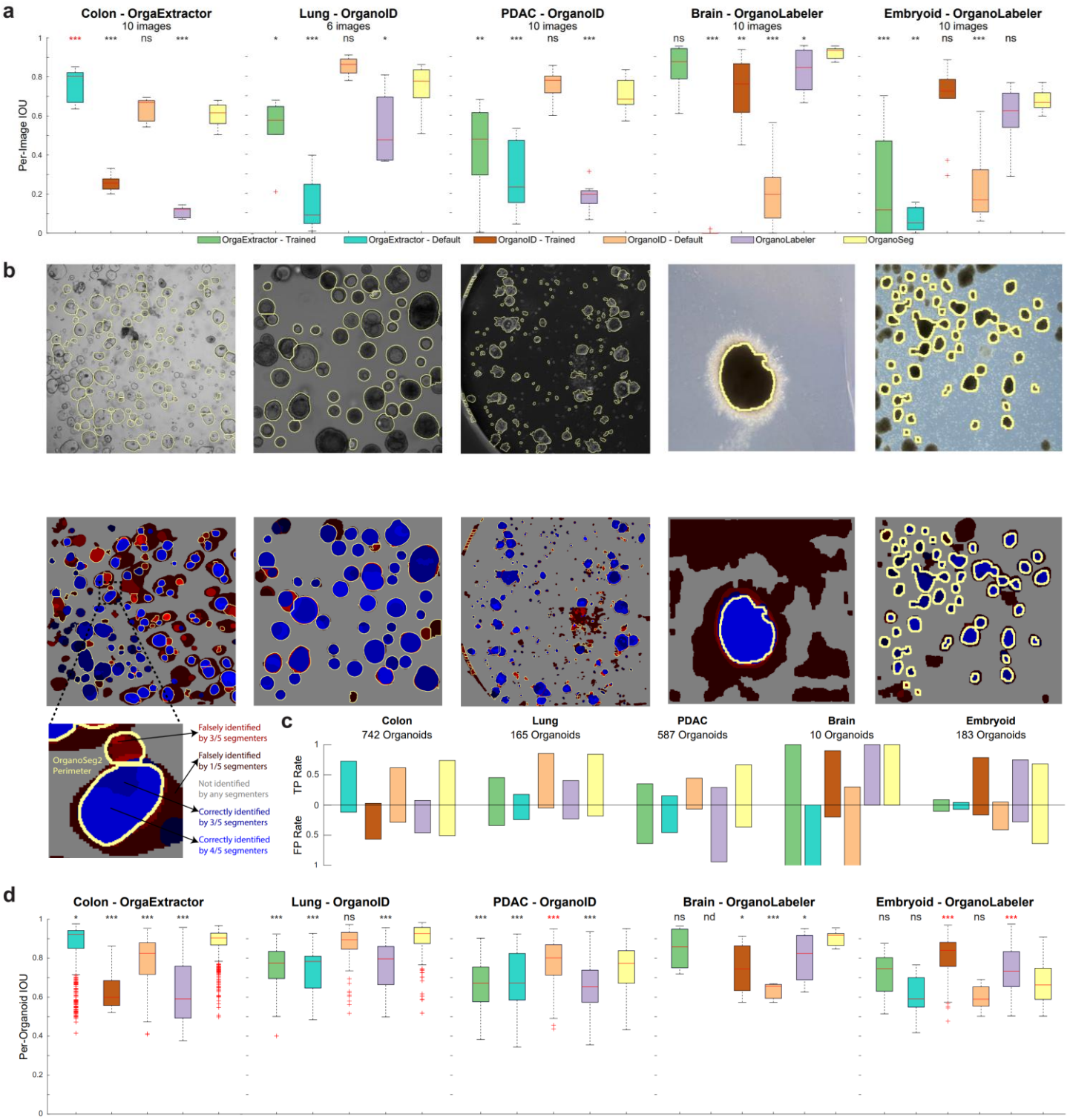


Figure 2. OrganoSeg2 compares favorably to alternative segmentation platforms. a) Boxplots of intersection-over-union (IOU) of whole-image segmentations for each segmenter across five organoid image types. ns – not significant, nd – no data, **P*<0.05, ***P*<0.01, ****P*<0.001, red asterisks indicate OrganoSeg2 was outperformed. **b)** Representative overlay images and segmentation masks with data aggregated from all segmenters. Yellow outline shows OrganoSeg2 segmentation. Bottom: Blue regions were identified by segmenters and manual segmentation, red regions were identified by segmenters but not manual segmentation. Increasing brightness indicates more segmenters identified this region, as indicated by the inset. **c)** Bar charts showing rates of true positive (positive y-axis) and false positive (negative y-axis) organoids, based on manual segmentation. Values are scaled to the total number of organoids in the manual segmentation for each group, indicated above the charts. **d)** Boxplots of IOU for individual organoid segmentations.

We also compared segmenting platforms for individual organoid segmentation quality, as the primary goal of each of these platforms is to accurately represent individual organoid morphology. These measurements also ignore thoroughness of labeling in manual segmentations, only focusing on the quality of organoids that appear in both manual and automated segmentations. We first matched organoids with sufficient overlap between manual and automated segmentation, denoting these as true positives (TP). Across all conditions, OrganoSeg2 correctly identified the most or nearly (within 15%) the most organoids. (Figure 2c). We then quantified IOU for each of the TP organoids, and OrganoSeg2 outperformed other segmenters in 13/21 comparisons, and was only outperformed in 3/21 (Figure 2d). Only one segmenter outperformed OrganoSeg2 on an external dataset (OrganoID segmenting embryoid bodies). Additionally, no segmenter outperformed OrganoSeg2 in both whole-image and individual-organoid comparisons. The weakest part of OrganoSeg2's performance was its high relative recognition of "false positives" compared to segmenters using their internal data sets. (Figure 2c). However, some of these false positives are due to inconsistencies in manual segmentation. Others are legitimate, such as well-edges or out-of-focus objects, which OrgaoSeg2 cannot distinguish as non-organoids. That said, these images were compared without any manual post-segmentation tuning. Using OrganoSeg2's improved post-segmentation tools, many of these false positives are easily removed manually by the user. In all, OrganoSeg2 is a generalizable option for collecting large quantities of organoid morphological data.

Application to Longitudinal Fluorescence Tracking

We sought to test OrganoSeg2's extended capabilities for downstream analysis by monitoring cell death in breast cancer organoids over the course of irradiation treatment. 5 cases of patient-derived organoids were cultured for 14 days, with exposure to 0 Gy, 1 Gy, or 6 Gy of radiation for 5 days. We monitored cell-death using two stains, Nucview 488 Caspase-3 and Saguaro LIVE-Dead stain, in separate cultures. We also introduced dilute DAPI into live cultures on day 14 to observe endpoint cell death based on compromised plasma membrane integrity and see if there was concordance between DAPI and the other fluorescent markers. The goal of this experiment was to see if organoid cultures responded differently to varying levels of irradiation, and if the data collected by OrganoSeg2 could be used to glean patient-to-patient heterogeneities in response to irradiation therapy.

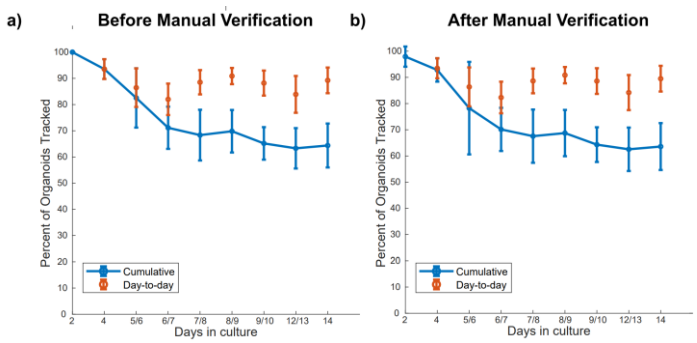


Figure 3. OrganoSeg2 tracks the majority of organoids across two-week organoid cultures a) Retention rate of organoids from day-to-day (orange) and across all organoids introduced by a given time point (blue). **b)** Retention rates after manual removing poor organoid traces.

To observe the development of cell death at the individual organoid level and pair observations between DAPI and the other fluorescent markers, we longitudinally tracked organoid cultures. Organoids were segmented using updated parameters and edge correction, registered, and matched by Euclidean distance. On average 80-95% of organoids were matched between consecutive images. Cumulatively, over 60% of the organoids that were introduced at some point in the 14-day culture were accounted for on the last day (Figure 3a). To ensure that these matchings were accurate, we added a manual verification option to the OrganoSeg2 tracking window, so that users could mark organoid traces as good or bad while seeing all segmentation results at once (Supplementary Figure 2). Organoid segmentation was sometimes faulty (e.g., poor edge recognition or organoid splitting), but the segmentation captured the correct organoid and would still be useful for fluorescence analysis. Therefore, we only removed organoid traces which contained obvious artifacts or distinct organoids. After removing traces via manual verification, the variance in organoid retention increased mildly but was minimally affected on average (Figure 3b).

After tracking, we quantified the cell-death fluorescence intensity of each organoid, seeking a two-state distribution of alive and dead organoids (Figure 4a). Fluorescent dye was added on the first day of irradiation (day 5), but the correlation of individual organoid fluorescent intensities between days 5 and 6 was drastically lower than all other pairs of consecutive days, so only days 6-14 were considered for fluorescence analysis (Supplementary Figure 3a). Initial quantification of fluorescence intensity over time showed noticeable fluorescent drift for both the Nucview and Saguaro stains, indicating that raw fluorescence data would not be suitable (Supplementary Figure 3b).

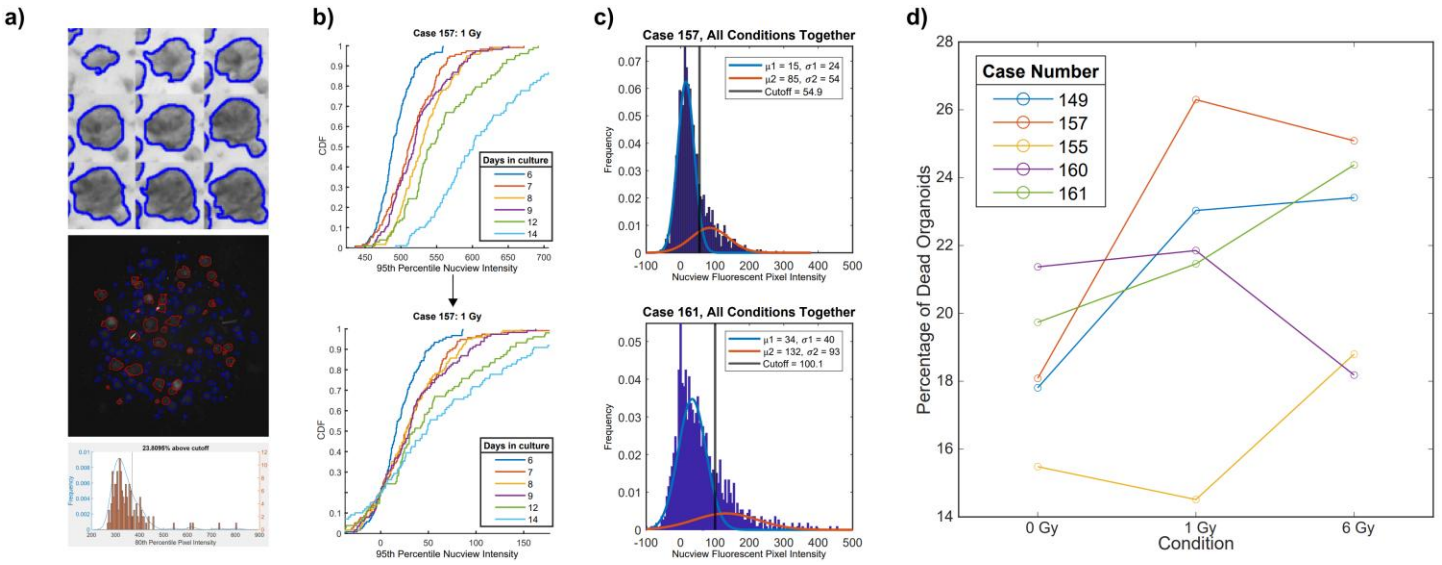


Figure 4. OrganoSeg2 facilitates longitudinal fluorescence analysis of breast cancer organoids. a) Individual organoid tracking (top) and fluorescence analysis (middle and bottom) as they are displayed in OrganoSeg2. **b)** Representative raw fluorescence data processed with a 20th percentile shift to account for differences between images. **c)** Representative aggregated data fit using a gaussian mixture model to determine a case-by-case cutoff. **d)** Total percentages of dead organoids by case and condition.

Additionally, most images had no clear two-state distribution of organoids, so applying a cutoff based on the in-app visualizations would be insufficient (Supplementary Figure 3c). We first addressed the fluorescent drift by shifting each image's data based on its 20th percentile organoid fluorescence (Figure 4b). This shift resulted in a steady progression in fluorescence intensity over time for Nucview. The Saguaro stain generally exhibited fading, except on day 12 when dye was refed, indicating that it may not be a strong option for continuous fluorescence monitoring. We decided to progress with longitudinal fluorescence quantification using only Nucview.

Individual images did not have enough organoids for a two-state distribution to emerge, but using the normalized fluorescence data, we aggregated data across days and conditions to find case-wide cutoffs. All cases exhibited a right-tailed distribution, but none were strongly bimodal (Supplementary Figure 4). We fit data with a Gaussian mixture model, where the lower, and generally larger, distribution represents the alive organoids. We used the 95th percentile of this distribution, and although this may sacrifice an overlapping portion of the dead organoid distribution, we estimated that this would provide enough dead organoids with only a few overlapping alive organoids. We successfully fit each case to a two-state distribution, defined case-wide cutoffs, and calculated the percent positivity for each case and condition (Figure 4c-d, Supplementary Figure 4). As expected, the average percent positivity increased with an increasing dose of irradiation, but these results were not significant given the limited sample size. However, heterogeneous behavior was observed between cases. For example, cases 149 and 157 showed large jumps from 0 Gy to 1 Gy but not 1 Gy to 6 Gy, while case 155 unexpectedly decreased at 1Gy before a large jump at 6 Gy. These results demonstrate the tolerance of our pipeline to different underlying distributions of fluorescence intensity, as well as its ability to assess variable treatment responses.

To supplement population-level fluorescence analysis, organoid tracking enabled comparisons of image-to-image fluorescence for individual organoids, which was useful for correcting, validating, and expanding on the results observed by Nucview. Firstly, we defined individual organoid traces of fluorescence positivity. In general, most organoids stayed positive for multiple days, but some went from positive to negative. We used organoid tracking to determine that this was due to fading dye (Figure 5a) and set an irreversible cutoff so that once an organoid was positive, it was considered dead for the remaining images. We then recalculated the percentage of death by case and condition, and found that fading dye had different effects on a case-by-case basis (Supplementary Figure 5) Secondly, the cultures containing Nucview also had DAPI measurements on the last day. Using the same pipeline as before, we defined case-by-case cutoffs and identified positive organoids for these endpoints. We then paired

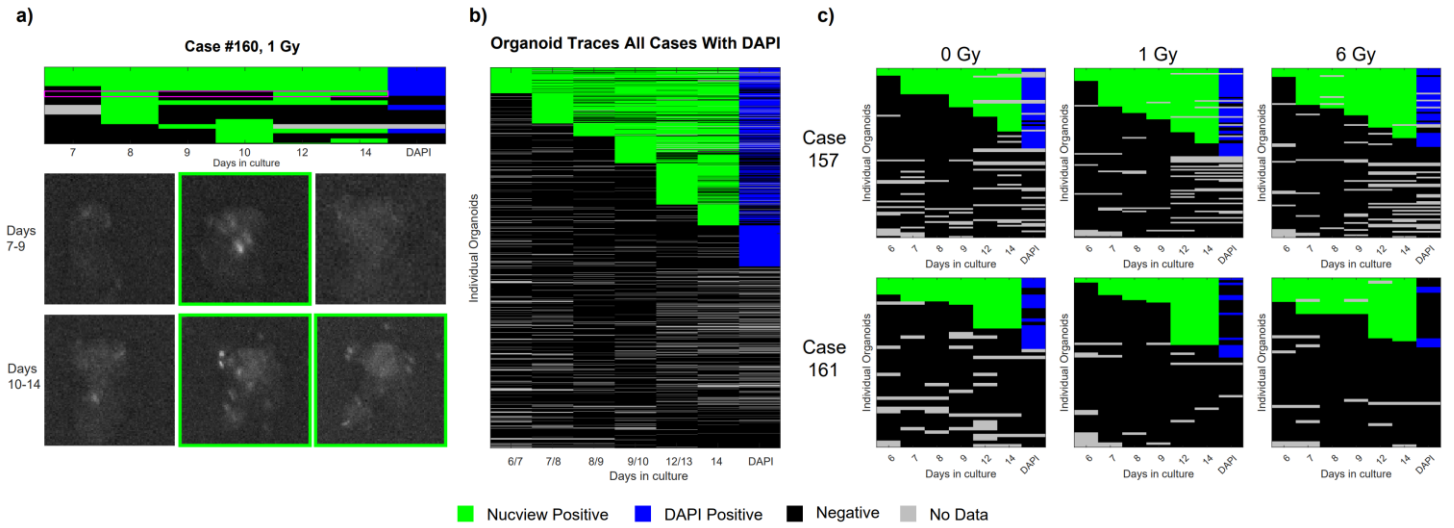


Figure 5. Individual organoid tracking provides a more detailed analysis of fluorescence results. a) Subset of individual organoid traces showing a switch in positivity (top) and corresponding images which display a clearly fluorescent organoid that faded and regained positivity (bottom) **b)** All organoid traces across cases and conditions show correspondence between Nucview and DAPI **c)** Two cases of organoid traces showing different patterns in the progression of organoid death.

DAPI data to the corresponding Nucview data. Out of 1374 organoids, 570 were Nucview positive at some point, 469 were DAPI positive, and 321 were positive by both (Figure 5b). This result is highly significant by a hypergeometric test ($P \approx 0$), indicating strong concordance between these two markers. Thus, individual organoid tracking enabled inspection of fluorescent data which adds confidence to the performance of our pipeline. Finally, organoid tracking provides a dynamic picture of the progression of organoid death. For example, case 157 and 161 have similar population levels of death, but case 161 appears to have minimal death after day 12 while case 157 has steadily increasing death at this point. Using individual organoid tracking, we were able to perform a more detailed analysis of fluorescence data.

Discussion

The improved OrganoSeg2 algorithm enhances segmentation results in direct comparisons and in applications to downstream analysis

We demonstrated that OrganoSeg2's use of edge correction and expanded parameter options creates segmentation results which are not only comparable, but better than other segmenters. As previously stated, a common critique of conventional techniques such as OrganoSeg is that parameter tuning is required to operate on different images or sets. However, these results indicate that for a given organoid image type, the same parameters provide high levels of accuracy. The only set which was subdivided into different parameters was the PDAC organoids, as they included a variety of image formats. Given images of similar formats, as would be expected from an experimental procedure, a uniform set of parameters is sufficient.

The high level of segmentation accuracy from OrganoSeg2 has downstream benefits in addition to capturing organoid morphology better. In the original breast cancer organoid experiments with OrganoSeg, a low intensity threshold was used to avoid incompletely segmented organoids. However, when these parameters were used for this experiment, it was clear that there were too many artifactual segmentations for tracking to be effectively applied. Applying a more stringent threshold has the side effect of removing some organoids from the analysis. However, the images we were working with had over a hundred organoids in most cases. Stricter parameters still captured a large quantity of these organoids and represent their development overtime more accurately, creating a higher quality data set.

Increasing the algorithmic complexity and number of parameters has an effect on usability. The parameters may be capable of creating better segmentations, but they are pointless if the user is unable to understand them enough to implement them. For example, OrganoLabeler uses parameters such as "factor," "clip limit," and "tile grid size." Documentation says these parameters relate to contrast adjustment and histogram equalization but provide no indication as to how adjusting the parameters may affect segmentation results. After experimenting with different levels, these parameters did not seem to have a linear effect

on segmentation sensitivity, making it difficult to adjust appropriately. Another issue arises when parameters interact with each other in a way that makes it difficult to discern their individual effects (i.e., multivariate processes). OrganoID, although primarily relying on machine learning, has the option to set parameters related to edge detection using Canny detection[24]. Canny detection requires setting a low and high threshold, and balancing these thresholds to get the desired result may take tedious combinatorial parameter tuning. With OrganoSeg2's parameters, the goal is to have each responsible for a distinct effect. For example, intensity threshold and edge threshold both increase sensitivity, but one affects the number of identified organoids while the other affects the extent of these organoids. Of course, these effects cannot be completely isolated, for example, the organoids identified by the intensity threshold are responsible for defining the boundaries between organoids which are preserved during edge correction. Still, restricting these parameters to separate processes and documenting their effects will help minimize complications.

Longitudinal fluorescence tracking demonstrates utility of OrganoSeg2 to recognize heterogeneities in response to radiotherapy

Using OrganoSeg2, we were able to create a pipeline for assessing organoid response to irradiation which demonstrated agreement with expected results and between two separate measures of organoid death. Given the limited sample size, we were not able to draw significant conclusions about the response to variable doses of irradiation, but the results generally show the expected trend of higher organoid death at higher doses. We also see that different cases of organoids may have varying levels of resistance, as certain cases see a jump in organoid death with any irradiation at all, while others require high doses of irradiation. Pairing a second fluorescent marker added confidence to these observations, since DAPI strongly agreed with Nucview in identifying dead organoids. In general, these results demonstrate the value of organoids to capture patient-to-patient heterogeneities that affect response to treatment, as well as the potential of OrganoSeg2 to accurately represent these heterogeneities.

The cell-death analysis was not done entirely in OrganoSeg2, but our pipeline should be easily adaptable for others using OrganoSeg2 for fluorescence analysis. Given the image-to-image variations in fluorescence and the necessity of aggregating all data for each case, we had to work with exported data rather than directly in the OrganoSeg2 GUI. Still, the pipeline of normalizing fluorescence between images and identifying a two-state distribution should be generalizable to most fluorescent data sources. By attaching our scripts for downstream analysis with OrganoSeg2, users will have a streamlined process to quantify their fluorescent data.

Limitations and Future Work

The machine learning platforms were largely ineffective without appropriate training data, so we attempted to train them to the best of our abilities. We obtained OrganoID models which were trained until convergence for each external dataset, but the results were generally loose around organoid boundaries (Supplementary Figure 6). This was likely a result of limited training data, given that OrganoID's default model was trained on 2000 images. These 2000 images were replicates of 50 original images which underwent a variety of transformations, so it would be informative to attempt training using a similar augmented dataset using the images from OrgaExtractor and OrganoLabeler. For OrgaExtractor, training was computationally expensive so only 10 epochs were executed for each training set. The results improved after training but were inconsistent between images. We will further assess OrgaExtractor quality by training until convergence using high-powered computing. Even so, the performance of these platforms on their internal datasets was not consistently better than OrganoSeg2, and we believe that this trend will hold if trained to the full extent on other datasets. This observation also underscores the amount of manual labor required to get these automated segmentation platforms operating to their full extent, making the parameter tuning required by OrganoSeg comparatively negligible.

The initial aims of this project included expanding the segmentation capabilities of OrganoSeg2 further, with an emphasis on improved splitting. Ultimately, this aim did not reach past qualitative stages of assessing different filters (e.g. Laplacian of Gaussian) on organoid aggregates. It would be beneficial to continue improving the OrganoSeg2 algorithm as this is a limitation that could be truly inhibiting in more challenging formats. That said, even the machine learning platforms appear to struggle with separating organoid aggregates. OrgaExtractor does not appear to have a mechanism for this, leaving somewhat jagged edges when two organoids are close. OrganoID identifies edge detection as an area of low confidence[18] and leaves several organoid aggregates together in the colon and lung images (Supplementary Figure 6). Given that machine learning is

not overwhelmingly preferable in this regard, it is worth developing conventional image processing techniques for organoid splitting.

We have only scratched the surface of what information may be obtained from the longitudinal fluorescence tracking. Future analyses will pair fluorescent and morphological data to see if certain organoid shape features correspond to early cell death. Using the individual organoid tracking, we may be able to observe if morphological changes occur before or after observed cell death on an individual organoid basis. If morphological changes precede cell death, we may be able to identify structural mechanisms which promote organoid death. With further experimentation, it is also possible to pair fluorescent markers for death with markers for proteins of interest. If the pipeline is successfully translated for multiple fluorescent markers, we would be able to identify if subpopulations of live and dead organoids have differential expression of these proteins, and how their expression develops over time. Since the Saguaro LIVE-Dead stain proved ineffective in our current application, we were unable to assess the translation of our pipeline to a second longitudinal fluorescent dataset. However, the underlying principle of searching for two distributions of organoids remains the same, and the transformations applied to the data were not based on any assumptions tied specifically to Nucview. Therefore, we expect that this pipeline will translate well to other fluorescent reporters.

Materials and Methods

Graphical User Interface Construction

The OrganoSeg2 GUI was built using MATLAB App Designer. Code for the main segmentation and metrics selection scenes were adapted from OrganoSeg made by Borten et al.[11] New windows were constructed for the brightness adjustments, individual organoid tracking, and fluorescence analysis features.

Segmentation Algorithm

The same adaptive thresholding algorithm from OrganoSeg was used to create the initial segments. Internal default parameters for image reconstruction, segmentation closing, border clearing, and aggregate splitting were changed to customizable parameters.

If edge correction is selected, this portion of the algorithm is applied after adaptive thresholding and post-segmentation processing, but before organoid splitting. The magnitude of the image gradient is taken using imgradient. If the user selects "Gradient Only" for edge correction, this gradient intensity at each pixel is compared to the user-defined edge threshold multiplied by the background intensity (taken using imfilter and fspecial with an average filter, neighborhood size = 500x500). Pixels which pass this threshold are added to the initial segmentation, and post-segmentation processing is applied again to fill holes and smooth the segmentation. If the user selects "Gradient Preserve Boundary," a background marker is created using a watershed transformation on the distance transformed initial segmentation. The same process is then applied as in "Gradient Only", except the pixels which lie on the background marker (dilated with imdilate, neighborhood size = 3x3) are set to zero so that separate organoids are not joined. If the user selects "Gradient + Watershed", the background marker is made as described in "Gradient Preserve Boundary", and the foreground marker is an eroded version of the initial segmentation (imerode, neighborhood size = 3x3). These markers are then used to perform marker-controlled watershed segmentation[25].

Selection of alternative segmentation platforms

A literature review was performed to find alternative segmentation platforms which were designed with organoids in mind and had similar capabilities for segmentation to OrganoSeg2. 16 platforms were identified which were either created specifically for organoids or explicitly tested on them in a proof-of-concept. Of these, four were excluded because they only output a bounding box of organoids as opposed to a direct outline of organoids, and are therefore limited in the morphological data they provide. Three more were excluded as they are designed for single-organoid images, requiring external programs to identify individual organoids in images of organoid cultures with multiple organoids. Three more were excluded as the software to run the program was not available. Of the six remaining, OrganoID, OrgaExtractor, and OrganoLabeler were chosen as three which had clear instructions, approachable implementations (i.e., options to execute without using a command line interface), and represented a mix of machine learning and conventional image processing techniques.

Edge Detection Comparison

OrgaoSeg2 was compared to OrganoSeg using spheroid images from the original OrganoSeg publication. For each spheroid, a target region around the spheroid perimeter was

defined using bwperim and imdilate (neighborhood size 5x5). For each spheroid in the manual segmentation, a spheroid in the automated segmentation was assigned to it if the intersection of pixels for the two spheroids was greater than 50% of the total pixels in each segmented object individually. Then, the MCC for each spheroid was calculated based on the fraction of perimeter pixels from the automated segmentation that overlapped with the target region. The distribution of MCCs was assessed using a Kolmogorov-Smirnov test.

Segmentation for Comparative Effectiveness Evaluation

Datasets of organoid images and manual segmentations were taken from the respective publications for each of the selected alternative platforms. The source and respective organoid type(s) are as follows: OrgaExtractor - Colon, OrganoID - PDAC and Lung, OrganoLabeler - Brain and Embryoid Body (EB). Images were grouped by organoid type for parameter-tuning and training. Colon and Lung+PDAC images were already grouped into training, validation, and testing groups, so the testing set of images were used for the comparison. The Brain and Embryoid images were not originally separated, so they were split into training/validation/testing sets of 111/12/10 and 139/16/10 respectively, with testing images selected to represent a diverse set of phenotypes.

OrganoSeg2 and OrganoLabeler require parameter-tuning, so separate parameter sets were selected for each group of organoids in an attempt to optimize segmentation accuracy. The PDAC organoids consisted of both phase-contrast and brightfield images, and also different levels of magnification. These images were separated into subgroups for parameter definition as it would be impractical to try and segment such different images with the same settings. Parameters for each image set are described in Supplementary Table 1. OrganoID also has optional parameters, which were adjusted for external images when necessary. Splitting was applied to all segmentations when given the option.

OrganoID and OrgaExtractor have pre-trained models using their respective datasets. Segmentation on external datasets with these pretrained models was inaccurate in most cases. New models were trained for each training set (Brain and Embryoid for both, Colon for OrganoID, and Lung+PDAC for OrgaExtractor). OrganoID was trained using default settings for 100 epochs, but all three models converged before 100 epochs. OrgaExtractor was only trained for 10 epochs due to computational constraints. Additionally, a subset of 50 images out of the 2000 from the OrganoID training dataset were used due to computational constraints.

Segmentation and training were performed as described in each platform's documentation, if possible. The OrganoLabeler code was modified to accommodate 16-bit unsigned integer images. The OrgaExtractor training code was modified to accommodate images smaller than 512x512 pixels. Code for each was modified so that inputs and outputs were handled similarly for each platform. Images which had a dimension size that was not a multiple of 8 caused errors when segmenting with OrgaExtractor. OrgaExtractor code was modified to resize images to alleviate the error, and was also modified to address other errors that arose during training.

Quantification of Segmentation Accuracy

Segmentation results were quantified in MATLAB by calculating the intersection-over-union (IOU) of whole-image segmentations and individual organoid segmentations. All segmentation outputs were converted to logical arrays and smoothed using imopen (neighborhood size = 5x5, used to remove artifacts in manual segmentations that likely arose from image size rescaling). For consistency across platforms, all segmentation outputs also had their borders cleared and all objects removed which were smaller than the smallest object in the corresponding manual segmentation (except in the case of brain organoids, as these each had one centrally located organoid which in some cases was unintentionally removed from segmentations with these adjustments). Segmentation outputs from OrgaExtractor which were resized during segmentation were scaled to their original size.

To calculate individual-organoid IOUs, organoids in the manual and automated segmentations were split into objects using bwconncomp. For each organoid object in the manual segmentation, an organoid in the automated segmentation was assigned to it if the intersection of pixels for the two organoids was greater than 50% of the total pixels in each segmented object individually. For each automated segmentation, "true positives" were counted as the number of organoids that were successfully matched to the manual segmentation, while all other organoid objects were counted as "false positives". P-values were calculated using a two-sample, two-tailed t-test for each segmenter in comparison to OrganoSeg2, for each organoid type.

Organoid Cultures

Breast cancer organoids were acquired and maintained in 14-day zero-passage cultures as described by Przanowska et al[20]. Starting on either day 5 or 6, organoids were treated with 0 Gy, 1 Gy, or 6 Gy of irradiation each day for 5 days. Cultures were fed with fluorescent reporters for cell death (Nucview 488 Caspase-3 or Saguaro LIVE-Dead stain) on the first and last day of irradiation. Brightfield and fluorescence microscopy images were taken either once or twice before irradiation, once on each day of irradiation, and twice after irradiation. On day 14, cultures were fed with DAPI and imaged. Five separate cases were conducted (cases 149, 155, 157, 160, and 161).

Longitudinal Tracking and Fluorescence Quantification

Organoid images were segmented with OrganoSeg2. The time course of images for each well were loaded into the tracking window in OrganoSeg2 and registered using the imregdemons algorithm in MATLAB, which was applied to the binary segmentation masks. For each pair of organoids across consecutive images, Euclidean distance is calculated, and the pairs with the smallest distance are matched, up to a user-defined maximum distance (50 pixels in this case). Organoid traces were manually inspected and excluded if they did not track the same organoid.

The images were then loaded into the fluorescence window in OrganoSeg2, and corresponding fluorescence images in blue, green, and red channels were loaded and assigned to the corresponding brightfield image. The brightfield segmentation mask was used to calculate a representative pixel intensity for each organoid (95th percentile for all cases except Case 155 DAPI, for which an 80th percentile was used since there were high levels of dead cellular debris).

Fluorescence Data Analysis

Fluorescence data for each organoid in each image was exported and normalized by subtracting the 20th percentile value on a per-image basis. For each case, a histogram of fluorescence data across all conditions was made (bin size = 5 for Nucview, bin size = 10 for DAPI), and a gaussian-mixture model was defined to match the frequency counts in the histogram. Gaussian-mixture fitting was done using the nlinfit function in MATLAB with a proportional error model. The 95th percentile of the lower distribution was used to define the case-wide cutoff for fluorescence positivity. Fluorescence positivity was then tracked by individual organoid, and once an organoid was observed to be positive, it was marked as positive for the rest of the images as the Nucview dye visibly faded over time. Organoids which were identified in less than half of the relevant images were removed from individual organoid traces.

End Matter

Author Contributions and Notes

C.J.W, N.L R.K.P, and K.A.J conceptualized software changes, C.J.W wrote software, N.L and K.A.J designed experiments, N.L performed experiments, C.J.W, N.L, and K.A.J analyzed data, C.J.W and K.A.J wrote paper.
The authors declare no conflict of interest.

Acknowledgments

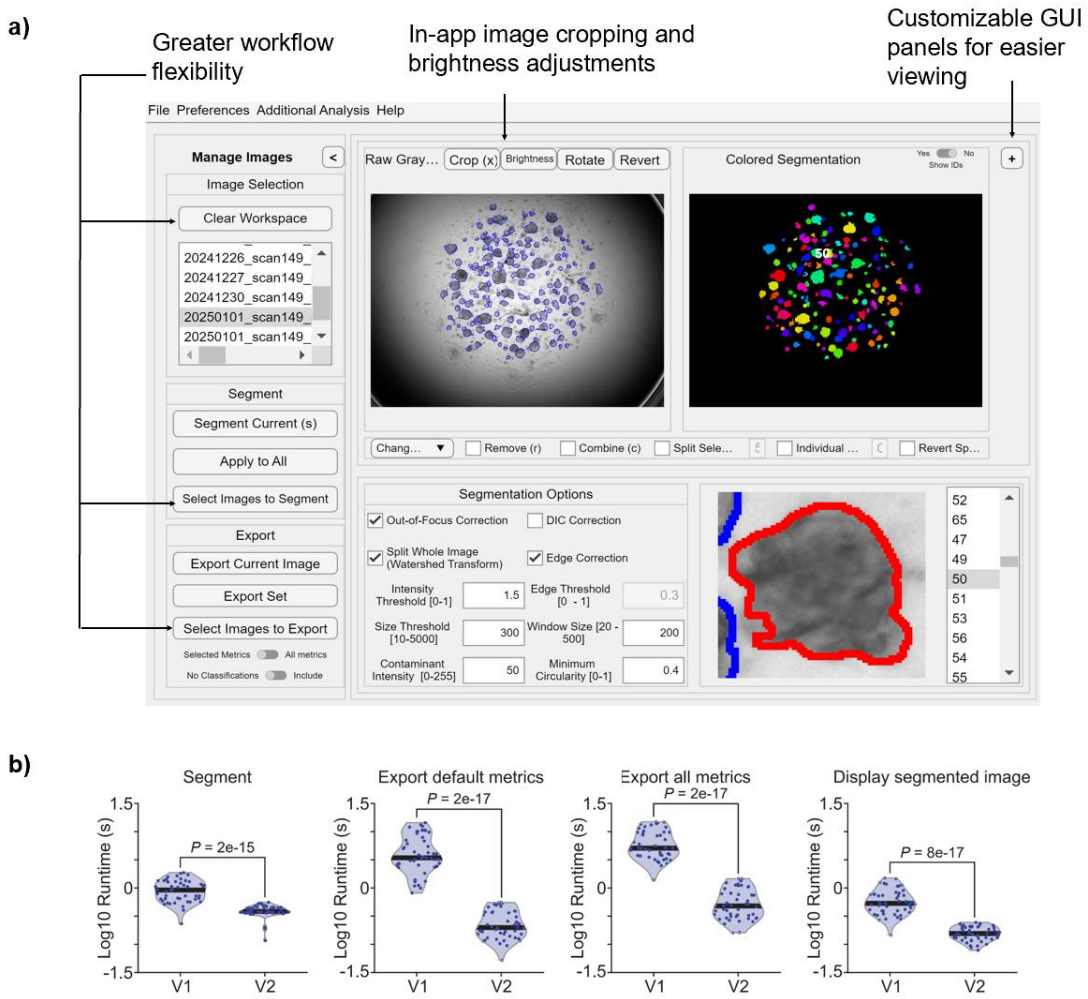
I would like to thank my advisors and all the members of the Janes Lab for their advice and support on the project. I would also like to thank Dr. Allen and the Capstone teaching team for their continuous feedback.

References

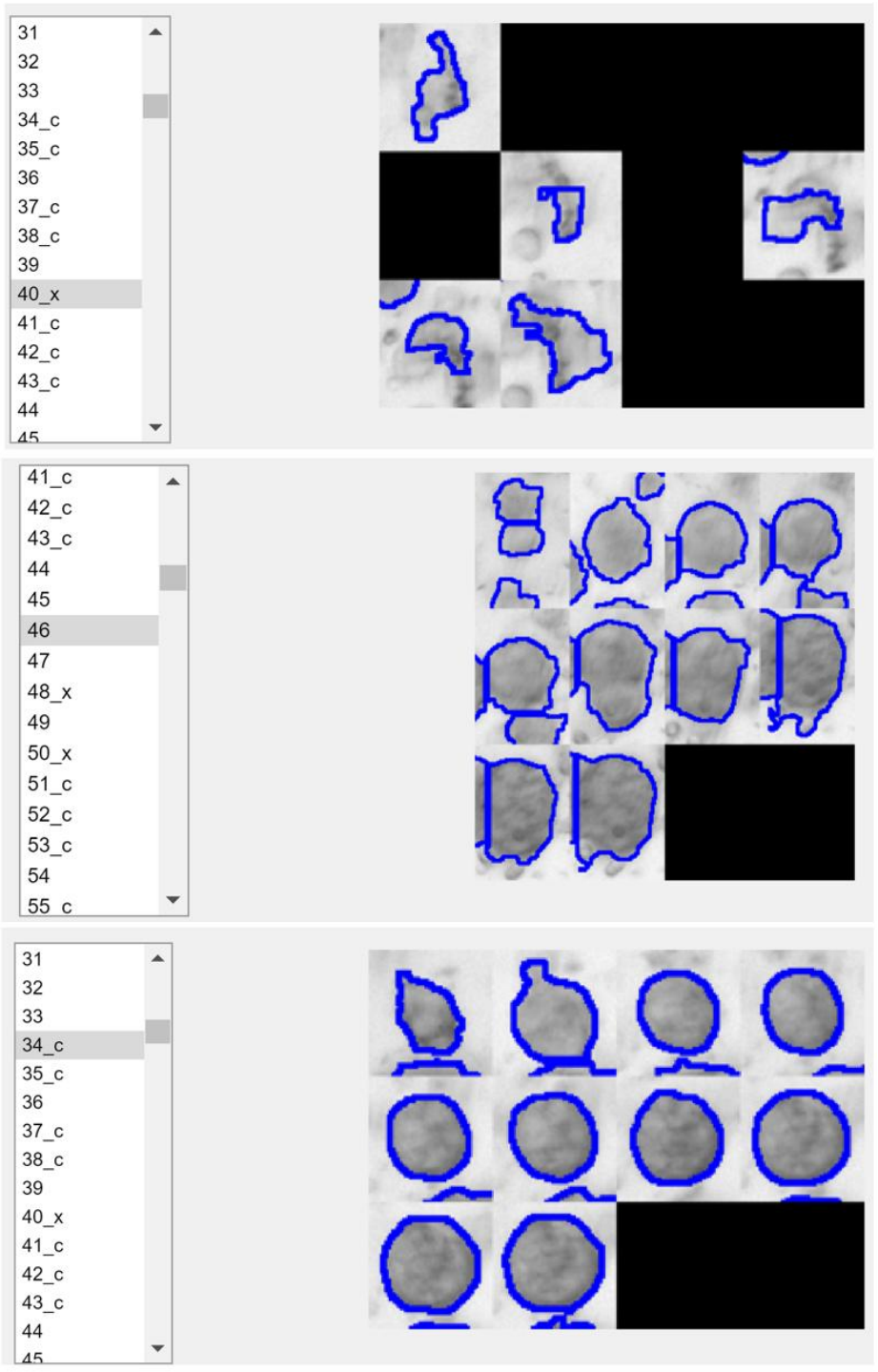
1. Hofer, M. & Lutolf, M. P. Engineering organoids. *Nat Rev Mater* **6**, 402–420 (2021).
2. Organoids | Nature Reviews Methods Primers. https://www.nature.com/articles/s43586-022-00174-y.
3. Pauli, C. *et al.* Personalized In Vitro and In Vivo Cancer Models to Guide Precision Medicine. *Cancer Discov* **7**, 462–477 (2017).
4. Kopper, O. *et al.* An organoid platform for ovarian cancer captures intra- and interpatient heterogeneity. *Nat Med* **25**, 838–849 (2019).
5. Yao, Y. *et al.* Patient-Derived Organoids Predict Chemoradiation Responses of Locally Advanced Rectal Cancer. *Cell Stem Cell* **26**, 17-26.e6 (2020).
6. Zhao, N. *et al.* Morphological screening of mesenchymal mammary tumor organoids to identify drugs that reverse epithelial-mesenchymal transition. *Nat Commun* **12**, 4262 (2021).
7. Cuyx, S. *et al.* Rectal Organoid Morphology Analysis (ROMA): A Diagnostic Assay in Cystic Fibrosis. *Journal of Visualized Experiments (JoVE)* e63818 (2022) doi:10.3791/63818.
8. Betge, J. *et al.* The drug-induced phenotypic landscape of colorectal cancer organoids. *Nat Commun* **13**, 3135 (2022).
9. CellProfiler: image analysis software for identifying and quantifying cell phenotypes | Genome Biology.

https://link.springer.com/article/10.1186/GB-2006-7-10-R100.
10. NIH Image to ImageJ: 25 years of image analysis | Nature Methods. https://www.nature.com/articles/nmeth.2089.
11. Borten, M. A., Bajikar, S. S., Sasaki, N., Clevers, H. & Janes, K. A. Automated brightfield morphometry of 3D organoid populations by OrganoSeg. *Sci Rep* **8**, 5319 (2018).
12. Morelli, M. *et al.* Metabolic-imaging of human glioblastoma live tumors: A new precision-medicine approach to predict tumor treatment response early. *Frontiers in Oncology* **12**, (2022).
13. Hashimi, M. *et al.* Antiviral responses in a Jamaican fruit bat intestinal organoid model of SARS-CoV-2 infection. *Nat Commun* **14**, 6882 (2023).
14. Gritti, N. *et al.* MOrgAna: accessible quantitative analysis of organoids with machine learning. *Development* **148**, dev199611 (2021).
15. Schröter, J. *et al.* A large and diverse brain organoid dataset of 1,400 cross-laboratory images of 64 trackable brain organoids. *Sci Data* **11**, 514 (2024).
16. Kahveci, B., Polatli, E., Bastanlar, Y. & Guven, S. OrganoLabeler: A Quick and Accurate Annotation Tool for Organoid Images. *ACS Omega* **9**, 46117–46128 (2024).
17. Keshara, R., Kim, Y. H. & Grapin-Botton, A. Organoid Imaging: Seeing Development and Function. *Annual Review of Cell and Developmental Biology* **38**, 447–466 (2022).
18. Matthews, J. M. *et al.* OrganoID: A versatile deep learning platform for tracking and analysis of single-organoid dynamics. *PLOS Computational Biology* **18**, e1010584 (2022).
19. Kassis, T., Hernandez-Gordillo, V., Langer, R. & Griffith, L. G. OrgaQuant: Human Intestinal Organoid Localization and Quantification Using Deep Convolutional Neural Networks. *Sci Rep* **9**, 12479 (2019).
20. Przanowska, R. K. *et al.* Patient-derived response estimates from zero-passage organoids of luminal breast cancer. *bioRxiv* 2024.03.24.586432 (2024) doi:10.1101/2024.03.24.586432.
21. Patel, P. & Jacobs, T. F. Tamoxifen. in *StatPearls* (StatPearls Publishing, Treasure Island (FL), 2025).
22. Petroni, G., Cantley, L. C., Santambrogio, L., Formenti, S. C. & Galluzzi, L. Radiotherapy as a tool to elicit clinically actionable signalling pathways in cancer. *Nat Rev Clin Oncol* **19**, 114–131 (2022).
23. Development of a deep learning based image processing tool for enhanced organoid analysis | Scientific Reports. https://www.nature.com/articles/s41598-023-46485-2.
24. Canny, J. A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **PAMI-8**, 679–698 (1986).
25. Marker-Controlled Watershed Segmentation. https://www.mathworks.com/help/images/marker-controlled-watershed-segmentation.html.

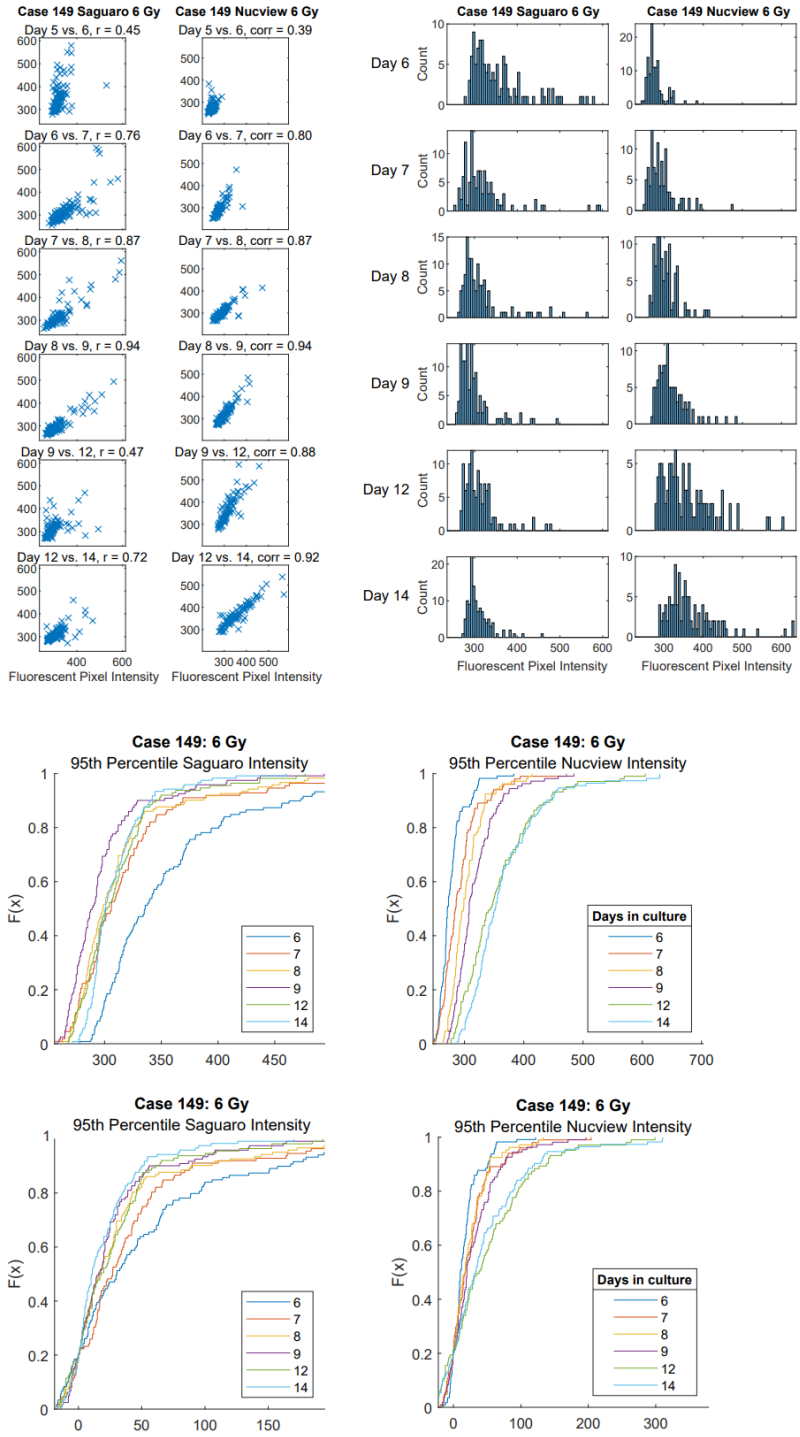
Supplementary Information



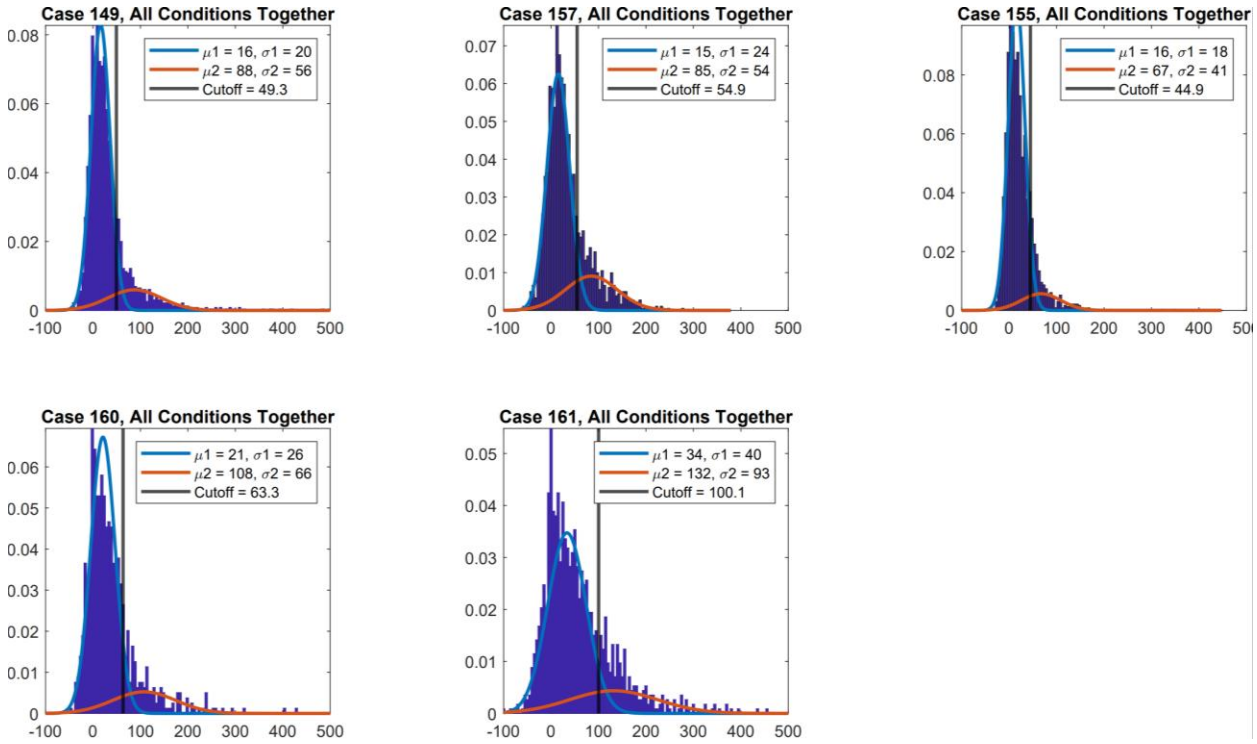
Supplementary Figure 1: OrganoSeg2 provides a more flexible and efficient GUI than OrganoSeg. a) OrganoSeg2 GUI with user-focused changes labeled. **b)** Runtimes for OrganoSeg(2) processes across 50 images



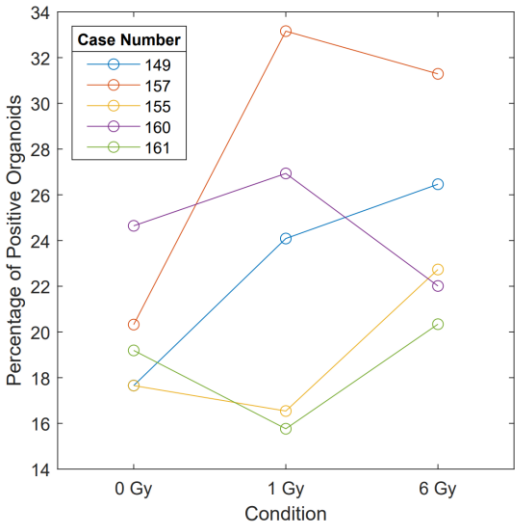
Supplementary Figure 2: Incorporation of manual verification into organoid tracking timeline. Examples of incorrect segmentation/tracking result, marked with "x" (top), accurate organoid tracking with faulty longitudinal morphological readout due to organoids joining (middle), and an almost perfectly correct segmentation, marked with "c" (bottom). Users are able to classify organoid traces according to these categories and export only desired categories. We used all traces except those marked with "x".



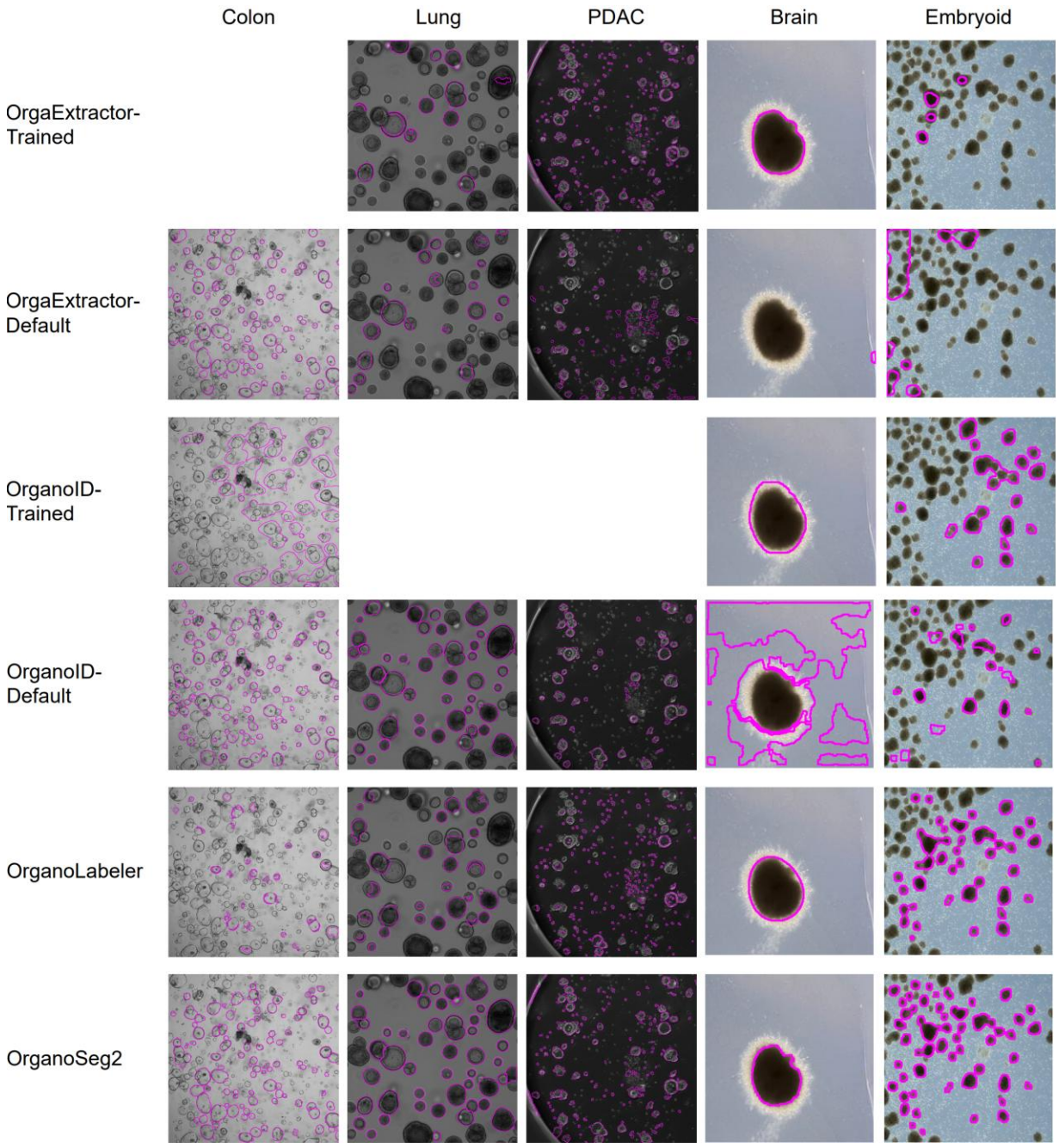
Supplementary Figure 3: Fluorescence analysis required post-processing to account for image to image variation. a) Correlation plots of individual organoid fluorescence between images. b) Histograms of fluorescence intensity across images for one case and condition. c) CDF plots before (top) and after (bottom) a per-image 20th percentile shift. **a-c)** Data is shown for both Saguaro Live-DEAD (left) and Nucview 488 Caspase-3 (right)



Supplementary Figure 4: Gaussian-mixture models fit to all cases despite a variety of underlying fluorescence distributions



Supplementary Figure 5: Population level cell death by case and condition after setting irreversible cutoff for each organoid



Supplementary Figure 6: Representative segmentations for all segmenting platforms and organoid types. Rows are different segmenters and columns are different organoid types. Blanks cells are for trained machine learning platforms using their internal dataset, since their default models were already trained on the dataset and do not require a trained entry.

	OrganoSeg2	OrganoLabeler	OrganoID - Default Model	OrganoID - Trained Model
Colon	I: 1 S: 200 EC: Yes, GW WS: 100 OOF: Yes MC: 0.5 CI: 120 IR: 10 RE: false CB: false	F: 1 B: 25 CL: 55 TG: 20 S: 100	Default	T: 0.9999
Lung	I: 1 S: 100 EC: Yes, GW WS: 300 OOF: Yes MC: 0.7 IR: 10 CB: false	F: 1.5 B: 75 CL: 100 TG: 50 S: 100	Default	NA
PDAC (Images 1,9,10)	I: 3 S: 300 EC: Yes, GO ET: 0.1 WS: 100 OOF: No MC: 0.4 CB: false II: true	F: -1 B: 25 CL: 150 TG: 30 S: 100	Default	NA
PDAC (Image 3)	I: 3 S: 300 EC: Yes, GO ET: 0.1 WS: 100 OOF: No MC: 0.4 CB: false RE: false II: true	F: -1 B: 35 CL: 500 TG: 5 S: 100	Default	NA

PDAC (Image 7)	I: 2 S: 500 EC: Yes, GO ET: 0.5 WS: 100 OOF: Yes MC: 0.4 CB: false II: true	F: -0.2 B: 25 CL: 100 TG: 10 S: 100	Default	NA
PDAC (Images 2,4,8)	I: 3 S: 500 EC: Yes, GO ET: 0.3 WS: 200 OOF: Yes MC: 0.4 WSS: 50 CB: false	F: 1.5 B: 35 CL: 100 TG: 10 S: 100	Default	NA
PDAC (Images 5, 6)	I: 3 S: 200 EC: Yes, GO ET: 0.15 WS: 200 OOF: Yes MC: 0.4 WSS: 50 CB: false	F: 2 B: 15 CL: 500 TG: 10 S: 100	Default	NA
Brain	I: 0.9 S: 3000 EC: Yes, GW WS: 100 OOF: Yes	F: 5 B: 95 CL: 55 TG: 1 S: 1	EL: 0.0005 EH: 0.01 S: 1000 ES: 2 T: 0.0005	T: 0.9999
EB	I: 1 S: 20 EC: No WS: 200 OOF: Yes IR: 1 CB: false	F: 0.1 B: 15 CL: 100 TG: 10 S: 10	EL: 0.4	T: 0.9999

Supplementary Table 1: Segmentation Parameters used in segmentation accuracy comparison