

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Automated Identification and Measurement of Haematopoietic Stem Cells in 3D Intravital Microscopy Data

Reema Adel Khorshed and Cristina Lo Celso

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/64089>

Abstract

Image analysis and quantification of Haematopoietic stem cells (HSCs) position within their surrounding microenvironment in the bone marrow is a fast growing area of research, as it holds the key to understanding the dynamics of HSC-niche interactions and their multiple implications in normal tissue development and in response to various stress events. However, this area of research is very challenging due to the complex cellular structure of such images. Therefore, automated image analysis tools are required to simplify the biological interpretation of 3D HSC microenvironment images. In this chapter, we describe how 3D intravital microscopy data can be visualised and analysed using a computational method that allows the automated quantification of HSC position relative to surrounding niche components.

Keywords: intravital microscopy, 3D image analysis, bone marrow visualisation, haematopoietic stem cell niche, object segmentation and classification

1. Introduction

Somatic stem cells have the extraordinary ability to maintain their own pool, while replenishing dead cells and regenerating tissues after injuries throughout our lifetime. Dividing stem cells have the potential to differentiate into other cell types such as blood cells, skin cells or brain cells, or maintain themselves through a process called “self-renewal”. The stem cells that form blood and immune cells are called haematopoietic stem cells (HSCs). These stem cells are responsible for turnover and maintenance of red blood cells, platelets and immune cells. During

differentiation, HSCs generate multi-potent and lineage-committed progenitor cells prior to attaining maturity, which results in the generation of billions of new blood cells every day [1].

The bone marrow (BM) is the main site, where HSCs and their immediate progeny reside, precisely contained in a regulated and a very complex environment called the niche. The niche microenvironment has a direct impact on the function of HSCs, determined by the interaction of the HSCs themselves with particular cellular and molecular components in their surroundings [1, 2].

To observe and understand such interactions, single-cell resolution intravital microscopy of fluorescent HSCs and niche components is a vital tool, allowing the study of stem cells, their behaviour and interactions during steady state, aging and disease [3, 4]. This method has been successfully utilised to directly observe the HSCs with their niche in mouse bone marrow. Fluorescent dyes were used to label HSCs prior to transplantation into transgenic reporter mice that expressed GFP under the control of an osteoblast-specific promoter. Osteoblast cells reside on the bone marrow endosteal surface and contribute to the formation of bone [5]. This kind of fluorescent niche component served as a fundamental tool for the visualisation of the HSC endosteal niche and resulted in multiple observations indicating that functional HSCs localise near osteoblastic cells [5–7]. Intravital microscopy of HSCs injected into wild type and other bone marrow stroma/haematopoietic reporters indicated the importance of α 5 β 1 integrin receptor subunit on homing HSCs [8], of interactions with nestin-positive mesenchymal stem and progenitor cells [9] and with regulatory T cells [10]. Time-lapse imaging has revealed that infection-exposed HSCs interact with larger endosteal niches [7, 11] and that aged HSCs interact with the bone marrow differently than young ones [12]. However, a clear understanding of how the coordinated action of multiple niche components regulates HSC fate is not clear, and considering the localization of HSCs relative to multiple surrounding cellular and structural constituents of the bone marrow microenvironment is the first step towards unravelling HSCs dynamics and signal exchanges.

Even though manual distance measurement of HSCs to other niche components is possible [1, 6, 13], it suffers from various limitations including the extensive time spent on conducting such manual analysis as well as intra- and inter-researcher inconsistencies and human error. Therefore, a specialised image analysis tool capable of completing such tasks will add the great benefits of producing consistent unbiased results, and will simplify the interpretation of the biological microenvironment events of the bone marrow niche.

During the last decade, several attempts have been made to develop image analysis approaches to automate the quantification of medical and biological images, from cell segmentation and detection to tumour diagnosis and classification [14–16]. However, the direct application of such methods to three-dimensional (3D) intravital microscopy images of the bone marrow is not possible. This is due to the complex structure of the tissue, decreased signal with increased depths, unpredictable scattering of both the excitation and emission photons and the irregular encasing bone structure, which restricts the resolution of intravital microscopy images when compared to *ex vivo* and soft tissue imaging.

Consequently, we developed an image analysis pipeline, to specifically tackle all the main challenges associated with intravital microscopy 3D images of bone marrow. The approach starts by detecting and segmenting HSCs and surrounding niche components, for instance, the bone and osteoblast. Then it classifies each segmented HSCs and assigns them into one of three classes, and finally it quantifies them based on their location to surrounding niche components. This analysis will allow the extraction of variable quantitative features, which could potentially reveal novel aspects of HSC biology, and especially their relation to the bone marrow microenvironment [11].

This chapter is organised as follows: first an introduction of the bone marrow and haematopoietic stem cells niche highlighting the importance of intravital microscopy and automated image analysis is provided (Section 2), followed by a detailed description of intravital microscopy of the bone marrow and the main challenges faced during the image analysis stage (Section 3). The automated image analysis tool is then described (Section 4), detailing how the three main segmentation parameters can be optimised for a particular cell or niche component (Section 5). Finally, a description of how a supervised machine learning classifier can be used to classify both HSCs and vasculature is presented (Section 6), followed by a brief description of how to obtain 3D positional measurements of HSCs to other niche components (Section 7).

2. Intravital microscopy of mouse bone marrow

In vivo imaging of stem cells is a growing field, providing unique insights of their behaviour and especially of their interaction with their surrounding microenvironment. This approach has been instrumental in producing new hypotheses and revealed a number of novel findings concerning the regulation of fate decision of multiple somatic stem cells [17, 18], including HSCs [1, 3, 4].

Nevertheless, how multiple niche components affect and regulate HSCs fate is still a significant question that scientists are attempting to answer. Recording the localization of HSCs relative to multiple surrounding cellular and structural constituents of the bone marrow microenvironment is the first step towards understanding such phenomena.

Confocal fluorescence and second harmonic generation microscopy of HSCs and their niche components in the mouse calvarium (top of the skull) bone marrow provides a powerful tool to observe cellular interactions and has been successfully used to detect and study fluorescently labelled HSCs and GFP expressing osteoblasts and other niche components in transgenic reporter mice [7, 11].

Confocal microscopy provides a better 3D optical resolution than epifluorescence microscopy because it restricts the light that reaches the photomultiplier through a pinhole. While in epifluorescence microscopy in-focus image objects are mixed with out-of-focus image information arising from regions outside the focal plane, in confocal microscopy the pinhole blocks out-of-focus signal. This technology has allowed major advances in the field of biological imaging, due to its cost-effective solution and ease of use. In addition, confocal microscopes

have improved sensitivity, resolution and intensity compared to epifluorescent microscopes. Confocal microscopy is also less phototoxic and can therefore be used to generate not only static but also dynamic data of living cells within their tissues, through time-lapse acquisition. All these qualities make confocal microscopes an ideal instrument for *in vivo* studies.

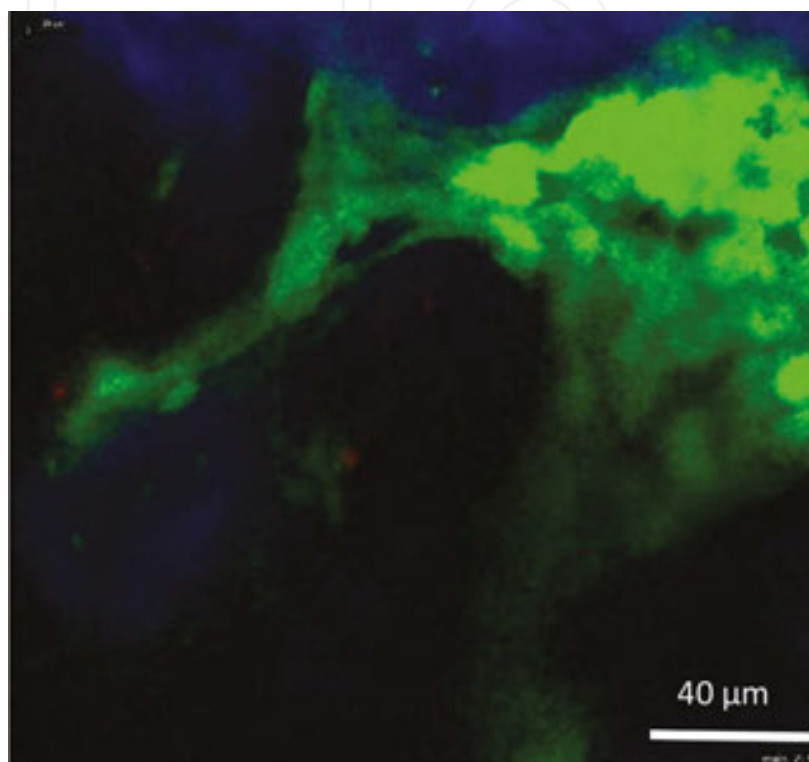


Figure 1. Maximum intensity projection of a 3D stack example of raw bone marrow *in vivo* images including DiD signal (red), GFP signal from osteoblastic cells (green) and bone collagen SHG signal (blue). Confocal microscopy was used to generate the signal for both DiD-labelled HSCs and osteoblasts using 633 and 488 nm lasers, respectively. Second harmonic generation (SHG) signal was obtained using 840-nm excitation of collagen to resolve bone structures.

Two-photon excitation of a fluorophore provides 3D optical sectioning similar to confocal imaging. However, it uses a wavelength roughly twice the length of the absorption peak of the specimen being imaged. In this chapter, we describe a particular application of two-photon excitation, named second harmonic generation (SHG). This takes advantage of a characteristic specific to certain molecules, such as collagen, which is the emission of photons at half of the wavelength of the exciting photons. Because collagen is one of the main components of bone, SHG signal allows detection of calvarium bone without the need to use specific transgenic fluorescent reporter mouse strains. An example of raw bone marrow *in vivo* images containing HSCs, osteoblast and SHG bone signal is presented in **Figure 1**. Two-photon excitation is also widely used to detect signal from multiple fluorophores, including GFP and tomato fluorescent protein [17], and signal obtained from two-photon excitation fluorescence microscopy can be analysed with the same computational algorithm we describe below.

3. Application and challenges of bone marrow intravital microscopy image analysis

The combination of confocal and SHG microscopy has greatly simplified high-resolution fluorescence imaging of animal tissues and organs. In Section 2, we have briefly described the advantages of the combination of confocal/SHG microscopy. Here, we describe the experimental set-up used and the challenges posed by the intravital images when they have to be further analysed. In the following sections, we describe in details how HSCs and bone marrow components can be automatically segmented, classified and measured.

To visualise HSCs, they were first labelled *ex vivo* using lipophilic membrane dyes such as 1,1'-Diocadecyl-3,3',3'',3'''-Tetramethylindodicarbocyanine (DiD) to generate a bright fluorescent signal [1, 5]. DiD-labelled HSCs were then injected into Col2.3GFP recipient mice, which allow the visualisation of osteoblasts as GFP-positive cells. Confocal microscopy was used to generate the signal for both DiD-labelled HSCs and osteoblasts using 633 and 488 nm lasers, respectively. Second harmonic generation (SHG) signal was obtained from 840-nm excitation of collagen to resolve bone structures. Acquisition setting (e.g. gain, laser powers, step and stack size) can vary between multiple users, leading to overall brighter/dimmer images and a range of field of view sizes and depths. Such variance imposes a challenge for automated image analysis. Therefore, developing an image analysis tool that can deal with the variability of setting preferences is a crucial point.

Another challenge imposed on image analysis from *in vivo* microscopy data is the decreased signal at increased depths. This results in non-uniform intensity and brightness of objects. In addition, light scattering caused by the surrounding tissue and bone restricts the resolution of *in vivo* microscopy of bone marrow compared to that of other soft tissues or *ex vivo* techniques.

To address these issues, we developed a local heterogeneity-based image segmentation (LH-SEG) approach [11] that employs multi-resolution segmentation [19] and mean intensity difference to neighbour thresholding. The approach works by comparing local morphological and intensity characteristics of objects, which most often are smaller than the cells and structures recorded. These detected objects are then grouped based on their homogeneity with other neighbouring objects within a defined distance. Because the method is applied to each 2D slide, it ensures reliable edge detection and segmentation across cells and structures with high intensity heterogeneity, despite the decreased signal (intensity) at increasing depths.

DiD labelling produces a bright fluorescent signal of HSCs [1, 5]. However, this and other related dyes have a number of limitations when associated with intravital microscopy:

1. The dye often diffuses from the labelled cells into the surroundings, and is diluted upon cell division, resulting in a loss of brightness and intensity of signal.
2. Labelling HSCs does not always provide a homogenous staining, which can result in some HSCs being brighter than others.
3. The dye often leads to background noise as result of cell debris and aggregate signal.

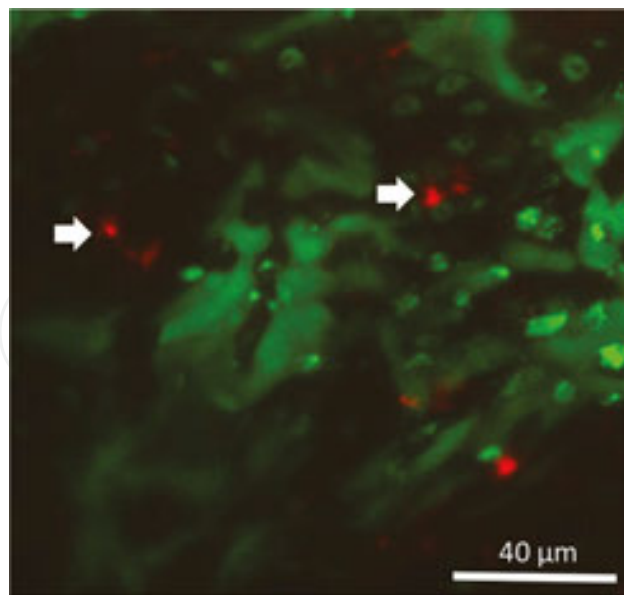


Figure 2. Maximum intensity projection of a 3D stack example representing the challenges for automated analysis of raw bone marrow *in vivo* images. DiD signal is in red. Arrows point at single DiD-labelled HSCs with varying intensity levels and sizes; the remaining red signal represents debris and background noise. Osteoblast cells (green) show non-uniform intensity levels too.

Figure 2 shows examples of the points listed above.

To overcome these issues, LH-SEG detects cells with variable intensities and signal brightness. Moreover, a machine learning classification protocol based on morphological and textural features recognises and classifies all segmented DiD signal, to differentiate real HSCs from cell debris and aggregate.

4. Image analysis

Image analysis is the process of extracting meaningful information from images, using manual or automated methods (the latter known as computer vision techniques) [20, 21]. The selection of the appropriate image analysis method for a certain type of images determines the success rate of the analysis. Therefore, understanding the challenges associated with the images acquired is the first step towards developing an effective image analysis solution.

Automated image analysis has a number of advantages over the human manual analysis. Human vision can easily be biased by pre-conceived concepts, affecting the output results and the rigorous testing of hypothesis. Manual analysis can also be time-consuming compared to automated analysis tasks, in which large datasets can be batch processed without the need of human monitoring, allowing users to perform other tasks while the analysis is being carried out.

Microscopy images are often complex, with a range of artefacts and background noise, which require variable image processing steps before any meaningful quantification can be extracted

from objects and region of interests. An overall image analysis protocol, including image processing, analysis and data output, needs to be designed and tailored according to the targeted image datasets. A general outline of the image analysis pipeline described in this chapter is provided in **Figure 3**.

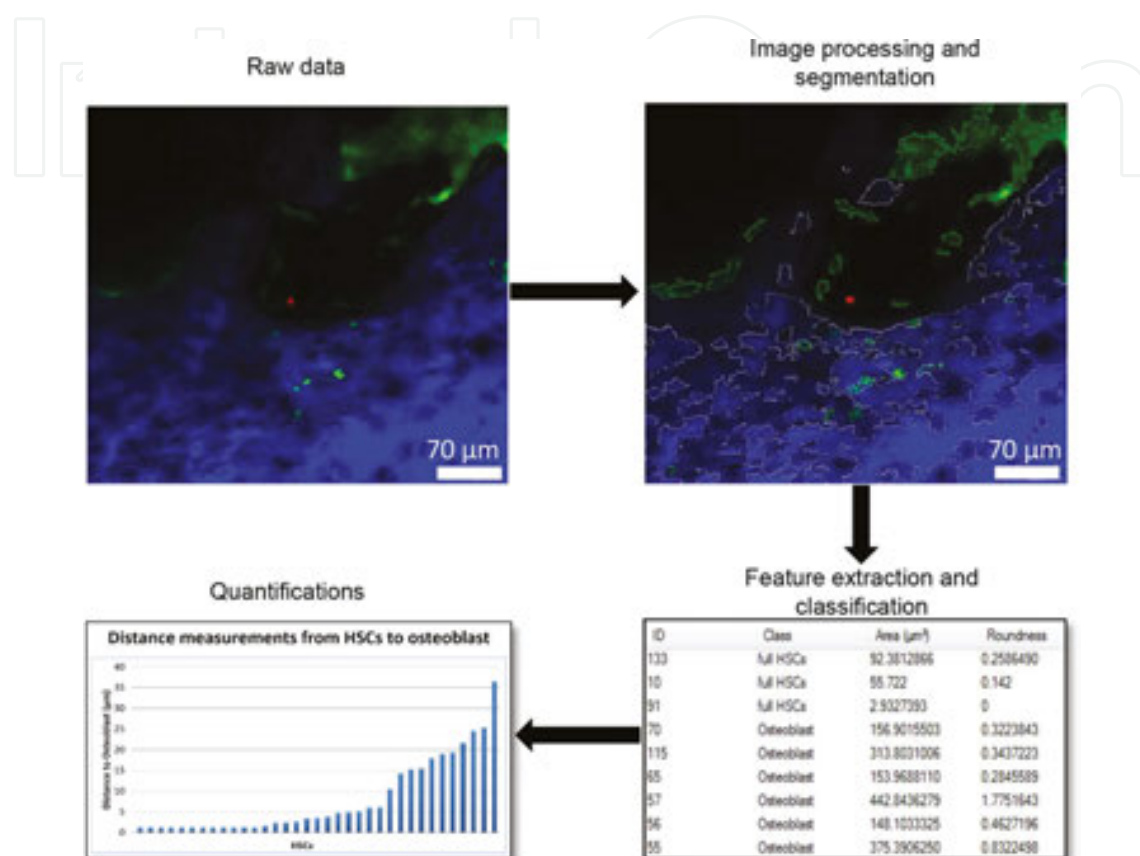


Figure 3. Outline of our image analysis protocol, starting from raw data and applying image processing and segmentation, feature extraction and classification, and finally distance measurements.

4.1. Image acquisition and processing

Image acquisition is the process of obtaining the raw image data using a microscope. Such data usually contain a number of imperfections, for example, due to oversaturation, out of focus signal and uneven excitation due to the irregular structure of the tissue itself.

To overcome these issues, filtering techniques are usually introduced to the image analysis as a pre-processing step. Smoothing and de-noising filters such as the median filter and Gaussian filter are applied to enhance image quality and reduces the noise introduced by the image acquisition process.

In the case of our *in vivo* bone marrow images, we applied Gaussian blur [22] to handle fluorescence intensity heterogeneity. The blur is typically used to smooth images detail and

reduce noise from uneven signal intensity. A small kernel size of (3×3) was used to reduce noise while avoiding over-smoothing of image objects.

4.2. Image segmentation

Object detection is a crucial step, as it determines the objects of interest in an image. This step ultimately governs the quality of quantifications extracted from each object. A very effective method of object detection is image segmentation. Segmentation is the process of dividing an image into smaller meaningful segments, by selecting a group of pixels to represent a region, or an object, contained within a border. Objects are segmented such that the pixels enclosed within one border share certain characteristics that define them as an object.

Successful segmentation is dependent on the method used, which critically needs to be tailored for each particular set of images. Images with a high level of contrast between objects of interests and background and with uniform intensities can be segmented using simple thresholding and edge detection approaches. More complex images with high intensity heterogeneity and intricate structures such as *in vivo* single cell resolution images require a more sophisticated segmentation approach.

The main challenges imposed upon segmentation of bone marrow *in vivo* images are the heterogeneity of fluorescent intensity as represented in GFP and DiD signal in **Figure 2**, the loss of signal with increasing depths and the unpredictable shape of stroma components. To overcome these issues, we have developed a two-step method of segmentation (LH-SEG). This method combines two powerful approaches: the multi-resolution segmentation and mean intensity difference to neighbourhood thresholding (MDN).

This segmentation approach was mainly developed to minimise artefacts due to loss of signal with increased depths. The first step of LH-SEG, multi-resolution segmentation, starts by dividing each 2D slice of the 3D stack into smaller segments grouped by their homogeneity in shape and texture. Homogeneity for each pixel is then calculated by selecting a scale parameter α . The parameter α is optimised for each object type (HSC, osteoblast and bone), taking into consideration their morphological and textural characteristics. Optimisation of this parameter for each object type is described in detail in Sections 4.3 and 5.

Adjacent homogeneous image segments are then merged based on their mean intensity difference to neighbourhood (MDN) threshold. MDN threshold calculates the difference between an image segment and its neighbouring image segments using the mean intensity values described in [11] as follows:

$$T_{\Delta_k}^-(v) = \frac{1}{w} \sum_{u \in N_v(d)} w_u (\bar{c}_k(v) - \bar{c}_k(u)) \quad (1)$$

where w is the image channel weight. Images are weighted by the distance between the segmented image objects, defined as follows:

$$w = \sum_{u \in N_v(d)} w_u \quad (2)$$

where v and u are two segmented image objects, N_v is the direct neighbour to the segmented image object v . u is defined as a direct neighbour to v if the minimum distance between them is less than or equal to d . d is the distance between neighbouring segments and defined as the radius of the segmented image object perimeter in pixel. w_u is the weight of the segmented image object defined by the difference of the mean intensity value between v and u in a given distance d . $\bar{c}_k$ is the mean intensity value of channel k . The appropriate MDN threshold for $T_{\bar{\Delta}_k}$ and distance featured d need to be selected for effective segmentation of each image object category (HSC, osteoblast and bone) (from [11]).

Figure 4 shows how, at the end of this process, segmented images on each z plane are then linked together to generate a 3D rendering of the whole stack.

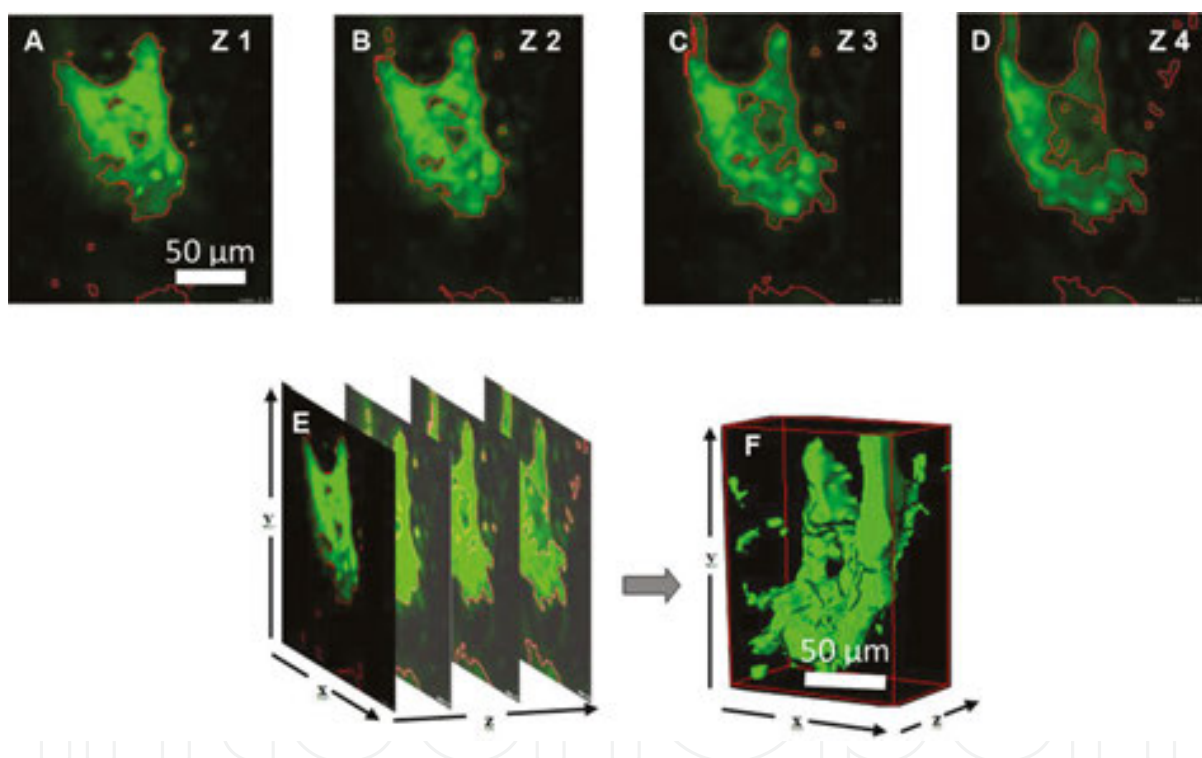


Figure 4. (A–D) segmented regions of osteoblast 2D slices using the LH_SEG method before they are merged (E) for the 3D rendering (F).

4.3. Parameter optimisation

Image segmentation is a fundamental step of image analysis, which ultimately determines the level of success of the overall quantification. Automated image analysis approaches can only perform sufficiently well when their parameters are optimally set for a certain type of images or dataset. In this section, we describe how the optimal settings can be selected for datasets containing intravital microscopy bone marrow 3D stacks. Three fundamental parameters are

targeted: (1) the multi-resolution segmentation scale parameter α ; (2) MDN threshold $T_{\bar{\Delta}_k}$; and (3) the MDN distance between neighbouring image segments d . The optimisation of parameters will be discussed in detail for each image object category (HSCs, osteoblast and bone) in Section 5.

The selection of the scale parameter α depends on the physical and textural structure of the objects. Different values of parameter α for the multi-resolution segmentation influence the output of the segmentation process. Selecting high scale parameters results in fewer, larger segments that can be bigger than the object observed, while lower scale parameter values result in smaller objects. To illustrate the consequences of using a range of parameter values for α on the resulting segmentation, we will present parameter values that would represent excessive cases of error when increasing or decreasing the scale parameter, as shown in Figure 5.

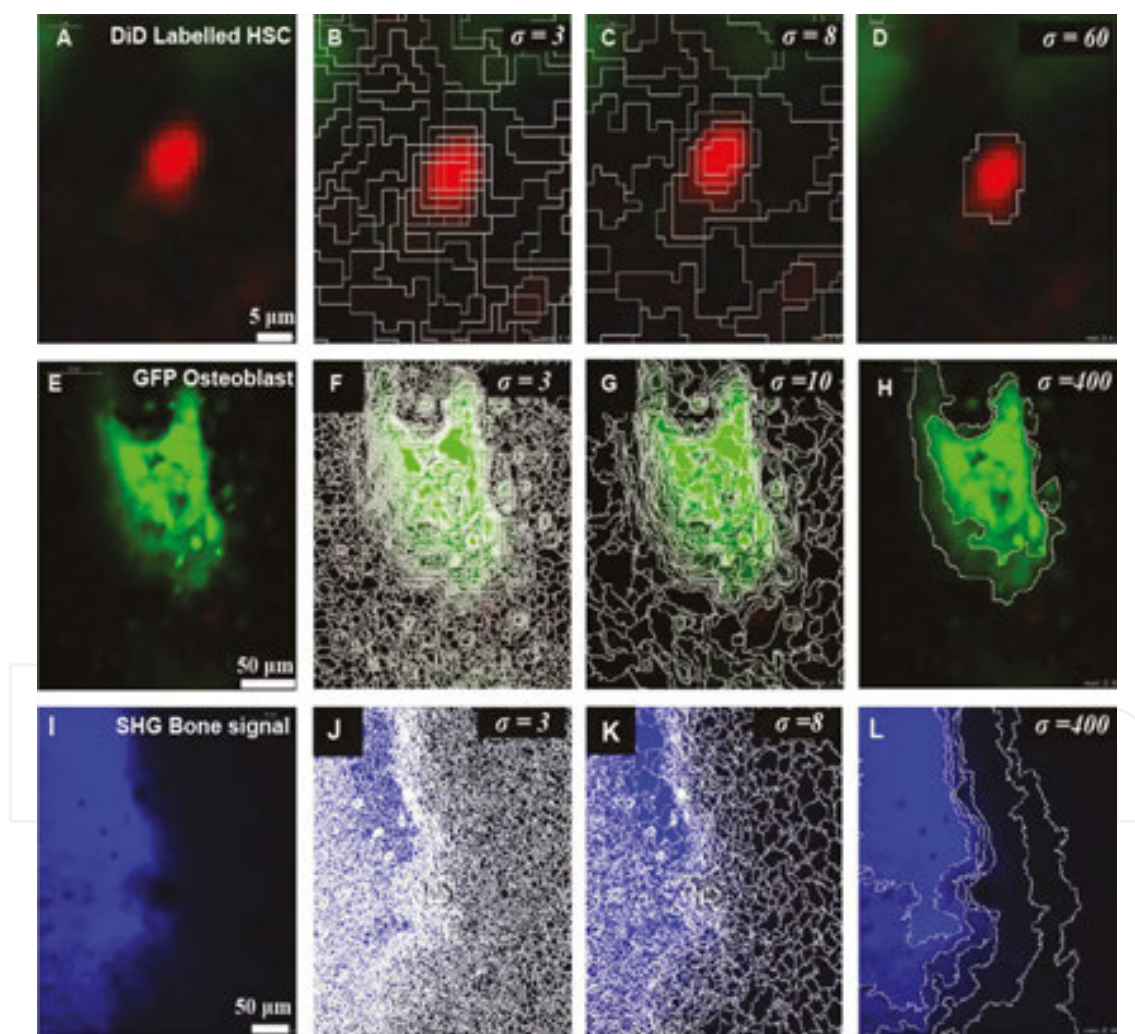


Figure 5. Optimisation of the multi-resolution segmentation parameter α for each image object category. Column (A–I) represent the data of each object category after smoothing with the convolution filter. Column (B–J) represent region segmentation error when decreasing the α parameter. Column (C–K) represent correct region segmentation with the optimal parameter α . Column (D–L) represent region segmentation error when increasing the α parameter.

MDN threshold has two parameters: MDN threshold value $T_{\bar{d}_k}$ and the distance to neighbouring image segments d . Results of all three parameters optimisation for each component type (HSCs, osteoblast and bone), together with examples of the consequences of using excessively high or low parameter values, are described in Section 5 and presented in **Figures 6** and **7**.

5. Parameter optimisation

In this section, we present examples of the results obtained when different parameter values for α , $T_{\bar{d}_k}$ and d are selected, to help the reader apply the same decision process to his/her own image dataset.

5.1. Multi-resolution parameter α

5.1.1. α for HSCs

Selecting a small value such as $\alpha = 3$ results in the region of the image corresponding to an HSC being divided in small segments (irregular boxes delimited by white lines, **Figure 5**) and increases the running time of the overall process. On the other hand, a scale parameter of $\alpha = 8$ leads to larger and fewer segments, covering the area occupied by the HSC and, once combined, describing its edges, therefore, serves as the optimal parameter for HSCs. Increasing the scale parameter α will result in even fewer, larger segments comprising areas of heterogeneous intensity and causing some background signal to be included in the HSC segment region (**Figure 5A–D**).

5.1.2. α for osteoblasts

Selecting a small value such as $\alpha = 3$ results in the osteoblastic cell region being divided in very small segments (irregular boxes delimited by white lines) and increases the running time of the overall process. A scale parameter of $\alpha = 10$ leads to fewer, larger segments, which cover the area occupied by the osteoblastic cells and combined describe their edges. Having fewer segments reduces the processing time. Therefore, $\alpha = 10$ serves as the optimal parameter for osteoblasts. Increasing the scale parameter α marginally did not have a major change in the segments produces, however, increasing the parameter extensively to e.g. $\alpha = 400$ results in larger segments comprising areas of heterogeneous intensity and causing some background signal to be included in the osteoblastic segment region (**Figure 5E–H**).

5.1.3. α for bone

Selecting a small value such as $\alpha = 3$ results in the bone region being divided into small segments (irregular boxes delimited by white lines) and increases the running time of the overall process. Similarly to osteoblasts, a scale parameter $\alpha = 8$ leads to fewer, larger segments covering the area occupied by the bone and bone cavities and, once combined, describing its edges. Again,

dealing with fewer segments reduces the processing time. Therefore, $\alpha = 8$ serves as the optimal parameter for bone. Increasing the scale parameter α marginally did not have a major impact on the segments produced, however, increasing the parameter extensively to e.g. $\alpha = 400$ results in larger segments comprising areas of heterogeneous intensity and causing some background signal to be included in the bone segment region **Figure 5I–L**.

5.2. MDN threshold value $T_{\bar{\Delta}_k}$

5.2.1. $T_{\bar{\Delta}_k}$ for HSCs

The optimisation of the MDN threshold value depends on the correct detection of HSC edges. To observe the effects of changing the MDN threshold on HSC segmentation, we set $\alpha = 8$ and $d = 30$, then change the parameter $T_{\bar{\Delta}_k}$. Selecting $T_{\bar{\Delta}_k} \geq 10$ resulted in some background signal being included in the HSC segment region, while $T_{\bar{\Delta}_k} \geq 68$ resulted in correct segmentation of the HSC region. On the other hand, selecting a threshold value of $T_{\bar{\Delta}_k} \geq 200$ resulted in restricting the selection of segments to only those which have $T_{\bar{\Delta}_k} \geq 200$ mean intensity difference value to their neighbouring segments and excluded HSC segments of lower intensity from the final HSC object. Hence selection of low threshold values for the MDN will result in increased segment region sizes, while selecting higher MDN threshold values will restrict the selection of segments to regions of smaller size (**Figure 6A–C**).

5.2.2. $T_{\bar{\Delta}_k}$ for osteoblasts

To observe the effects of changing the MDN threshold on osteoblastic cell segmentation, we set $\alpha = 8$ and $d = 30$, then change the MDN threshold parameter $T_{\bar{\Delta}_k}$. $T_{\bar{\Delta}_k} \geq 0$ resulted in some background signal being included in the osteoblastic cell segment regions, while $T_{\bar{\Delta}_k} \geq 8$ resulted in correct segmentation of the osteoblastic cell regions. $T_{\bar{\Delta}_k} \geq 100$ resulted in restricting the selection of segments to only those which have $T_{\bar{\Delta}_k} \geq 100$ mean intensity difference to their neighbouring segments and excluded osteoblast segments of lower intensity from the final osteoblastic cell objects. Hence selection of low threshold values for the MDN will result in increased segment regions, while selecting higher MDN threshold values will restrict the selection of segments (**Figure 6D–F**).

5.2.3. $T_{\bar{\Delta}_k}$ for bone

To observe the effects of changing the MDN threshold on bone segmentation, we set $\alpha = 8$ and $d = 70$, then change the MDN threshold parameter $T_{\bar{\Delta}_k}$, while $T_{\bar{\Delta}_k} \geq 0$ resulted in some background signal being included in the bone segment region. $T_{\bar{\Delta}_k} \geq 4$ resulted in correct segmentation of the bone region, $T_{\bar{\Delta}_k} \geq 80$ resulted in restricting the selection of segments to only those which have ≥ 80 mean intensity difference to their neighbouring segments and excluded bone segments of lower intensity from the final bone object. Hence selection of low

threshold values for the MDN will result in increased bone segment regions, while selecting higher MDN threshold values will restrict the selection of segments (**Figure 6G–I**).

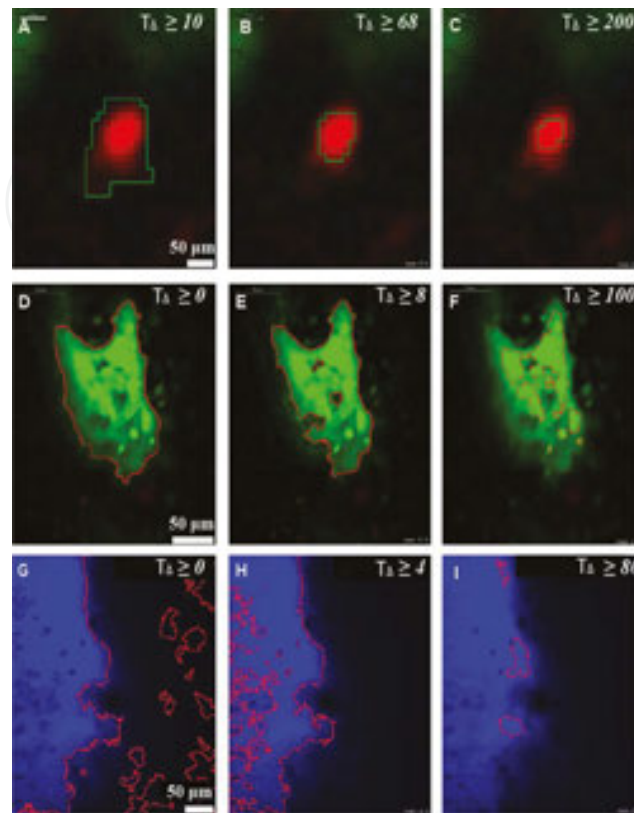


Figure 6. Optimisation of the MDN threshold parameter $T_{\bar{\Delta}}$ for each image object category. Column (A–G) represent object segmentation error when decreasing the $T_{\bar{\Delta}}$ parameter. Column (B–H) represent correct object segmentation with the optimal parameter $T_{\bar{\Delta}}$. Column (C–I) represent object segmentation error when increasing the $T_{\bar{\Delta}}$ parameter.

5.3. Distance to neighbour d parameter

5.3.1. d for HSCs

To illustrate the process of selecting the optimal distance to neighbour d parameter, we use varying distance d between neighbouring segments as follows:

We set the optimal multi-resolution scale parameter $\alpha = 8$ and MDN $T_{\bar{\Delta}_k} \geq 68$, then select the neighbourhood size. Small neighbourhood size such as $d = 2$ restricts the selection to fewer segments, specifically to those within a two pixels radius from the centre of the HSC region. As a result, the edges of the HSC are excluded from the final object. On the other hand, increasing the neighbouring distance to $d = 30$ resulted in correct selection of segments belonging to the HSC and elimination of segments containing background signal. Increasing the neighbouring distance further to about $d = 120$ resulted in some background signal being included in the final HSC region (**Figure 7A–C**).

5.3.2. d for osteoblasts

To show the effects of changing the neighbourhood sized d on osteoblast segmentation, we use the optimal multi-resolution scale parameter $\alpha = 8$ and MDN threshold to $T_{\Delta_k} \geq 8$, then select the neighbourhood sized d . Selection of small neighbourhood size such as $d = 2$ restricted the selection of segments, resulting in fragmented segments scattered across the osteoblastic cell regions and failed to detect actual osteoblast regions. On the other hand, $d = 2$ resulted in correct selection of segments belonging to the osteoblasts and elimination of segments containing background signal. Increasing the neighbouring distance to $d = 120$ resulted in some background signal being included in the final osteoblast region (**Figure 7D–F**).

To demonstrate the effect of the selection of the value of parameter d on low intensity osteoblast regions, we selected images from lower slices of the same 3D stack of images (**Figure 7G–I**).

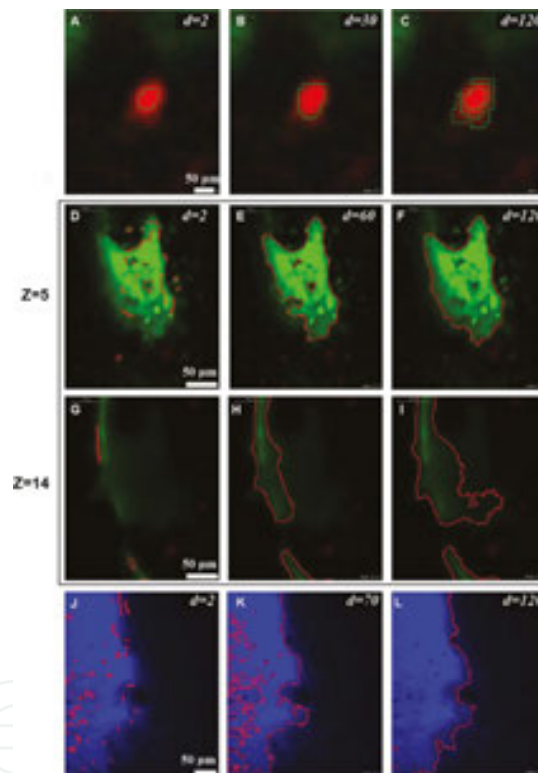


Figure 7. Optimisation of Distance to neighbour d parameter for each image object category. Column (A–J) represent object segmentation error when decreasing the d parameter. Column (B–K) represent correct object segmentation with the optimal parameter d . Column (C–L) represent object segmentation error when increasing the d parameter. Row (G–I) represent images from lower slices of the same 3D stack of images (D–F).

5.3.3. d for bone

To show the effect of changing the neighbourhood size d on bone segmentation, we use the optimal multi-resolution scale parameter $\alpha = 8$ and MDN threshold to $T_{\Delta_k} \geq 4$, then select the neighbourhood sized d . Selection of small neighbourhood size such as $d = 2$ restricted the

selection of segments, resulting in fragmented segments scattered across the bone regions and failed to detect actual bone regions. As a result, the edges of the bone are excluded from the final object. On the other hand, $d = 70$ resulted in correct selection of segments belonging to the bone and elimination of segments containing background signal. Increasing the distance to $d = 120$ resulted in some background signal being included in the final bone region and missing small detail of bone cavities (**Figure 7J–L**).

Note: Parameter selection can be further observed on the YZ and XZ dimensions to ensure accurate selection of the optimal parameters (**Figure 8**). To optimise the parameters for further niche components and cell types (e.g. vasculature), the same steps can be followed as described for the HSCs, osteoblasts and bone.

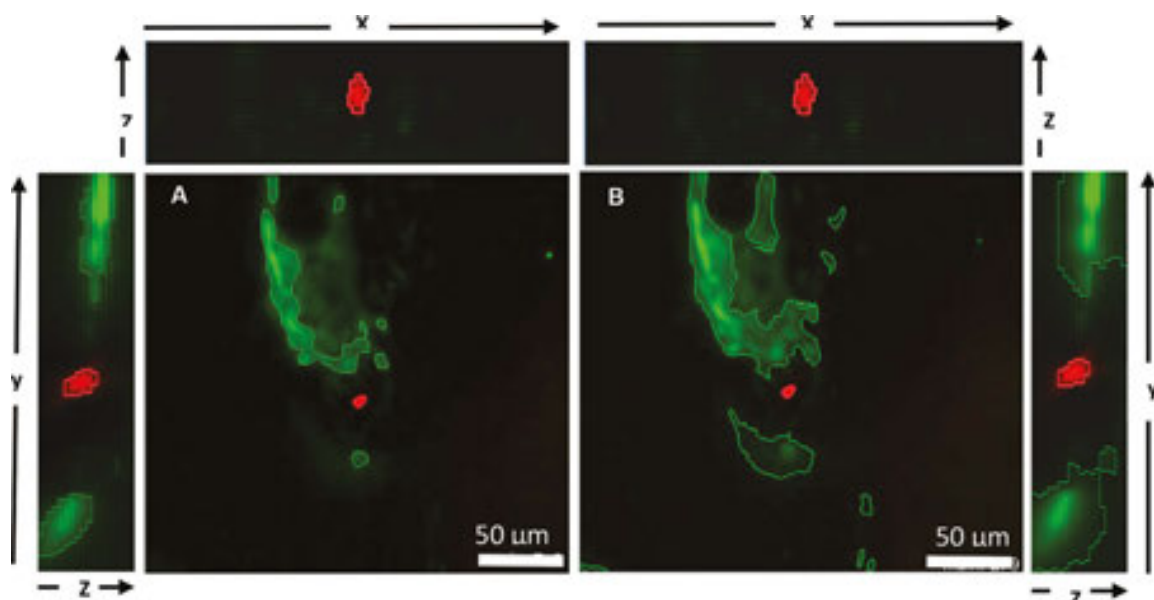


Figure 8. Examples of different dimensions view to observe the segmentation results following selecting different parameter values for α , T_{Δ_k} and d .

6. Machine learning

Machine learning (ML) has become a valuable artificial intelligence tool, increasingly used for analysis of complex image data [23, 24]. ML serves two main objectives: classification and regression. Classification approaches are performed when a computer is given a set of options and is expected to divide them into a subset of categories. Regression is the process of selecting the appropriate response to a particular situation from a set of possible responses. Furthermore, ML classifiers can be categorised into two main models: supervised and unsupervised classifiers. Unsupervised approaches utilise a clustering technique, where ML looks for resemblances across data, and then splits the same data into clusters. Clusters are then used to define the classes. Conversely, supervised approaches require a training set, where the ML

learns information by extracting a set of features from that particular set, and the ML classifier is then expected to classify a new set of data and categorise it into a k number of classes based on the discriminative features extracted from the training set. Supervised classifiers are capable of identifying a set of complex features, suitable for classifying heterogeneous imaging data such as HSCs and vasculature from *in vivo* bone marrow 3D images, and then use them to perform such classification.

6.1. Decision tree classifier

The decision tree (DT) classifier is a structured approach that builds classification models from an input dataset to predict the output of an unknown dataset. The DT classifier starts the classification procedure structurally from the top node and uses the feature vectors to split it into further nodes [25, 26]. The selection of the feature vectors is measured by the purity of a particular subset, such that if a pure subset is produced, the splitting stops, otherwise splitting continues until a pure subset is allocated. Impurity of a specific subset is calculated by the entropy and is defined in [27] as follows:

$$E(S) = -\sum_{c \in C} p(c) \log_2 p(c) \quad (3)$$

where S is the dataset for which the entropy is calculated, c the set of classes in set S , $p(c)$ the proportion of the number of elements belonging to class c to the number of elements in set S . Pure subsets will generate a value of 0, while impure subsets will generate a value of 1.

$$I(S, F) = \sum_i \frac{|S_i|}{|S|} E(S_i) \quad (4)$$

where F is a feature. When a feature F splits the set S into subsets S_i , the average entropy is computed and the sum compared to the entropy of the original set S ; from [28].

$$Gain(S, F) = E(S) - I(S, F) = E(S) - \sum_i \frac{|S_i|}{|S|} E(S_i) \quad (5)$$

In this chapter, we describe the application of the DT classifier for classifying DiD signal (potential HSCs) and vasculature structures. The DT classifier was selected in this study due to its computational simplicity and illustrative attributes represented in the DT output. The classifier also has the advantage of being able to select discriminative features without a prior step of feature optimisation, allowing the incorporation of a wide range of features for testing at the training stage.

6.2. Training and testing the classifier

6.2.1. HSCs

LH-SEG provides a sufficient level of DiD objects detection. However, not all the DiD objects are HSCs. This is due to cell debris and the diffusion of the dye, which could lead to aggregates producing signal similar in shape and intensity to that of the HSCs.

In order to identify HSCs, we propose the use of DT classifier. We first trained our classifier to distinguish three classes of DiD objects: Class-1 represents HSCs distinguished by their smoother, round surface and high intensity as observed for quiescent HSCs; Class-2 comprises HSCs that have less rounded shapes and present small uropod-like protrusions [27], as previously observed in time-lapse images of migratory HSCs [7]; and Class-3 contains DiD objects that are not HSCs and characterised by their exceedingly uneven morphology. A selection of objects was manually selected for each category and fed into the classifier as training set. The exact number of objects needed for training each class depends on the variability of the structures contained, such that objects with regular, predictable characteristics such as Class-1 HSCs would require fewer objects for training, while irregular objects containing high variability such as class-3 would require a higher number of objects for training to ensure the classifier collects sufficient numbers of characteristics to handle the complexity and inconsistency of objects. After training, the classifier is tested on different datasets from those used for training with a 3-fold cross-validation approach. Classification results provided a high accuracy for all classes and successfully classified the DiD object to the three proposed classes as described in [11] and **Figure 9A and B**.

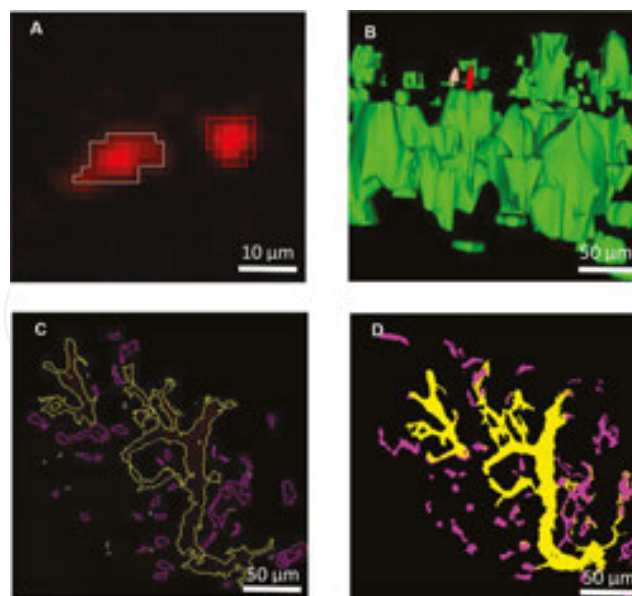


Figure 9. (A) Classification results of DiD objects using our proposed approach: two DiD-labelled HSCs were classified as HSC class 1 (red) and 2 (moccasin). (B) 3D rendering of DiD HSCs represented in (A) inside their microenvironment. (C) Classification results of vasculature using our proposed approach, showing small sinusoids (magenta) and large sinusoids (yellow). (D) 3D rendering of the 3D stack containing the slice shown in (C).

6.2.2. Vasculature

Blood vessels inside the bone marrow are another component that plays a vital role in the maintenance of HSCs. Interactions between vasculature and HSCs have become one of the main headlines for many recent studies [29, 30]. As demonstrated in our previous study, segmentation of vasculature of *in vivo* images was possible using the LH-SEG [11]. However, the automated identification of the variable components of the BM vascular network is a challenging task due to the complex and interlinked structure of blood vessels. Therefore, the selection of effective thresholds for classifying the vasculature into different regions based on their appearance and morphology is a challenging task. To overcome these issues, we propose a DT Classifier to define distinct features of the different types of vasculature and classify them into four categories based on their morphological and topological characteristics inside the BM space [28]. Blood vessels share mutual features in relation of their complexity and interlinking physical structure. Consequently, we manually select different 2D segmented from the vasculature region to create the training sets for four classes of blood vessels: (1) large sinusoids: larger blood vessels in comparison to other sinusoids and positioned mostly towards the periphery of the BM space; (2) small sinusoids: blood vessels smaller in section when compared to the large sinusoids, spread across bone marrow space; (3) the central sinus: a large venous vessel situated in the centre of the bone marrow space; and (4) the bifurcation of the central sinus, which results from the central sinus branching out towards either sides of the BM space in its frontal area. The number of samples was optimised for each class to avoid under-fitting and over-fitting the classifier. Three-fold cross validation was used in this experiment. Morphological and topological features were fed to the classifier for extraction of discriminative values. Classification results provided a high accuracy for all classes and successfully classified all vessels into one of the four classes as described in [23] and presented in **Figure 9C and D**. This demonstrates the applicability of ML classification approaches to 3D *in vivo* images of the bone marrow, which escalates the throughput of intravital imaging and our understanding of the complexity of the HSC and their interaction with multiple niche components.

7. 3D positional measurements

Following segmentation and classification of DiD-labelled HSCs, HSCs position relative to their nearest osteoblast and bone is measured as described in [11], and could be measured to their nearest vessel or other segmented bone marrow components. The overall automated distance measurements from HSCs to osteoblast and bone were equivalent to previously published data obtained using manual distance measurements and described in [5, 6]. Further quantitative measures such as the morphological and textural characteristics of each segmented and classified object can also be obtained using this tool.

8. Conclusion

In this chapter, we describe an approach for the visualisation, analysis and quantification of HSCs within their surrounding microenvironment in the bone marrow using intravital imaging and image analysis approaches.

Confocal fluorescence and SHG microscopy combined facilitate the collection of large amounts of image data of HSCs and their niche components. They specifically allowed the first observations of the cellular interactions of *ex vivo* labelled HSCs with GFP expressing osteoblasts and bone, and were further used to uncover several other details of HSC-niche interactions.

An image analysis approach combining image segmentation (LH-SEG) and machine learning classification detects and defines the edges of HSCs, osteoblast, bone and blood vessels. Subsequently, detected HSC and vasculature objects are fed into a decision tree classifier for categorising the detected signal into further classes. HSC signal was categorised into three classes representative of their morphological and intensity characteristics, proving high levels of accuracy for detection of HSCs and elimination of DiD debris and background signal. Furthermore, vasculature in the bone marrow was categorised into four classes based on their intrinsic morphological and position features. This will simplify the analysis of this component and allow simultaneous analysis of other components, increasing our understanding of the complexity of the HSC niche. After all the biologically important objects have been segmented and classified, it is then possible to derive quantitative information from these features and further translate it into a deeper understanding of the biological events observed.

This automated image analysis pipeline was particularly designed to address the major challenges imposed on *in vivo* 3D images of the bone marrow using confocal and SHG microscopy and provided high level of accuracy in both segmentation and classification of HSCs and their niche components. Moreover, optimisation of the segmentation parameters is fundamental for extracting resourceful and precise quantitative features from all segmented and classified objects. However, this method can be virtually applied to any complex 3D imaging data, especially those generated by confocal and two photon microscopy, such as *ex vivo* whole mount immunofluorescence of whole bones, and intravital microscopy data from other tissues and organs.

In conclusion, once the correct parameters are selected as we describe, this approach is applicable to the analysis of a broad range of 3D and intravital microscopy images from other tissues and organs, leading to a faster pace of discovery not only in the haematopoiesis and bone marrow field, but also in other somatic stem cell studies, and developmental and cancer biology studies. Further development of the described tool can be expanded to more programming languages and software platforms. This will accelerate the productivity of the method and allow an even wider range of applications.

Author details

Reema Adel Khorshed* and Cristina Lo Celso

*Address all correspondence to: r.khorshed@imperial.ac.uk

Department of Life Sciences, Imperial College, London, UK

References

- [1] C. Lo Celso, C. P. Lin, and D. T. Scadden, "In vivo imaging of transplanted haematopoietic stem and progenitor cells in mouse calvarium bone marrow," *Nat Protoc*, vol. 6, pp. 1–14, Jan 2011.
- [2] S. J. Morrison and D. T. Scadden, "The bone marrow niche for haematopoietic stem cells," *Nature*, vol. 505, pp. 327–34, Jan 16 2014.
- [3] L. Ritsma, S. I. Ellenbroek, A. Zomer, H. J. Snippert, F. J. de Sauvage, B. D. Simons, et al., "Intestinal crypt homeostasis revealed at single-stem-cell level by in vivo live imaging," *Nature*, vol. 507, pp. 362–5, Mar 20 2014.
- [4] A. Kohler, V. Schmithorst, M. D. Filippi, M. A. Ryan, D. Daria, M. Gunzer, et al., "Altered cellular dynamics and endosteal location of aged early haematopoietic progenitor cells revealed by time-lapse intravital imaging in long bones," *Blood*, vol. 114, pp. 290–8, Jul 9 2009.
- [5] C. Lo Celso, H. E. Fleming, J. W. Wu, C. X. Zhao, S. Miake-Lye, J. Fujisaki, et al., "Live-animal tracking of individual haematopoietic stem/progenitor cells in their niche," *Nature*, vol. 457, pp. 92–6, Jan 1 2009.
- [6] S. W. Lane, Y. J. Wang, C. Lo Celso, C. Ragu, L. Bullinger, S. M. Sykes, et al., "Differential niche and Wnt requirements during acute myeloid leukemia progression," *Blood*, vol. 118, pp. 2849–56, Sep 8 2011.
- [7] N. M. Rashidi, M. K. Scott, N. Scherf, A. Krinner, J. S. Kalchschmidt, K. Gounaris, et al., "In vivo time-lapse imaging of mouse bone marrow reveals differential niche engagement by quiescent and naturally activated haematopoietic stem cells," *Blood*, May 21 2014.
- [8] G. B. Adams, I. R. Alley, U. I. Chung, K. T. Chabner, N. T. Jeanson, C. Lo Celso, et al., "Haematopoietic stem cells depend on Galpha(s)-mediated signalling to engraft bone marrow," *Nature*, vol. 459, pp. 103–7, May 7 2009.
- [9] S. Méndez-Ferrer, T. V. Michurina, F. Ferraro, A. R. Mazloom, B. D. MacArthur, S. A. Lira, et al., "Mesenchymal and haematopoietic stem cells form a unique bone marrow niche," *Nature*, vol. 466, pp. 829–834, 2010.

- [10] J. Fujisaki, J. Wu, A. L. Carlson, L. Silberstein, P. Putheti, R. Larocca, et al., "In vivo imaging of Treg cells providing immune privilege to the haematopoietic stem-cell niche," *Nature*, vol. 474, pp. 216–219, 06/09/print 2011.
- [11] R. A. Khorshed, E. D. Hawkins, D. Duarte, M. K. Scott, O. A. Akinduro, N. M. Rashidi, et al., "Automated Identification and Localization of haematopoietic Stem Cells in 3D Intravital Microscopy Data," *Stem Cell Rep*, vol. 5, pp. 139–53, Jul 14 2015.
- [12] H. Geiger and K. L. Rudolph, "Aging in the lympho-haematopoietic stem cell compartment," *Trends Immunol*, vol. 30, pp. 360–5, Jul 2009.
- [13] A. Sanchez-Aguilera, Y. J. Lee, C. Lo Celso, F. Ferraro, K. Brumme, S. Mondal, et al., "Guanine nucleotide exchange factor Vav1 regulates perivascular homing and bone marrow retention of haematopoietic stem and progenitor cells," *Proc Natl Acad Sci USA*, vol. 108, pp. 9607–12, Jun 7 2011.
- [14] G. Chung and L. Vese, "Image segmentation using a multilayer level-set approach," *Computing and visualisation in Science*, vol. 12, pp. 267–285, 2009/08/01 2009.
- [15] S. Y. Yeo, X. Xie, I. Sazonov, and P. Nithiarasu, "Level set segmentation with robust image gradient energy and statistical shape prior," in *IEEE International Conference on Image Processing*, pp. 3397–3400, 2011.
- [16] K. Nandy, J. Kim, D. P. McCullough, M. McAuliffe, K. J. Meaburn, T. P. Yamaguchi, et al., "Segmentation and quantitative analysis of individual cells in developmental tissues," *Methods Mol Biol*, vol. 1092, pp. 235–53, 2014.
- [17] P. Rompolas, K. R. Mesa, and V. Greco, "Spatial organization within a niche as a determinant of stem-cell fate," *Nature*, vol. 502, pp. 513–518, 10/24/print 2013.
- [18] L. Ritsma, S. I. J. Ellenbroek, A. Zomer, H. J. Snippert, F. J. de Sauvage, B. D. Simons, et al., "Intestinal crypt homeostasis revealed at single-stem-cell level by in vivo live imaging," *Nature*, vol. 507, pp. 362–365, 03/20/print 2014.
- [19] G. Mallinis, N. Koutsias, M. Tsakiri-Strati, and M. Karteris, "Object-based classification using Quickbird imagery for delineating forest vegetation polygons in a Mediterranean test site," *ISPRS J Photogramm Remote Sens*, vol. 63, pp. 237–250, 3//2008.
- [20] F. Long, J. Zhou, and H. Peng, "Visualisation and analysis of 3D microscopic Images," *PLoS Comput Biol*, vol. 8, p. e1002519, 2012.
- [21] S. Uchida, "Image processing and recognition for biological images," *Dev Growth Differ*, vol. 55, pp. 523–549, 2013.
- [22] B. Moon, "A Gaussian smoothing algorithm to generate trend curves," *Korean J Comput Appl Math*, vol. 8, pp. 507–518, 01 2001.
- [23] E. Mjolsness and D. DeCoste, "Machine learning for science: state of the art and future prospects," *Science*, vol. 293, pp. 2051–5, Sep 14 2001.

- [24] C. Sommer and D. W. Gerlich, "Machine learning in cell biology — teaching computers to recognize phenotypes," *J Cell Sci*, vol. 126, pp. 5529–39, Dec 15 2013.
- [25] C. Agarwal and A. Sharma, "Image understanding using decision tree based machine learning," in *International Conference on Information Technology and Multimedia*, pp. 1–8, 2011.
- [26] O. Aydemir and T. Kayikcioglu, "Decision tree structure based classification of EEG signals recorded during two dimensional cursor movement imagery," *J Neurosci Methods*, vol. 229, pp. 68–75, May 30 2014.
- [27] M. F. Krummel and I. Macara, "Maintenance and modulation of T cell polarity," *Nat Immunol*, vol. 7, pp. 1143–9, Nov 2006.
- [28] R. A. Khorshed and C. Lo Celso, "Machine learning classification of complex vasculature structures from *in-vivo* bone marrow 3d data," (*ISBI*), *IEEE International Symposium on Biomedical Imaging*, Prague, 2016.
- [29] A. Mendelson and P. S. Frenette, "haematopoietic stem cell niche maintenance during homeostasis and regeneration," *Nat Med*, vol. 20, pp. 833–846, 08//print 2014.
- [30] Daniel J. Nolan, M. Ginsberg, E. Israely, B. Palikuqi, M. G. Poulos, D. James, et al., "Molecular signatures of tissue-specific microvascular endothelial cell heterogeneity in organ maintenance and regeneration," *Dev Cell*, vol. 26, pp. 204–219, 29 July 2013.