Published in final edited form as:

IEEE Trans Med Imaging. 2016 March ; 35(3): 819–829. doi:10.1109/TMI.2015.2497285.

Automatic Stem Cell Detection in Microscopic Whole Mouse Cryo-imaging

Patiwet Wuttisarnwattana,

Department of Computer Engineering, Chiang Mai University, Chiang Mai, Thailand, and
Biomedical Engineering Center, Chiang Mai University, Chiang Mai, Thailand

Madhusudhana Gargasha,

BioInVision, Inc., Mayfield Village, OH, USA

Wouter van't Hof,

Cell Processing Facility, Cleveland Cord Blood Center, Cleveland, OH, USA

Kenneth R. Cooke, and

Department of Pediatric Oncology, Johns Hopkins University, Baltimore, MD, USA

David L. Wilson

D.L. Wilson is with Department of Biomedical Engineering, Case Western Reserve University, Cleveland, OH, USA, Department of Radiology, University Hospitals of Cleveland, Cleveland, OH, USA and BioInVision, Inc., Mayfield Village, OH, USA

Patiwet Wuttisarnwattana: Patiwet@eng.cmu.ac.th; Madhusudhana Gargasha: madhu.gargasha@bioinvision.com; Wouter van't Hof: wvanthof@clevelandcordblood.org; Kenneth R. Cooke: kcooke5@jhmi.edu; David L. Wilson: dlw@case.edu

Abstract

With its single cell sensitivity over volumes as large as or larger than a mouse, cryo-imaging enables imaging of stem cell biodistribution, homing, engraftment, and molecular mechanisms. We developed and evaluated a highly automated software tool to detect fluorescently labeled stem cells within very large (~200GB) cryo-imaging datasets. Cell detection steps are: preprocess, remove immaterial regions, spatially filter to create features, identify candidate pixels, classify pixels using bagging decision trees, segment cell patches, and perform 3D labeling. There are options for analysis and visualization. To train the classifier, we created synthetic images by placing realistic digital cell models onto cryo-images of control mice devoid of cells. Very good cell detection results were (precision=98.49%, recall=99.97%) for synthetic cryo-images, (precision=97.81%, recall=97.71%) for manually evaluated, actual cryo-images, and <1% false positives in control mice. An α -multiplier applied to features allows one to correct for experimental variations in cell brightness due to labeling. On dim cells (37% of standard brightness), with correction, we improved recall (49.26%→99.36%) without a significant drop in precision (99.99%→99.75%). With tail vein injection, multipotent adult progenitor cells in a graft-versus-host-disease model in the first days post injection were predominantly found in lung, liver, spleen, and bone marrow. Distribution was not simply related to blood flow. The lung contained clusters of cells while other tissues contained single cells. Our methods provided stem cell

distribution anywhere in mouse with single cell sensitivity. Methods should provide a rational means of evaluating dosing, delivery methods, cell enhancements, and mechanisms for therapeutic cells.

Index Terms

biodistribution; cell; cell detection; cryo-imaging; fluorescent imaging; machine learning; optical imaging; segmentation; stem cell homing; image processing; visualization

I. INTRODUCTION

There are many pre-clinical and clinical studies of stem cell therapies across many different diseases and conditions, including treatment of graft-versus-host disease, autoimmune diseases, diabetes mellitus, multiple sclerosis, cardiac ischemia, osteoarthritis, cancers, and more [1–3]. In nearly all studies, knowledge of stem cell homing, biodistribution, longevity, optimal dosing strategies, and mechanism of action is lacking [4, 5]. Many imaging modalities, including MRI [6], PET [7], SPECT [8], bioluminescence (BLI) [9], and intravital imaging [10] have been used to image stem cells in animal models, but all have limitations as to sensitivity, quantification, and/or field of view.

Cryo-imaging is an imaging technology that enables cellular tracking throughout an entire mouse with single cell sensitivity. A cryo-imaging system (CryoViz™, BioInVision), consists of a fully automated system for repeated physical sectioning and tiled microscope imaging of a tissue block face, providing anatomical brightfield and molecular fluorescence, 3D microscopic imaging with single cell (5–20 μm diameter) resolution and sensitivity over large volumes [11–14]. Cryo-imaging has been used to investigate stem cell biology and regenerative medicine applications [13, 15, 16], among other applications such as phenotyping by spatial mapping of fluorescent reporter gene expression in small animals and embryos [17, 18], validation of other biomedical imaging modalities [19], tumor and metastasis characterization [20–22], etc. Our work here is focused on stem cell applications.

There are challenges to the development of automated processing for detecting stem cells in large volumes of cryo-image data. First, to detect single stem cells, high resolution (10 $\mu\text{m} \times 10 \mu\text{m} \times 40 \mu\text{m}$ slice thickness) cryo-imaging is needed, giving >200 GB of image data per mouse. Such an enormous data set cannot be loaded into memory on a contemporary workstation, requiring piecewise processing. Computation time is an issue limiting the types of algorithms which can be used. Second, autofluorescence from tissues, organs, and matter in the gastrointestinal tract can challenge automated cell detection algorithm, leading to potential false positive detections. Third, the brightness of cells can vary from one experiment to the next with more or less labeling. Fourth, in cryo-imaging, exceedingly bright cells can be visible in multiple image slices due to the presence of subsurface fluorescence, although this can be mitigated with processing [23–26]. Fifth, in a typical experiment, stem cells are really quite sparse, with 99.998% of a mouse occupied by other tissues. This requires that processing must have extraordinary specificity, simply because of the extraordinary number of other voxels present. Methods developed herein address these issues.

Since the detection of scattered stem cells in whole mouse cryo-imaging is a unique problem, there are only a few relevant image processing publications. There have been previous reports of computer algorithms for detecting and segmenting cells in low resolution microscopic imaging. For example, retinal nuclei detection using sombrero filter [27, 28], overlapping spots segmentation using weighted image matching [29], and low resolution object quantification using model-based segmentation [30]. However, these algorithms were specifically developed for detecting objects in relatively homogenous background, and the field of view was limited as compared to stem cell cryo-imaging data. Groups have employed a section-and-image approach to detect fluorescent microspheres in the heart vasculature of small animals to study blood flow [31–35]. Detection would have been easier because studies were limited to a single organ with low auto-fluorescence and because microspheres tend to be uniformly brighter than labeled cells. Steyer from our group [14] proposed a 3D algorithm to detect fluorescently-labeled stem cells from cryo-images based on hysteresis thresholding and color channel ratios. Although an accuracy of >98% was obtained in specific tissues for brightly labeled cells, we found the algorithm to be sensitive to noise and not applicable to the current whole mouse data with typically dimmer cells.

In this paper, we describe a novel machine learning-based approach for detecting fluorescently labeled stem cells in cryo-imaging data. Stem cells used in our study were multipotent adult progenitors cells (MAPCs) [5]. According to the literature, the cell shape is small and triangular and the size is about 10–20 μm . We carefully assess the robustness and accuracy of the algorithm using multiple approaches. We account for varying intensity of cells from one experiment to another without retraining of the classifier. We also demonstrate some of our visualization and analysis tools. The example application is stem cell therapy in a graft-versus-host disease mouse model.

II. Stem cell detection algorithm

Our stem cell detection algorithm takes as input a stack of 2D multi-band fluorescence cryo-images and creates an output, segmented 3D label volume with each fluorescent cell patch, typically consisting of a single cell, identified by a unique number. Representative 2D brightfield and fluorescence images from a whole mouse experiment with red fluorescence, quantum dot labeled stem cells are shown in Fig 1, respectively. One can see the red cell signals in many organs, including liver (Fig 1c). In this section, we describe the algorithm for processing multispectral fluorescent image data. Readers unfamiliar with cryo-imaging might want to first read about image acquisition in Experimental Methods.

Stem cell detection algorithm steps are: (1) Preprocess to create accurately tiled 3D images. (2) Eliminate regions that are immaterial to subsequent applications such as fur, embedding medium, and food remnants in the GI-tract. (3) Extract features using multiple spatial filters on each fluorescent channel. (4) Identify numerous candidate pixels for stem cell detection using simple feature rules. (5) Classify candidate pixels as either cell or background using a supervised classifier. (6) Segment cell patches from fluorescent volume. (7) Apply 3D-connected-component labeling to the stack of images to label cell patches. Results are labeled cell(s); optionally, cell counts per cluster; and spatial coordinates of detected cells. Detected cells are analyzed using a variety of 3D visualization and quantitative analyses,

including whole mouse 3D visualization, cell densities and cell spacing in volumes of interest, and cell counts within segmented organs or other tissue regions. Each step is described in more detail below. Pseudocode for stem cell detection is summarized in Fig 2.

A. Preprocessing

During the acquisition of 2D images on the imaging workstation, significant real-time image preprocessing occurs. Any non-uniform illumination pattern is compensated using a reference image of a white card for color and a fluorescence card for fluorescence. For tiled acquisition, individual tiles are registered and stitched together. A typical intensity-compensated, stitched 2D whole-mouse image is 9,500 by 3,400 pixels. These final images are then automatically aligned to each other along the z-axis to correct for either operator adjustment of field of view due to change in tissue size or minor misalignments due to any repositioning error of the digitally controlled stage. Because the system is mechanically accurate, alignments and registration is robustly done over a small search window. With this approach, we obtain tiled-images of a whole mouse without error [13]. 2D and 3D image data can be viewed in real time as images are acquired, allowing one to visualize anatomy and determine if one wants to pause sectioning to acquire optional histological tissue sections.

B. Elimination of immaterial regions

To reduce the possibility of false positive detections and reduce processing time, we digitally eliminate immaterial background regions and regions of high autofluorescence which could be misconstrued as stem cells and which are typically of no interest in stem cell experiments. The latter include food remnants in the gastro-intestinal tract (GI-tract), fur, lacrimal gland, ovaries, etc. We eliminate such regions and the external embedding medium by creating binary (0 and 1) masks where the background is comprised of these immaterial regions. Processing steps are illustrated in Fig 3. We remove the non-fluorescent OCT gel embedding medium by applying a threshold (T_{OCT}) to the green fluorescence channel (Fig 3a), creating a 2D binary image of the mouse foreground and OCT background (M_1 , Fig 3b). This process inadvertently sets some low fluorescence regions within the mouse to background. These latter regions are filled in using a morphological hole-filling algorithm (M_2 , Fig 3c) [36]. To remove fur and skin, we morphologically erode the result with a disk structuring element (SE), having a diameter equal to fur and skin layer thickness (R_{fur}) (M_3 , Fig 3d). Additionally, this operation removes any small, rare noise “islands” outside the mouse body as are obtained with debris. Next, food remnants inside the GI-tract can be very highly autofluorescence even when mice are on alfalfa-free chow. To detect and remove such regions, we employ a hysteresis thresholding technique [37] with empirically determined high and low thresholds, T_{high} and T_{low} , respectively (M_4 , Fig 3e). Threshold values are interactively set by the user at run-time in representative slices, and final results are relatively insensitive to selections. After binary image subtraction ($M_3 - M_4$), the final body mask is shown in Fig 3f. This 2D processing approach allows us to process each tiled-image in computer memory at full resolution.

C. Feature filters: sombrero “matched” filter and top-hat transform

To detect bright fluorescent cells, we used spatial filters “well-matched” to the bright spot cell feature. We used inverted Laplacian of Gaussian ($-LoG$), sometimes called a sombrero or Mexican hat filter. Sombrero filter parameters consist of δ_{filter} and size of the sombrero kernel (W_{kernel}). In preliminary experiments, parameters were adjusted to maximize filter output against the tissue background by examining cell signal profiles. As recommended [38, 39], the size of the sombrero kernel (W_{kernel}) was set to a support of $4w$ where w is the diameter of the central positive region of the sombrero. This minimizes distortion introduced by truncation. To avoid phase shift, we used a symmetric odd-sized kernel. This zero mean filter is designed to resemble the stem cell signal and provides a very high response to single and clustered stem cells while eliminating background signal. The formula describes the size of the sombrero kernel can be written as: $W_{kernel} = \text{round-to-the-nearest-odd-number}(4w)$ where $w = \sqrt{2}\sigma_{filter}$. We spatially filtered both the red and green fluorescent channels, giving a filtered red image $fR(x,y)$ (Fig 4c) and a filtered green image $fG(x,y)$ (Fig 4d).

In addition, we extracted features based on non-linear morphological operations. To each channel of fluorescent images, we applied a morphological gray-scale top-hat transform [40] using a suitable SE to obtain images $tR(x,y)$ (Fig 4e) and $tG(x,y)$ (Fig 4f). For an image I , the top-hat transform, tI , is defined as $tI = I - (I \circ SE)$ where SE is a 2D flat disk with radius R_{filter} . Importantly, the top-hat transform gives a minimal response to large step edges such as organ boundaries (Fig 4f) as compared to sombrero filtering alone (Fig 4d). This unique feature of the morphological operation greatly reduces false positive detections at edges. The values of δ_{filter} and R_{filter} were chosen so as to provide strong responses to stem cell signal with minimal noise response.

To compensate for changes in cell brightness due to variations in cell labeling from one experiment to the next, we introduced an α -multiplier to allow the classifier to detect new cells with different fluorophore concentration. The α -multiplier is simply a gain on the filter responses:

$$p^{(i)} = \alpha \times \begin{bmatrix} fR^{(i)} \\ fG^{(i)} \\ tR^{(i)} \\ tG^{(i)} \end{bmatrix} \quad (\text{Eq. 1})$$

where $p^{(i)}$ is a feature vector of pixel i and $fR^{(i)}$, $fG^{(i)}$, $tR^{(i)}$ and $tG^{(i)}$ are the corresponding filtered values of the pixel i . The α adjustment should be determined manually by dividing the cell intensity of the reference dataset (I_{ref}) by that of a new dataset (I_{new}). For example, if the new cell intensity decreases by half, α is 2. We used the adjustment below:

$$\alpha = \begin{cases} I_{ref}/I_{new} & I_{new} < I_{ref} \\ 1.0 & I_{new} \geq I_{ref} \end{cases} \quad (\text{Eq. 2})$$

Note that increasing α is similar to increasing exposure time. We analyze α in detail later.

D. Identification and classification of candidate pixels

Processing is done with consideration to the sparseness of cells. A volume of tiled-fluorescent images contains about 25 billion pixels as compared to 1 million voxel-sized cells used in a typical experiment. Therefore, we adopt a 2-pass technique by applying a fast processing method to identify candidate pixels before classifying them using a machine learning algorithm into “cell” or “background” class categories. This way, we greatly reduce computational time as compared to classifying each pixel.

Rules for determining candidate pixels are based on the following observations. (1) The red fluorescently labeled cell signal is highly responsive to the filters as discussed earlier. Only pixels with red-filtered values above thresholds (T_{fR} , T_{tR}) are considered candidates. (2) The red-labeled cells have a red filtered intensity greater than the green filtered intensity. Rules are codified below.

$$p_{candidate}^{(i)} \leftarrow \left\{ p^{(i)} \mid \left[\left(\alpha \times f R^{(i)} \right) > T_{fR} \right] \cap \left[\left(\alpha \times t R^{(i)} \right) > T_{tR} \right] \cap \left[t R^{(i)} > t G^{(i)} \right] \right\} \quad (\text{Eq. 3})$$

Parameters T_{fR} and T_{tR} are selected to “over-call” stem cells so as to create candidate group with few false negatives. We proposed two methods to estimate these parameters. First, we linked the parameters to noise in the data, e.g., $T_{fR} = (k) \times \text{Variance of } fR(x,y)$ and $T_{tR} = (k) \times \text{Variance of } tR(x,y)$, where k is a small positive number. Second, we manually adjusted the parameters using representative images and the corresponding detection result in an interactive visualization. One optimizes parameters until all the cell pixels are included in the candidate group (Suppl. Fig 1). We rely on subsequent processes to remove the false positive background pixels.

In the second step, we employed supervised machine learning classification to label the candidate pixels $p_{candidate}^{(i)}$ as either “cell” or “background” pixel. Each pixel had the four filtered values as features (Eq. 1). For classification, we used bagging decision trees [41]. Briefly, bagging decision trees classification is developed based on a bootstrap aggregating method where each decision tree is constructed from bootstrap replicas of the training data. To classify a pattern, each decision tree makes a vote on the pattern and the result is the majority of the votes. Principal advantages are ease of use with only a small number of easily tuned parameters, speed, and robustness to training noise. To select the optimal number of trees in the bagging decision tree classifier, we plotted the “out-of-bag” error [42] over the number of grown classification trees (Suppl. Fig 2). The “out-of-bag” error typically decreases with the number of trees and then flattens. As recommended, we chose this number to be the number of trees. For other parameters regarding bagging decision trees, we used the default parameters which came with Matlab^(TM) 2014b Statistics Toolbox (Mathworks, Inc.). More about classification training procedure is described later.

E. 2D segmentation of cell patches and 3D labeling

We next segment cells and clusters of cells using the detected pixels. Sometimes more than one pixel is labeled cell, especially when there is optical blurring or, less frequently, multiple cells are clumped together. A multiple pixel entity that belong to one cell or a cell cluster is

called a “cell patch”. Pixel detection algorithm in Step 4 may not detect all pixels that belong to a single cell patch (Fig 5a). This requires additional image processing. Steps are: (1) Morphologically dilate with a disk structuring element having radius $R_{segment}$, the binary image of detected pixels to completely cover all pixels of the cell patch (Fig 5b). (2) Apply top-hat transformation with a disk structuring element with radius R_{filter} to the red channel in order to remove autofluorescence background (Fig 5d), (3) Multiply results from steps (1) and (2) to obtain an image containing gray scale intensity values of cell pixels (Fig 5e).

Once all 2D images in the stack are processed using steps above, 3D-connected-component analysis (3D-CCA) [43] is employed to provide a unique label for each 3D cell patch. Since the whole mouse volume will not fit into a typical computer memory, we processed images in chunks, typically consisting of 200 tiled-images. Results from each chunk were accumulated appropriately. Numbers of cells connected at the interface between any two chunks was subtracted from the total count to eliminate double counting. Optionally, one can use next image processing technique [24] as part of the workflow to reduce effects of subsurface fluorescent in the event of very bright cells.

III. Algorithm training using synthetic images

For algorithm training and initial testing, we created synthetic images by adding realistic model cells to cryo-images from control mice devoid of fluorescently labeled cells. To match measured, sometimes irregular asymmetric real cell intensity profiles, we created a model cell patch by adding multiple Gaussians with small spatial offsets. The cell patch model was parameterized by the origin (spatial location) of the model cell in the fluorescent volume (x, y, z) , a Gaussian spread (δ_{cell}), an integrated cell patch intensity (I_{cell}), a random number of Gaussians per cell patch (n), and offset of each Gaussian $((dx_j, dy_j) | j = 1, \dots, n)$.

To create synthetic images, steps were: (1) Load representative red and green fluorescence images of a control mouse devoid of fluorescent cells into memory. (2) Mask immaterial tissues (OCT, fur, and food remnants) to avoid putting model cells into these regions. (3) Randomly pick N locations in the volume. These locations defined *origins* of the model cells $((x_i, y_i, z_i) | i = 1, \dots, N)$. (4) Add model cells to the red fluorescence images at the defined *origins*. To create model cells, steps were: (4.1) Generate a random number n_i from a Poisson distribution with small mean value λ , where n_i is the number of Gaussian per cell patch. (4.2) Generate small random offsets $((dx_j, dy_j) | j = 1, \dots, n_i)$ from a Gaussian random number generator having standard deviation δ_{offset} and zero mean. (4.3) Generate Gaussian intensity distribution $(G_j(x_i + dx_j, y_i + dy_j, z_i; \delta_{cell}) | j = 1, \dots, n_i)$ where δ_{cell} represents standard deviation, (x_i, y_i, z_i) represent origin of the model cell, and $((dx_j, dy_j) | j = 1, \dots, n_i)$ represent spatial offsets for the Gaussians. (4.4) Sum all n_i Gaussian intensity distributions and scale the result so that the intensity integrates to I_{cell} , and create a cell $(cell_i(x_i, y_i, z_i))$. (4.5) Repeat step (4.1)–(4.5) for $i = 1, \dots, N$. (5) To mimic a small amount of crosstalk of red quantum dot signal into the green channel, we multiplied the cell by β , the ratio of green to red intensity, and created green cells. (6) Write out color fluorescent synthetic images. Model parameters were chosen to best represent imaged cells qualitatively (Fig 6) and quantitatively (Suppl. Fig 3). Pseudocode and parameters are given in the supporting data file (Suppl. Fig 4 and Suppl. Table 1).

To train the algorithm, we created a synthetic image volume by adding model cells at random locations in images of a control mouse devoid of fluorescent stem cells. We chose 500 MB tiled images which included a range of tissues. So-called immaterial regions were excluded. We then applied sombrero filtering and top-hat transform to create four feature *values* (tR , tG , tR , tG) at each voxel. From known centers of model stem cells, we collected voxels positive for cells. We ran the candidate voxel processing step described above to identify candidates known to be negative for stem cells. Together, these represent the training data for the bagging decision trees (Matlab^(TM) 2014b Statistics Toolbox, Mathworks, Inc.) classification algorithm. The number of trees (*numTrees*) was empirically optimized to be in range 20–30 (Suppl. Fig 2). Other parameters were set to default values. Pseudocode and parameters regarding the classification training and cell detection algorithm are given in the supporting data file (Suppl. Fig 2 and Suppl. Table 2). Independent algorithm assessment will be discussed later.

IV. Image Analysis and Visualization Software

We developed an interactive user interface to provide qualitative and quantitative stem cell information. Quantitative measurements include total cell count in a whole mouse, cell counts in segmented organs, cell density (cells per tissue volume, where a volume could be a rectangular solid or more complicated), histogram of integrated cell patch intensity, cell patch size, and mean inter-cellular distance. We also developed interactive 3D visualization that enables one to zoom in to a particular organ of interest or zoom out for global biodistribution visualization. Automatic multiscale visualization enables alternate views between high and low resolution volume renderings depending on the zoom level and the region of interest within a mouse-sized volume. This feature overcomes memory limitations. When one renders a whole mouse on a screen with a limited number of display pixels, cell patches consisting of one or a few voxels will be lost. To remedy this, we adaptively dilate segmented cells based on zoom level such that almost all cells are visible across varying zoom levels.

V. Experimental Methods

A. Cryo-imaging

Cryo-imaging (CryoVizTM, BioInVision, Inc.) [11–16, 21, 24] was used to acquire image data for our experiments. CryoVizTM consists of a digital cryo-microtome (DigitomeProTM, BioInVision Inc.) housed within a –20C freezer chamber, a microscopic imaging system consisting of a high numerical aperture (NA) objective and a low-noise, cooled camera, a robotic XYZ positioner, and control system. CryoVizTM can section accurately between 2–200 μm section thickness and create tiled block face images with down to $\sim 5 \mu\text{m}$ pixel size. Fluorescence images for this study were acquired using a 510 nm long pass green fluorescent protein (long-pass GFP) filter (Exciter: HQ470/40x, Dichroic: Q495LP, Emitter: HQ500LP, Chroma, Rockingham, VT) followed by a RGB LCD filter. “Red” and “green” fluorescence images were obtained with these optical filters. Brightfield color images were also acquired with white light illumination. For whole mouse imaging, we sectioned at 40 μm with 10.5 μm pixels in plane, enabling single cell sensitivity. Micronscale resolution over

a large field-of-view was enabled by CryoViz™ Preprocessor (BioInVision, Inc.), which implements real-time 3D tiling and volume preview. Preprocessing runs on a high-end workstation with 128 GB RAM, 12-core Intel Xeon processor at 3.0 GHz, and a Windows 7 64-bit operating system.

B. Standard cell brightness

Cell brightness can vary from one experiment to the next due at least to variations in cell labeling. For simulations and analyses, we adopt a typical cell “standard.” This is a red quantum dot labeled cell with peak brightness of 32 on a 0–255 scale with an integrated intensity of 96. Empirically, integrated intensity was approximately 3 times the number of peak intensity. Standard imaging conditions are: exposure time of 2 s, 0.63x objective, 10x zoom optics, 10.5 μm pixel resolution, and using CryoViz™ light source and the filter sets for red quantum dot imaging. We relate the standard cell intensity to calibration microspheres (AlignFlow™ 515–660 nm). Under standard imaging conditions, the standard cell integrated brightness is roughly 17% of the calibration microspheres.

C. Animal protocol and preparation

The mouse protocol used in the experiment was approved by CWRU’s IACUC. Our exemplary stem cell experiment was based on a stem cell therapy with multipotent adult progenitor cells or MAPCs (MultiStem™, Athersys, Inc.) in a graft-versus-host disease (GVHD) mouse model. For this study, inbred B2D2F1 mice under allogeneic bone marrow transplantation (alloBMT) in a graft-versus-host disease (GVHD) condition were used. Briefly, the mice were lethally irradiated and then rescued using bone marrow from allogeneic donors (haplo-identical C57Bl/6). Additionally, purified splenic T-cells from the same donor were administered in order to induce GVHD. Approximately half a million MultiStem™ cells were labeled with red quantum dots (Qtracker® 625, Life Technologies). For the stem cell mice, the stem cells were then delivered via tail vein at 24 hours following alloBMT. For control mice, *unlabeled* MultiStem™ cells were injected. We allowed sufficient time (24 hours post injection) for the cells to circulate and home before sacrificing mice. For this paper, we report results on eight representative mice from a larger study on GVHD. Animals were anesthetized with isoflurane and euthanized by carbon dioxide. After sacrifice, animals were embedded in OCT medium (Tissue-Tek, Sakura Fintek USA Inc.) inside a custom freezing apparatus, snap frozen in liquid nitrogen, and mounted on to the CryoViz™ specimen stage for imaging.

D. Assessment of algorithm performances

To assess the stem cell detection algorithm performance, we conducted two experiments. In the first experiment, we used synthetic images consisting of background fluorescent images with model cells superimposed. We randomly selected 70 tiled-fluorescent images from whole mouse data (≈ 700 images totally) of a control mouse, having $9,414 \times 3,380$ pixels, each. Our dataset consisted of 35,069 randomly placed, model cells with averaged cell peak intensity of 30 on 255 grayscale intensity. The representative pixels from these model cells defined *cell pixels*. The pixels from the control images that passed *candidacy rules* (Eq. 3) defines *background pixels*. Data were partitioned into training and testing sets using 5-fold cross-validation. To assess classification, true positive (TP), false positive (FP), false

negative (FN), precision and recall [44] were recorded across the partitions and averaged. Second, we tested our algorithm against a smaller dataset created with *expert selection*. The dataset was derived from 10 images of a stem cell mouse and another 10 images from a control mouse. To define *cell pixels*, analysts identified 2,518 *cell pixels* in the images. If any observed cell signal consists of multiple pixels (cell patch), analysts may click at any pixel in the cell patch. The selected pixel was then adjusted automatically to the peak intensity pixel which represents the *cell pixel*. Analysts were told to discard any cell signals that appear as subsurface fluorescence [24]. *Background pixels* were the candidate pixels from the *control images* (10 images). Candidate pixel identification is described previously. The representative control images were selected by their visual similarity to the stem cell images. Without modification, the classifier trained on synthetic images was applied to these *expert-selected dataset*. Precision and recall were recorded.

In another experiment, we simply ran the detection software without modification on whole mouse data, stem cell mice ($n=4$) and control mice ($n=4$), and recorded detections. All of the detections in the control mice would be false positives, since there were no labeled cells present. By comparing to mice with fluorescent cells, we can estimate the percentage of FPs in term of false discovery rate (FDR). False discovery rate was calculated by dividing the number of false positive detections in a control mouse by the average number of positive detections in mice with stem cells. Implicitly, this assumes that the number of false positives in a control mouse and mouse with stem cells are equivalent. This is a reasonable assumption because both stem cell and control mice were prepared exactly the same. The only difference was that cells injected to the control mouse were not fluorescent. We acquired images and processed both in exactly the same manner.

E. Assessment of the effect of α -multiplier on detection of low contrast cells

We performed experiments to determine the ability to detect low contrast cells using the α -multiplier correction. To evaluate sensitivity of the software, we created synthetic images with model cells of different integrated intensity and then ran the cell detection software with and without gain adjustment (Eq. 2). We generated 430,000 synthetic cells with known levels of integrated intensity (I_{cell}), and placed them within brain tissue of a control mouse volume. Although brain tissue is quite homogenous, it tends to be have more false positives than other tissues of interest due to its autofluorescence. Test datasets were created by varying I_{cell} from 24 to 108, and *cell* and *background* pixels were collected. The classifier was initially trained to recognize a baseline cell intensity I_{cell} of 96. Cell detection software was then applied to all test datasets with $\alpha=1.0$ and classification results were assessed. For each new datasets with dimmer or brighter cells, α was adjusted prior to applying stem cell detection, as described previously (Eq. 2). We evaluated only recall as the metric on this experiment.

Additionally, we determined the optimal range of α -multiplier on a fixed dataset, as assessed using precision, recall and F_1 score [44]. Brain tissue synthetic images were generated with model cells having a fixed integrated intensity level. The classifier was repeatedly tested on the same dataset with α -multiplier varied from 0.1 to 6.0. Precision, recall, and F_1 scores were determined. The testing dataset consisted of equal number of *cell* and *background*

pixels, about 430,000 for each class. The classifier was trained to recognize model cells with intensity $I_{\text{cell}} = 96$ but tested with model cells with $I_{\text{cell}} = 36$.

VI. Results

The algorithm performed well on synthetic images containing model cells with typical cell intensity (peak 30 on 255 grayscale). Based on 35,069 model cells and 70 control tiled-images, the precision and recall were $98.49 \pm 0.07\%$ and $99.97 \pm 0.01\%$, respectively (Table 1). The performance were computed on 538 ± 24 FPs and 1 ± 2 FNs. The numbers represent mean $\pm$ standard deviation of the detection performance on 5 testing sets (5-fold cross validation). To extrapolate the detection performance to a mouse experiment, we assumed that (1) a typical mouse experiment receives 1 million labeled cells, and (2) cryo-imaging system produces 700 tiled-fluorescent images. By these assumptions, we estimated number of FPs and FNs to be 5,380 and 29 respectively. These yield estimated precision and recall in a whole mouse to be 99.46% and 99.97% respectively. Next, we used the classifiers to test against manually detected cells in another mouse. In this experiment, analysts picked 2,518 cells in 10 tiled-images containing real cell signals. The cells were used as *cell pixels* for the classification assessment. *Background pixels* were derived from 10 representative control images. The average cell peak intensity of the dataset was about 30 on 255 grayscale, but there was normal variation in cell intensity and patch size. Despite this additional variability, precision and recall on this new testing set were $97.81 \pm 0.18\%$ and $97.71 \pm 0.51\%$ respectively (Table 1). This includes 55 ± 5 FPs and 58 ± 13 FNs. By using previous extrapolation method, we estimated number of FPs and FNs in a whole mouse experiment to be 3,850 and 22,875 respectively. The results yield estimated precision and recall in a mouse experiment to be 99.61% and 97.71% respectively. In this experiment, we found that the analysts tended to pick subsurface fluorescent signals as the ground truth for algorithm assessment. This led to more FNs in the human detection experiment as compared to the number of FNs in the synthetic image experiment. However, by having a large number of TP in the calculation, the recalls on both experiments were very high ($>97\%$).

There were few false positive detections in control mice devoid of fluorescent stem cells. We detected $3,463 \pm 1,501$ false positives in the whole mouse data ($n=4$). As compared to 290,144 number of detection in the stem cell group (averaged, $n=4$), the false discovery rate (FDR) is about 1.2%. In tissues of most interest (liver, lung, spleen, bone marrow and kidneys), false positive detections were extremely low (Table 2). We found only 110 ± 71 FPs in liver, yielding FPR of $0.07 \pm 0.04\%$. Similar results were obtained in other tissues. Considering the whole mouse, the majority of FPs were found in other tissues with high autofluorescence. These included GI-tract inadequately masked and urinary bladder (49% of total FPs), lacrimal glands (16%), skin (13%), oral and nasal cavities (9%), ears (3%), female reproductive tract (2%) and others (8%). Since these tissues are anatomically identifiable, one can manually remove these tissues which would be immaterial in many stem cell experiments. Using a 3D segmentation and editing tool, one can do this cleaning in about 30 minutes. From our experience, this technique will greatly reduce the number of FPs, leaving only about 10% of the 3,463 identified above, giving an FPR of $\approx 0.1\%$ for the entire mouse.

We demonstrate stem cell visualization and analyses for a typical mouse. The 3D biodistribution is shown in Fig 7, which is also shown in a movie in the Suppl. Movie 1. Our software includes an interactive visualization tool which enables one to zoom within such a volume and examine the distribution of cells in great detail.

It was found that α -multiplier should allow one to process cryo-images from experiments with dimly labeled stem cells. Without α adjustment, we observed reduced recall as the cell intensity went down. However, when we adjusted α -multiplier as suggested in Eq. 2 and repeated the experiment, recall substantially improved (Fig 8). For instance, when we tested the algorithm with 63% reduced cell intensity ($I_{\text{cell}} 96 \rightarrow 36$), recall reduced from 99.99% to 53.72%. By proportionally setting α -multiplier to 2.7, the recall improved from 53.72% to 99.98%.

To determine the effect of α -multiplier on recall and precision, we varied the α -multiplier when processing dim cells at a fixed intensity. In this experiment, the cell brightness also reduced by 63% ($I_{\text{cell}} = 36$ as compared to the training set with $I_{\text{cell}} = 96$). The result show that the recall remained high for $\alpha > 1.5$, but rapidly dropped below 90% when $\alpha < 1.3$. With regards to precision, precision decreased with increasing α (Fig 8a). Precision remained high for $\alpha < 3.0$, but dropped below 90% for $\alpha = 3.5$, and then rapidly decreased to a minimum for $5.2 \leq \alpha \leq 6.0$ due to noise amplification by the adjustment. By considering without α -adjustment ($\alpha = 1.0$), recall and precision were 49.26% and 99.99%, respectively. With α -adjustment ($\alpha = 2.7$), we improved recall (49.26% $\rightarrow$ 99.36%) without a significant drop in precision (99.99% $\rightarrow$ 99.75%) (Fig 9a). F_1 score, which is derived from precision and recall, can be used to determine the optimal range of α -multiplier for stem cell detection (Fig 9b). The F_1 score reveals an optimal range ($F_1 > 95\%$) for $1.5 \leq \alpha \leq 3.5$.

Many cells in the lung were found in clusters consisting of a few cells or of a large number of cells (Fig 10a). A cell patch from the lung was large and probably represents over 5 cells based upon a probability density function (PDF) derived from the cell histogram of cell patch size (Fig 10d) as well as 3D histogram of cell patch size and the integrated intensity (Suppl. Fig 6a). Some cell clusters were much bigger than this. Cell patches from distant organs such as liver (Fig 10b) and spleen (Fig 10c) were much smaller. A PDF of cell size shows that cell patches in lung (Fig 10d) were much larger than those in other tissues such as liver (Fig 10e) and spleen (Fig 10f). We found that means of the patch size in lung, liver and spleen were 15.64 ± 18.01 voxels, 3.12 ± 4.87 voxels, and 3.26 ± 3.40 voxels respectively (mean $\pm$ standard deviation). We also found that medians of the patch size in lung, liver and spleen were 9, 2 and 2 voxels, respectively. Furthermore, by partitioning detected cells according cell size using a threshold (Fig 10d–f), we categorized detected cells into two groups, large cell size group (size > 5 voxels, green area in the figures) and small cell size group (size ≤ 5 voxels, blue area in the figures). The proportion of large cell clusters in lung was 64% whereas the proportion of large cell clusters in liver and spleen were 14% and 23%, respectively. This suggests that majority of red quantum dot signals in lung were from clusters of cells whereas the signals from other tissues were likely from isolated cells. Moreover, we performed 3D histogram using cell integrated intensity and patch size from lung, spleen and liver (Suppl. Fig 6). The results were consistent with the previous findings.

VII. Discussion

We developed robust, highly automated software for detecting and quantifying fluorescently labeled stem cells anywhere within a whole mouse from large, microscopic cryo-imaging datasets having single cell sensitivity. This is important because in many stem cell and regenerative medicine studies, surprisingly little is known about mechanisms. An initial step in understanding mechanisms is the determination of cell biodistribution, homing, fate, and engraftment. Our methods can provide this information with heretofore unavailable sensitivity and accuracy. In experiments reported here, cells were labeled with red quantum dots (Qtracker[®] 625, Life Technologies). However, in other experiments, we have labeled cells with fluorescent dyes and fluorescent proteins including GFP [22], Cy5 [45], CFSE [46], as well as RFP and Wasabi (unpublished). The downside of labeling with an exogenous fluorophore (quantum dots or dye) is that the signal can dissipate over time or possibly show up in macrophages as stem cells are scavenged. For longer term studies of engraftment, we recommend the use of fluorescent proteins where cells can stably create fluorescent proteins *in vivo* for weeks and even months. Potentially, one can improve single cell detection by using near-infrared fluorophores such as iRFP [47] if it is sufficiently bright. The advantage is that the near-infrared spectral region has minimal autofluorescence to confound detection.

There are some significant innovations in this report. For the first time, we report software which can robustly detect and quantify single stem cells throughout a mouse. Processing time for our large, >200 GB data sets is about 3 hours, three times faster than the much less robust, semi-automatic algorithm previously reported [15]. For comparison, we estimate that a fully manual analysis would require >6 months of extraordinary effort. Very many stem cells studies have examined cell distributions using histology. Our proposed methodology can replace such exceedingly laborious efforts. For example, one can cryo-image and analyze an entire excised liver in a few hours instead of probably weeks with histology. This is the first time that a machine learning algorithm has been used with cryo-image data. Although relatively simple features were used, results are excellent. Our use of synthetic images for training and assessment is somewhat novel, and appropriate, as very similar results were obtained against manually analyzed cryo-images. Determination of the number of false positives in control mice was shown to be an easy, reliable assessment method. In addition to stem cell quantification in tissues, we have found visualization to be equally important and desired by our stem cell biology collaborators. We developed interactive 3D visualization that enables one to examine biodistribution in a mouse and then zoom to particular regions to examine densities or to zoom even further and see individual cells. All of this can be done within the context of volume rendered color brightfield data giving full anatomical context using specialized rendering reported earlier [12]. We encourage readers to examine movies in supplemental material. To enable multi-scale visualization of large data sets (>200GB), much larger than computer or video card memory, we developed specialized methods described earlier.

Assessment results showed that the algorithm was accurate and robust across the entire mouse. We assessed the efficacy of the algorithm and classification using cross validation on synthetic images with known model cells on cryo-images and with cells detected by expert analysts. Precision and recall from these experiments were >97%. When we analyzed false

positive detection in control mice devoid of fluorescent stem cells, there were remarkably few false positives. On average, 3,500 false positives ($n=4$) were found in whole mice. Of these 90% were found in some few highly autofluorescent, easily recognizable tissues (e.g., GI tract and lacrimal glands), which can be effectively masked out with about half an hour of manual editing, if desired. This leaves only about 350 false positives in those organs of most interest in stem cell therapies, e.g., brain, liver, spleen, etc. On average, in the liver only about 0.07% of detections were deemed false positives.

We can also compare numbers of cells detected to those injected. In experiments, we injected about 500,000 labeled cells and detected about 290,000 fluorescent clusters at 24 hours post injection. Since there are a large number of multi-cell clusters in the lung, we added a correction of $N_L(C_c-1)$, where N_L is the number of clusters in the lung and C_c is the approximate average number of cells per cluster. This gives about 400,000 cells ($N_L \approx 53,000$, $C_c \approx 3$), still leaving a goodly number of cells left unaccounted. Potential explanations include incomplete injections; cells of varying intensity at the time of injection; some of which might not be detected; missing dim cells in thick tissue slices ($40 \mu\text{m}$); cell death in the 24 hours between injection and imaging; and probably more. Future experiments should help determine the size of these potential effects. Despite the “missing” cells, high precision and recall indicate that it is possible to make statements about absolute cell counts and densities in specific tissues. Relative counts between tissues are probably even more stable.

In our experience, cell brightness can vary from one experiment to the next due to variations in cell labeling and possibly other experimental variables. Since it would be quite disadvantageous to train a classifier for each new cell population, we developed and tested the α -multiplier approach where a single classifier model was used and filter features from a new experiment were multiplied by α prior to classification. This mimics what would happen with a longer exposure time or with brighter cells. As α increases, recall is improved as cells are detected, but at some point precision decreases due to an increase in number of false positives from autofluorescent noise. For the case of very dim cells, 17% of the standard integrated brightness, we found that there is a wide sweet spot for α between 1.5 and 3.5 where precision and recall were both 1.0. In the text, we offer a scheme for determining appropriate α values.

Since there are occasional batches of cells too weakly labeled for accurate detection, we need a quick method for labeling assessment. Simulations and experimental experience suggest that average integrated cell brightness under standard imaging conditions should be at least 0.4 times that of the standard cell or 13 on 255 grayscale peak intensity. This might seem conservative based upon simulations in Fig 7. However, unlike the simulations, actual cell brightness follows a distribution, and we must capture cells less bright than the average. Given a new batch of cells, we freeze some labeled cells on a glass slide and image the slide using cryo-imaging. Imaging the cells under standard conditions allows us to compare brightness to the standard cell. Experimentally, we have found results to be predictive of those in vivo.

Histograms of cell patch size and integrated intensity reveals that signals in the lung were typically from clusters of cells while those from other tissues were from single cells (Fig 10

and Suppl. Fig. 6). Owing to the systemic tail vein injection, a majority of cells were trapped within the lung in this short-term study (Suppl. Fig 7). In the lung, clusters consisted of a few cells to probably as many as 3–5 cells (Fig 10a). In addition, high quantum dot signals were detected at the tail injection site probably indicating some cell leakage from the vein into tissue. We also speculate that stem cells might respond to the injury so that they tend to stay at the injured site. Following filtration in the lung, cells at distant organs were almost all single isolated cells. Regardless of high cardiac output to brain, muscle and kidneys, we found that there was low stem cell density in these tissues. We typically found only 50 cells out of 500,000 injected in the brain, indicating that the blood brain barrier was relatively impervious to cells in our experiments. We conclude that the stem cell bio-distribution is not simply related to blood flow and that cells preferentially deposit in specific organs.

In conclusion, our new software for stem cell detection allows us to accurately detect and quantify cells anywhere in the entire whole mouse volume with single cell sensitivity. This work is significant as it provides answers to the pervasive question in stem cell regenerative medicine, “Where did the cells go?” We believe that the cryo-imaging together with the proposed software will provide a rational means of evaluating dosing, delivery methods, cell enhancements, and mechanisms for therapeutic cells. Hopefully, it will have a significant impact on the advancement of stem cell therapy and regenerative medicine.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The research was supported by an Ohio Third Frontier, WPP award (D.L.W., co-I), National Center of Regenerative Medicine Pilot Grant (K.R.C.), the Ohio Board of Regents (K.R.C.), the Meredith Cowden Foundation (K.R.C.), Ministry of Science and Technology of Thailand Scholarship (P.W.), Chiang Mai University Fellowship Program (P.W.), and the National Institute of Health through R42-CA124270 & R41HD063241-01 (D.L.W.).

References

1. Barry FP, Murphy JM. Mesenchymal stem cells: clinical applications and biological characterization. *Int J Biochem Cell Biol.* Apr.2004 36:568–84. [PubMed: 15010324]
2. Salem HK, Thiemermann C. Mesenchymal stromal cells: current understanding and clinical status. *Stem Cells.* Mar 31.2010 28:585–96. [PubMed: 19967788]
3. Uccelli A, Moretta L, Pistoia V. Mesenchymal stem cells in health and disease. *Nat Rev Immunol.* Sep.2008 8:726–36. [PubMed: 19172693]
4. Sudres M, Norol F, Trenado A, Gregoire S, Charlotte F, Levacher B, et al. Bone marrow mesenchymal stem cells suppress lymphocyte proliferation in vitro but fail to prevent graft-versus-host disease in mice. *J Immunol.* Jun 15.2006 176:7761–7. [PubMed: 16751424]
5. Auletta JJ, Cooke KR, Solchaga LA, Deans RJ, van't Hof W. Regenerative stromal cell therapy in allogeneic hematopoietic stem cell transplantation: current impact and future directions. *Biol Blood Marrow Transplant.* Jul.2010 16:891–906. [PubMed: 20018250]
6. Walczak P, Kedziorek DA, Gilad AA, Barnett BP, Bulte JW. Applicability and limitations of MR tracking of neural stem cells with asymmetric cell division and rapid turnover: the case of the shiverer dysmyelinated mouse brain. *Magn Reson Med.* Aug.2007 58:261–9. [PubMed: 17654572]
7. Wolfs E, Struys T, Notelaers T, Roberts SJ, Sohni A, Bormans G, et al. 18F-FDG labeling of mesenchymal stem cells and multipotent adult progenitor cells for PET imaging: effects on ultrastructure and differentiation capacity. *J Nucl Med.* Mar.2013 54:447–54. [PubMed: 23353687]

8. Chin BB, Nakamoto Y, Bulte JW, Pittenger MF, Wahl R, Kraitchman DL. ¹¹¹In oxine labelled mesenchymal stem cell SPECT after intravenous administration in myocardial infarction. *Nucl Med Commun.* Nov.2003 24:1149–54. [PubMed: 14569169]
9. Wang H, Cao F, De A, Cao Y, Contag C, Gambhir SS, et al. Trafficking mesenchymal stem cell engraftment and differentiation in tumor-bearing mice by bioluminescence imaging. *Stem Cells.* Jul. 2009 27:1548–58. [PubMed: 19544460]
10. Myers JT, Barkauskas DS, Huang AY. Dynamic Imaging of Marrow-Resident Granulocytes Interacting with Human Mesenchymal Stem Cells upon Systemic Lipopolysaccharide Challenge. *Stem Cells Int.* 2013; 2013:656839. [PubMed: 23606861]
11. Gargsha M, Qutaish M, Roy D, Steyer G, Bartsch H, Wilson DL. Enhanced Volume Rendering Techniques for High-Resolution Color Cryo-Imaging Data. *Proc Soc Photo Opt Instrum Eng.* 2009; 7262:72655V.
12. Gargsha M, Qutaish MQ, Roy D, Steyer GJ, Watanabe M, Wilson DL. Visualization of color anatomy and molecular fluorescence in whole-mouse cryo-imaging. *Comput Med Imaging Graph.* Apr.2011 35:195–205. [PubMed: 21041065]
13. Roy D, Steyer GJ, Gargsha M, Stone ME, Wilson DL. 3D cryo-imaging: a very high-resolution view of the whole mouse. *Anat Rec (Hoboken).* Mar.2009 292:342–51. [PubMed: 19248166]
14. Steyer GJ, Roy D, Salvado O, Stone ME, Wilson DL. Cryo-Imaging of Fluorescently-Labeled Single Cells in a Mouse. *Proc Soc Photo Opt Instrum Eng.* Jan 1.2009 7262:72620W–72620W8.
15. Steyer GJ, Dong F, Kanodia L, Roy D, Penn M, Wilson DL. Detection and quantification of fluorescent cell clusters in cryo-imaging. *Int J Biomed Imaging.* 2012; 2012:698413. [PubMed: 22481905]
16. Wilson D, Roy D, Steyer G, Gargsha M, Stone M, McKinley E. Whole Mouse Cryo-Imaging. *Proc Soc Photo Opt Instrum Eng.* Jan 1.2008 6916:69161I–69161I9.
17. Powell KA, Wilson D. 3-dimensional imaging modalities for phenotyping genetically engineered mice. *Vet Pathol.* Jan.2012 49:106–15. [PubMed: 22146851]
18. Roy D, Gargsha M, Steyer GJ, Hakimi P, Hanson RW, Wilson DL. Multi-scale characterization of the PEPCK-C mouse through 3D cryo-imaging. *Int J Biomed Imaging.* 2010; 2010:105984. [PubMed: 20467563]
19. Faramarzalain A, Prabhu D, Abdul-Aziz A, Wang W, Chamie D, Yamamoto H, et al. Ex vivo cryoimaging for plaque characterization. *JACC Cardiovasc Imaging.* Apr.2014 7:430–2. [PubMed: 24742895]
20. Burden-Gulley SM, Qutaish MQ, Sullivant KE, Lu H, Wang J, Craig SE, et al. Novel cryo-imaging of the glioma tumor microenvironment reveals migration and dispersal pathways in vivid three-dimensional detail. *Cancer Res.* Sep 1.2011 71:5932–40. [PubMed: 21862632]
21. Qutaish MQ, Sullivant KE, Burden-Gulley SM, Lu H, Roy D, Wang J, et al. Cryo-image analysis of tumor cell migration, invasion, and dispersal in a mouse xenograft model of human glioblastoma multiforme. *Mol Imaging Biol.* Oct.2012 14:572–83. [PubMed: 22125093]
22. Burden-Gulley SM, Qutaish MQ, Sullivant KE, Tan M, Craig SE, Basilion JP, et al. Single cell molecular recognition of migrating and invading tumor cells using a targeted fluorescent probe to receptor PTPmu. *Int J Cancer.* Apr 1.2013 132:1624–32. [PubMed: 22987116]
23. Krishnamurthi G, Wang CY, Steyer G, Wilson DL. Removal of subsurface fluorescence in cryo-imaging using deconvolution. *Opt Express.* Oct 11.2010 18:22324–38. [PubMed: 20941133]
24. Steyer GJ, Roy D, Salvado O, Stone ME, Wilson DL. Removal of out-of-plane fluorescence for single cell visualization and quantification in cryo-imaging. *Ann Biomed Eng.* Aug.2009 37:1613–28. [PubMed: 19513848]
25. Kim B, Naemura T. Blind Depth-variant Deconvolution of 3D Data in Wide-field Fluorescence Microscopy. *Sci Rep.* 2015; 5:9894. [PubMed: 25950821]
26. Sarder P, Nehorai A. Deconvolution methods for 3-D fluorescence microscopy images. *Ieee Signal Processing Magazine.* May.2006 23:32–45.
27. Byun JY, Verardo MR, Sumengen B, Lewis GP, Manjunath BS, Fisher SK. Automated tool for the detection of cell nuclei in digital microscopic images: Application to retinal images. *Molecular Vision.* Aug 16.2006 12:949–960. [PubMed: 16943767]

28. Usaj, M.; Torkar, D.; Miklavcic, D. Automatic cell detection in phase-contrast images for evaluation of electroporation efficiency in vitro. 11th Mediterranean Conference on Medical and Biological Engineering and Computing; 2007; 2007. p. 851-855.
29. Noordmans HJ, Smeulders AWM. Detection and characterization of isolated and overlapping spots. *Computer Vision and Image Understanding*. Apr.1998 70:23–35.
30. Bornfleth H, Satzler K, Eils R, Cremer C. High-precision distance measurements and volume-conserving segmentation of objects near and below the resolution limit in three-dimensional confocal fluorescence microscopy. *Journal of Microscopy-Oxford*. Feb.1998 189:118–136.
31. Krueger MA, Huke SS, Glenn RW. Visualizing Regional Myocardial Blood Flow in the Mouse. *Circulation Research*. Apr 26.2013 112:E88–U24. [PubMed: 23513055]
32. Bernard SL, Ewen JR, Barlow CH, Kelly JJ, McKinney S, Frazer DA, et al. High spatial resolution measurements of organ blood flow in small laboratory animals. *Am J Physiol Heart Circ Physiol*. Nov.2000 279:H2043–52. [PubMed: 11045936]
33. Hakimzadeh N, van Horsen P, van Lier MG, van den Wijngaard JP, Belterman C, Coronel R, et al. Detection and quantification methods of monocyte homing in coronary vasculature with an imaging cryomicrotome. *J Mol Cell Cardiol*. Nov.2014 76:196–204. [PubMed: 25179912]
34. van Horsen P, Siebes M, Hoefer I, Spaan JA, van den Wijngaard JP. Improved detection of fluorescently labeled microspheres and vessel architecture with an imaging cryomicrotome. *Med Biol Eng Comput*. Aug.2010 48:735–44. [PubMed: 20574721]
35. Spaan JA, ter Wee R, van Teeffelen JW, Streekstra G, Siebes M, Kolyva C, et al. Visualisation of intramural coronary vasculature by an imaging cryomicrotome suggests compartmentalisation of myocardial perfusion areas. *Med Biol Eng Comput*. Jul.2005 43:431–5. [PubMed: 16255423]
36. Hasan MM, Mishra PK. Improving morphology operation for 2D hole filling algorithm. *International Journal of Image Processing (IJIP)*. 2012; 6:635–646.
37. Canny J. A Computational Approach to Edge-Detection. *Ieee Transactions on Pattern Analysis and Machine Intelligence*. Nov.1986 8:679–698. [PubMed: 21869365]
38. Zhao, FW.; deSilva, CJS. Use of the Laplacian of Gaussian operator in prostate ultrasound image processing. *Proceedings of the 20th Annual International Conference of the Ieee Engineering in Medicine and Biology Society*; 1998. p. 812-815.
39. Sotak GE, Boyer KL. The Laplacian-of-Gaussian Kernel - a Formal Analysis and Design Procedure for Fast, Accurate Convolution and Full-Frame Output. *Computer Vision Graphics and Image Processing*. Nov.1989 48:147–189.
40. Jackway P. Improved morphological top-hat. *Electronics Letters*. 2000; 36:1194–1195.
41. Breiman L. Bagging predictors. *Machine Learning*. Aug.1996 24:123–140.
42. Bylander T. Estimating generalization error on two-class datasets using out-of-bag estimates. *Machine Learning*. Jul-Sep;2002 48:287–297.
43. Borgfors G, Nystrom I, diBaja GS. Connected components in 3D neighbourhoods. *Scia '97 - Proceedings of the 10th Scandinavian Conference on Image Analysis*. 1997; 1 and 2:567–572.
44. Goutte C, Gaussier E. A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. *Advances in Information Retrieval*. 2005; 3408:345–359.
45. Zhou Z, Qutaish M, Han Z, Schur RM, Liu Y, Wilson DL, et al. MRI detection of breast cancer micrometastases with a fibronectin-targeting contrast agent. *Nat Commun*. 2015; 6:7984. [PubMed: 26264658]
46. Wuttisarnwattana P, Raza SA, Eid S, Cooke KR, Wilson DL. Novel T lymphocyte proliferation assessment using whole mouse cryo-imaging. 2014 pp. 90381U-90381U-9.
47. Filonov GS, Piatkevich KD, Ting LM, Zhang J, Kim K, Verkhusha VV. Bright and stable near infra-red fluorescent protein for in vivo imaging. *Nature biotechnology*. Jul 17.2011 29:757–761.

Fig. 1. Cryo-imaging enables detection of red quantum dot labeled stem cells anywhere within a whole mouse with single cell sensitivity. It affords brightfield contrast showing anatomical details (a) as well as co-registered molecular fluorescence (b). Upon zooming into a region of the fluorescence image (yellow box in (b)), one can see individual cells in liver (c).

```
Preprocess untiled fluorescent images after cryo-imaging acquisition
For each tiled fluorescent cryo-image
    Eliminate immaterial regions such as OCT, skin, fur and GI-tract
    Filter image with sombrero filter and top-hat transform
    Identify candidate pixels using thresholding rules
    Classify candidate pixels using the supervised classifier (bagging decision trees)
    Segment cell patches
End
Label each detected cell patch in the 3D volume using chunk processing
```

Fig. 2.
Stem cell detection pseudocode

Fig. 3.

Algorithm for eliminating immaterial regions. On the input fluorescent image (a), following a threshold for removing low intensity regions (b), we apply hole-filling (c) to restore regions that were unintentionally removed. We remove skin and fur using a morphological erode (d), and highly autofluorescent regions using hysteresis thresholding (e). The final binary mask after removal of immaterial regions is shown in (f).

Fig. 4.

Feature extraction using image filtering. We apply a sombrero filter the red (a) and green (b) channels of input image to obtain corresponding filtered images (c,d). Top-hat transformation results in output images (e,f) with organ edges suppressed. There are bright cell signals in the lung which are highly responsive to the filter (a, c, e).

Fig. 5.

Cell patch segmentation. The binary image output by the classifier (a) is dilated to “cover” side lobes and identify the full spatial extent of cluster (b). This dilated image is used to mask the top hat transformed red channel (d), yielding the desired result (e) containing gray scale intensities of segmented clusters.

Fig. 6.

Model cell signals were created to mimic real stem cell signals using the proposed algorithm. Model cells were created randomly by adding Gaussians with different offsets. We chose example model cells (c, f, i) that mimic actual cell signals: dim cell (a, b), bright cell (d, e), and cell cluster (g, h).

Fig. 7. 3D visualization of stem cell bio-distribution in a whole mouse. Detected stem cells were dilated and rendered using surface rendering (yellow beads). Anatomical brightfield data are also shown. With tail vein injection, stem cells in a graft-versus-host-disease model were predominantly found in lung, liver, spleen, and bone marrow.

Fig. 8.

Performance improvement of low contrast cells with α -multiplier. The α -multiplier adjustment results in nearly perfect recall even with data sets containing stem cells of much lower intensity as compared to the ones used for training. In this experiment, the classifier was trained to detect cell with $I_{\text{ref}} = 96$, then it was tested with synthetic dataset with $I_{\text{new}} = 24, 36, 48, 60, 96, 108$ with and without alpha adjustment.

Fig. 9.

Detection performance as a function of alpha multiplier. Precision decreases with increasing alpha ($\alpha > 3.0$) due to amplified noise (a, blue dot line). Inversely, recall increases with increasing alpha (a, red dot line). In this experiment, F₁ score reveals that optimal range ($F_1 > 95\%$) is $1.3 < \alpha < 3.5$ (b).

Fig. 10.

The lung contained clusters of cells while other tissues contained single cells. Due to tail vein injection, detected cells were single cells in most of the tissues except those in lung. In most case, detected cells in lung (a) were bigger than the detected cells in other tissues such as liver (b) and spleen (c). Probability density function (PDF) of cell size shows that the detected cells in lung (d) were larger than those in other tissues such as liver (e) and spleen (f). Green area in the PDFs represent proportion of detected cells with large cell patch size (>5).

TABLE I

Algorithm Assessment Results

Detection Performance	Synthetic Images	Human Detected Images
Precision	98.49 $\pm$ 0.07%	97.81 $\pm$ 0.18%
Recall	99.97 $\pm$ 0.01%	97.71 $\pm$ 0.51%

TABLE II

Number of Detection in the Organs of Interest

Organs of interest	Number of detection in stem cell mice (n=4)	Number of detection in control mice (n=4)	False discovery rate
Liver	164,056 $\pm$ 60,767	110 $\pm$ 71	0.07 $\pm$ 0.04%
Lung	52,992 $\pm$ 8,625	43 $\pm$ 30	0.08 $\pm$ 0.06%
Spleen	5,591 $\pm$ 3,185	1 $\pm$ 2	0.02 $\pm$ 0.03%
Bone marrow	8,910 $\pm$ 5,119	38 $\pm$ 27	0.42 $\pm$ 0.30%
Kidneys	1,504 $\pm$ 879	13 $\pm$ 14	0.88 $\pm$ 0.95%

Mean $\pm$ standard deviation