

Imperial College London

MRes Project 2 Dissertation

(Standard project)

Combining machine learning algorithms for cell segmentation of intravital microscopy video data to study neutrophil behaviour in the spleen

Mingze Gao

CID 01737263

Supervisors:

Dr. Joram M. Posma

Department of Metabolism, Digestion and Reproduction

Faculty of Medicine, Imperial College London

Dr. Katia De Filippo

National Heart & Lung Institute

Faculty of Medicine, Imperial College London

This research report is submitted in partial fulfilment of the requirements for the degree programme of MRes in Biomedical Research (Data Science stream).

August 17, 2020

Statement of Originality

I hereby certify that this thesis and the research to which it refers are the product of my own work, conducted during *Apr. 2020 ~ Aug. 2020* of MRes in Biomedical Research (Data Science Stream) at Imperial College London. Any ideas or quotations from the work of other people, published or otherwise, or from my own previous work are fully acknowledged in accordance with the standard referencing practising of the discipline.

Acknowledgements

I would like to thank my supervisors, Dr. Joram Matthias Posma and Dr. Katia De Filippo, for their continued support from informatics, statistical analysis and biological experiments aspect. Their guidance and feedback in the past few months were invaluable in achieving the project's aims and objectives. Gratitude also goes to Dr. Katia De Filippo for providing intravital microscopy dataset in both video and structured data format.

Abbreviations

CD11a	Integrin Alpha L, one component of Lymphocyte Function-associated Antigen 1		
fastER	Fast Segmentation with Extremal Regions		
Mask-RCNN	Mask Region-based Convolutional Neural Network		
BM	Bone Marrow	ML	Machine Learning
DL	Deep Learning	MOT	Multiple Object Tracking
GPU	Graphics Processing Unit	RoI	Region of Interest
IoU	Intersection over Union	RPN	Region Proposal Network
IVM	Intravital Microscopy	SVM	Support Vector Machine

Abstract

Background: Neutrophil plays an important role in the immune system and their immune response to infections in the spleen were described recently. However, the detailed migration information of neutrophils and their interaction with other splenic cells have not been fully elucidated. This project aims to use machine learning (ML) and deep learning (DL) algorithms to analyse intravital microscopy (IVM) video data, then investigate neutrophil's migration and interaction behaviour.

Methods: Two image segmentation algorithms fastER and Mask-RCNN are applied on IVM. Results are filtered, merged and the performance is compared with 'Imaris', a widely used commercial software designed to analyse microscopy data. Migration and interaction details are then defined, quantified and analysed using statistical metrics based on the trajectory of each monitored neutrophil, both before and after the injection of CD11a monoclonal antibody.

Results: The performance of ML/DL algorithms outperform Imaris in number of neutrophils detected. Migration results exhibit region heterogeneity as neutrophils' behaviour from one part of the spleen changed significantly after treatment whilst for another part, there was no statistical difference. We also observed that more neutrophil-B cell interactions took place before the treatment than after. However, the duration for a single interaction has not been affected.

Discussion: This research work is a proof of concept that with ML/DL algorithms, more cells could be identified and more accurate data could be provided than Imaris. The effect of blocking on neutrophils might be region-dependent thus it is essential to know which area in the spleen the experiments were conducted. We attempted to study contact-dependent interactions between splenic cells which is beneficial in understanding immune responses and cellular communications further. However, future works are needed to refine the multiple object tracking algorithm and revise the features as well as thresholds used in this project because some are arbitrarily defined.

Availability: Python code, access via <https://github.com/MichaelGMZ/MRes-2020DS-Project2>

Contents

1 Introduction	4
1.1 Neutrophil function & cell motility	4
1.2 Neutrophil in the spleen and interactions with B cell	4
1.3 Intravital microscopy	5
1.4 Image data processing	6
1.4.1 Object segmentation algorithms	6
1.4.2 Object tracking techniques	7
1.5 Aims, Hypothesis & Objectives	8
2 Materials & Methods	9
2.1 IVM Datasets	9
2.2 IVM video data segmentation	9
2.2.1 fastER	9
2.2.2 Mask-RCNN	11
2.3 Hierarchical clustering	13
3 Algorithm Development	14
3.1 Neutrophil segmentation	14
3.2 Multiple neutrophils' motion tracking	15
3.3 Migration feature extraction	16
3.4 Definition of interaction	17
4 Results	19
4.1 Methods performance comparison	19
4.2 Neutrophil migration analysis	20
4.3 Neutrophil-B cell interaction	24
5 Discussion	26
References	29

1 Introduction

1.1 Neutrophil function & cell motility

Neutrophils are the most abundant leukocyte circulating in blood with an estimation of 10^7 neutrophils in mice and 10^{11} in humans, which account for approximately 60% of white blood cells [1]. Neutrophils are released daily from the bone marrow (BM) and they play a vital role in acute inflammation, chronic disease and provide the first line of defense against pathogens with multiple intra- and extracellular killing mechanisms [2].

The response to infection and neutrophil recruitment are highly dependent on the migration ability of these innate immune cells to affected tissues [3]. Various adhesion molecules are expressed on the surface of immune cells, namely cluster of differentiation (CD), which are crucial for cells' life cycle. Among them, integrins are essential for neutrophils' signalling, adhesion and migration ability [4]. Integrins are formed by two subunits, α chain and β chain, the latter one can be classified further to three subfamilies, β_1 , β_2 and β_3 . Specifically, CD11a (Integrin Alpha L) together with β_2 subunit form the lymphocyte function-associated antigen-1 (LFA-1) that plays a key role when leukocyte and lymphocyte migrate from bloodstream to tissues [5]. Previous research has shown that the adhesion molecules like CD11a is used by monocytes for migration across endothelium [6]. However, when and how CD11a sustain cell migration within the tissue is still unclear, as sometimes lymphocyte migrate also in the absence of CD11a [7] which suggests cell migration is a complex function that might not be controlled, activated and affected by a single molecule player.

1.2 Neutrophil in the spleen and interactions with B cell

Neutrophils are produced and retained within the bone marrow. Together with the BM, the spleen is one of the major reservoirs of mature neutrophils. The spleen is an immunological site with a morphology similar to a large lymph node. Among the most common functions, the spleen filters blood, cleans blood-borne pathogens as well as aged erythrocytes and it acts as the primary site of adaptive immune response [8]. Produced in bone marrow, B cells are also part of the immune system and they

can migrate to spleen and other lymphoid tissues where they eventually become mature. B cells play a vital role in immunopathogenesis and secrete antibodies that contribute to tissue injury via multiple mechanisms [9].

Recent literature has elucidated that splenic neutrophils could generate immunoglobulin defense through the activation of B cells [10], such as somatic hypermutation and antibody production. This kind of involvement is shown by biomedical mechanism such as expression of regulatory molecules or cytokines [10]. Such involvement could be treated as the proof of concept that neutrophil-B cell interaction exists in innate immune system.

However, quantified migration behaviour of splenic neutrophils in the absence of CD11a and neutrophil recruitment has not been fully recognised, with little work emphasised on the neutrophil-B cell interactions especially the duration of interaction to our best knowledge [11]. Gaining information in this research area is essential to understand immune mechanisms further and how internal/external factors may regulate splenic immune cells' activation, interaction and migration during infections.

1.3 Intravital microscopy

Intravital microscopy (IVM) is a sub-type of microscopy which allows the observation of biological processes in deeply anaesthetised live animals at a high resolution [12]. Due to the immense physiological complexity of tissues and organs, the characteristics of experimental subjects might not be accurately captured *in vitro*. Therefore, IVM could be the ideal tool to study the dynamic process of cell motility within tissues and monitor how the cell behaviours change during treatment. Recently, IVM has been applied to study the process of hematopoietic stem cells migrate from the BM to the spleen during leukaemia [13]. The same technique has allowed to identify that 2 populations, immature ($Ly6G^{low}$) and mature ($Ly6G^{high}$) neutrophils within the spleen exhibit different migratory behaviours [11].

Commercial software **Imaris v8.3** (Bitplane AG) is widely used for automated cell analysis from microscopy data. However, a recent study showed that Imaris suffers from high false positive rate occasionally and tends to be more conservative regarding poorly detected cells (cell detected, but

not resolved as a single object or merged with another cell) [14]. Therefore, other segmentation algorithms could be implemented in cell detection stage if they have better performance.

1.4 Image data processing

1.4.1 Object segmentation algorithms

Object segmentation is a technique related to image processing and computer vision that deals with detecting instances of semantic objects of certain classes [15]. It can be applied to various fields, such as autonomous driving, video surveillance, object tracking and cell recognition.

Algorithms for object segmentation and detection can generally be classified as either (traditional) machine learning (ML) or deep learning (DL) based approaches. For the former, features are identified first, then classification methods are applied to identify target regions. On the other hand, DL methods are an end-to-end solution without the need to pre-define the features and are typically based on convolutional neural network (CNN) architectures.

Pattern recognition techniques can be used to perform image segmentation, such as Support Vector Machines (SVM) [16]. This algorithm tries to find a hyperplane that best separates two classes. It can minimise the empirical classification error and maximise external margin between two classes simultaneously. The user-friendly cell segmentation tool "fastER" [17], used in the preliminary stage of this project (see Section 2.2.1), is based on SVM. Beside classification based techniques, algorithms based on thresholds can perform same task by finding the edge pixels and eliminate the noise. This type of methods comprises edge-based algorithms that related to the edge information and its structures as well as regional-based algorithms using the concept that quantifiable features inside a structure tend to be homogeneous [18] compared with external features.

For DL algorithms, they generally follow the similar concept as traditional methods in which the frameworks can be categorised into two types, whilst the first regards object detection as a regression or classification problem and, the second is based on region proposal generation [19]. However, with neural networks as its internal structure, it can outperform ML methods in several fields. Recent research has shown a CNN based classifier achieved promising results for the segmentation of liver

images [20]. Furthermore, U-net was devised as an extension of CNN which has achieved satisfactory performance on different biomedical segmentation applications [21].

1.4.2 Object tracking techniques

Object tracking is a subsequent procedure of object detection, which can be defined as the process of identifying the location of the moving object (a cell for example) and the ability to determine whether it is the same object in previous and subsequent frames. Existing tracking algorithm can be mainly classified as three categories: filtering-based sampling approaches, model-based evolution approaches and detection based association approaches [22].

Kalman filter is a filtering-based sampling approach which utilises a series of observed measurements to estimate unknown variables over frames [23]. Through the iterative variable prediction and update step the algorithm can be applied in real time and the estimation of unknown variable becomes more and more accurate compared with real value. Among the model-based evolution approaches, mean shift tracking is a probabilistic approach and creates a confidence map based on the information of the tracking object from previous frames. This algorithm works well if the features used to define multiple mean-shift kernels is known and the velocity of moving object is slow [24].

The last approach segments and locates objects first and then uses data association methods to match the tracklets (a fragment of the trajectory followed by a moving object) with correct tracking identification number, named "tracking by detection". Data association problem refers to matching correct regions of interest in a series of frames. There are two basic ways of finding correspondent region of two subsequent frames. One is centroid based tracking method which utilises the Hungarian algorithm to assign labels and minimises the Euclidean distance of pairwise objects, the other one is an intersection over union (IoU) based tracking method which choose the maximum IoU of pairwise objects of two consecutive frames to reconstruct associated tracklets.

1.5 Aims, Hypothesis & Objectives

The overarching aim of this project is to implement ML/DL based algorithms to analyse splenic IVM video data and use statistical metrics to quantify the migration characteristics of neutrophils and neutrophil-B cell interaction behaviours under different experimental conditions.

In light of previous research results in literature [25, 26, 27], the first hypothesis is that the ability of neutrophils to migrate within the spleen is affected by the Integrin LFA-1 (CD11a). Due to the fact that neutrophil can switch phenotype accordingly and generate various sub-populations with unique functions [11, 27], the second hypothesis is that neutrophil sub-populations respond differently during treatment.

Specifically, the objectives of this project are to:

- (a) Detect neutrophils on raw IVM video data using ML and DL cell segmentation methods.
- (b) Compare the performance of ML/DL algorithms with commercial software Imaris.
- (c) Devise a multiple object tracking (MOT) algorithm to track the trajectory of detected neutrophils.
- (d) Analyse the structured data regarding the migratory behaviours of splenic neutrophils and their sub-populations, both before and after CD11a blocking.
- (e) Investigate splenic neutrophils' interaction with B cells and effect on neutrophils' behaviour.

2 Materials & Methods

2.1 IVM Datasets

The spleens of live mice were imaged with IVM making use of fluorescent labelling monoclonal antibody (mAb) Ly6G that recognise neutrophils and mAb B220 that recognise B cells. We have six datasets in total, shown in Table 2.1. The main focus will be analysing cell behaviours using 'CD11a' related datasets. It was previously observed that the ratio of neutrophils to B-cells increased after the injection of FTY720 and it is used here as additional test sets to evaluate the performance and generalisability of the cell detection model.

Table 2.1: "CD11a" and "FTY720" corresponding to spleens of mice are injected intravenously with CD11a monoclonal antibody and immunomodulating drug Fingolimod. Imaris file column means whether Imaris detection results exist, this can be used to compare performance with ML/DL methods. Region column indicates whether experiments are conducted at same part of spleen.

Index	Treatment	Imaris File	Video Duration	Region
1	Before injection of CD11a antibody	+	6s	1
2	After injection of CD11a antibody	+	6s	1
3	Before injection of CD11a antibody	-	14s	2
4	After injection of CD11a antibody	-	12s	2
5	Before injection of FTY720	+	6s	3
6	After injection of FTY720	+	6s	3

2.2 IVM video data segmentation

In this project, two state-of-art segmentation methods fastER [17] and Mask-RCNN [28] were applied on IVM video data directly. This approach has two primary benefits, (1) compared with commercial software Imaris, not only the location of cell centroid but also morphology features can be captured, (2) two methods can be complementary to each other to make the best of segmentation results as either one has their own advantages and drawbacks.

2.2.1 fastER

fastER is a cell segmentation tool with robust and computationally friendly performance developed for microscopy images - where 'ER' stands for Extremal Regions, a tree of nested candidate regions.

fastER supports various cell types with no extra parameters to be tuned, thus this is an easy-to-use software even for non-experts [17]. The internal and fundamental algorithm used by fastER to estimate the likelihood of each candidate region to be a cell is SVM, a discriminative classifier perfectly for two-class problem (image background and cells in our experiment).

SVM is a supervised learning algorithm that is mainly used in classification. Supervised learning means training data are labelled with proper class in advance, then based on selected features of training data, SVM builds a model to assign new data to a corresponding category. SVM can solve linear and non-linear cases by spatial transformation (low-dimensional to high-dimensional space), then kernel functions are used to reduce the computation costs of high-dimensional calculation and map the result into low dimensional space.

Essentially, it tries to use limited support vector (sample data) to find a decision hyper-plane that can separate two classes with maximum margin. As illustrated in Fig. 2.1(A), $w \cdot x - b = 0$ represents the optimal hyper-plane where w is normal vector, b is intercept and $\frac{2}{\|w\|}$ is the objective function that needs to be maximised. This can be converted to a constrained optimisation problem as follows,

$$\min_{w,b} \frac{1}{2} |w^2| \quad s.t. \quad y_i(w \cdot x_i + b - 1) \geq 0, i = 1, 2, \dots, N \quad (2.1)$$

where y_i represents the category of x_i . To solve this equation, generally a Lagrange multiplier α_i is introduced to create a Lagrange function and transform the optimisation function into,

$$\min_{\alpha} \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) - \sum_{i=1}^N \alpha_i \quad s.t. \quad \sum_{i=1}^N \alpha_i y_i = 0, \alpha_i \geq 0, i = 1, 2, \dots, N \quad (2.2)$$

In fastER interface, the user should define positive and negative sample regions from microscopy image that best fit cell and background label. Then a feature vector which contains nine features of a candidate region is extracted by fastER to build the model, including volume, axis length, average intensity, average gradient, eccentricity etc. This feature vector is the input of SVM with Gaussian kernel model used to train relevant parameters. In the end testing data-sets are provided for prediction purpose, a segmentation example from our dataset is provided in Fig. 2.1(B).

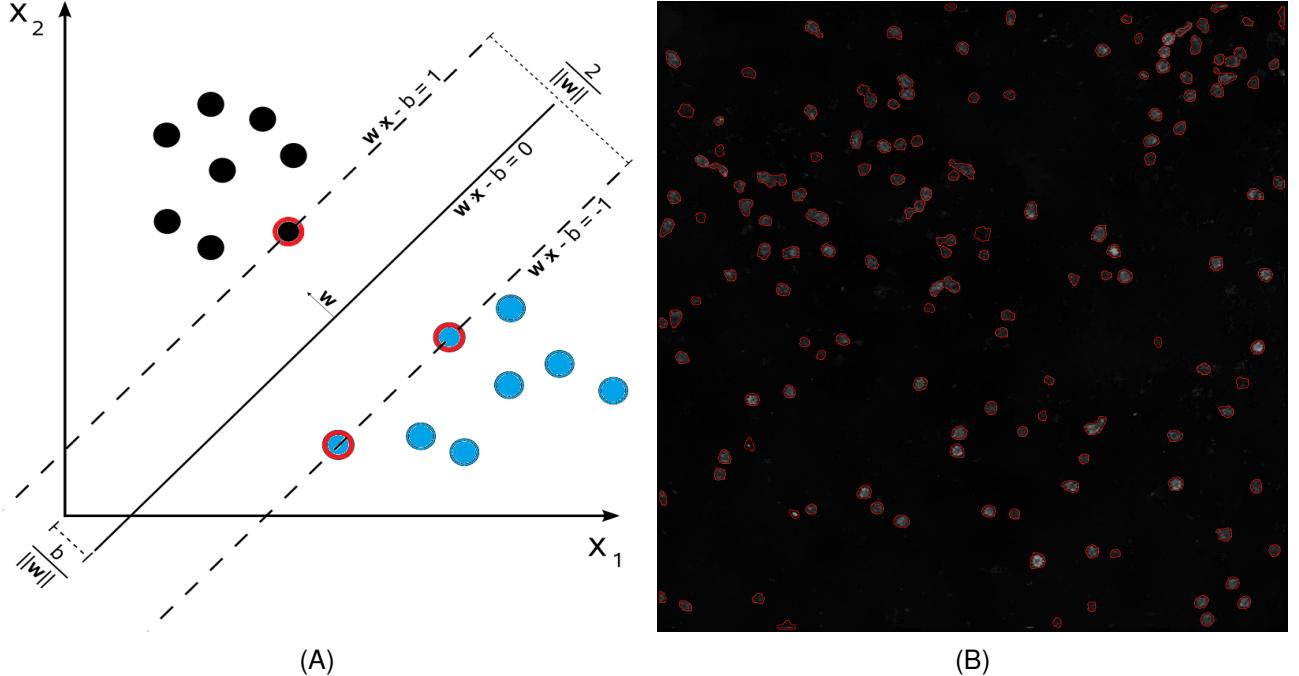


Figure 2.1: (A) Illustration of SVM maximum separation hyper-plane and margin between two classes. Data with red contour are support vectors. Axes represent feature vector. (B) fastER segmentation result depicted by red contour on 1st frame of dataset 3.

Since the intrinsic methodology of fastER is a non-deep learning method, it does not require extra hardware such as GPU and therefore reduces training time significantly. A graphical user interface makes it user-friendly and enables interactive feedback in real time with the possibility of adding new labels if segmentation errors occur or monitoring segmentation quality as labels increased.

2.2.2 Mask-RCNN

Mask-RCNN is a Region-based Convolutional Neural Network instance segmentation algorithm developed by He *et al.* in 2017 [28]. Instance segmentation will further assign different instances to individual objects based on the outcome of semantic segmentation. With a satisfying outcome, this algorithm has been applied into various fields, such as autonomous driving and human pose estimation [28]. This algorithm is built on top of its previous version, Faster-RCNN and an object mask will be returned in addition to bounding box and class label (Fig. 2.2(A)). At a high level, the network structure contains the following modules, Backbone Model, Region Proposal Network (RPN), Region of Interest (RoI) classifier and Segmentation Masks.

Backbone model is a network architecture used to extract features from images and Mask-RCNN uses ResNet101. Early layers of backbone detect low level features such as edges and corners

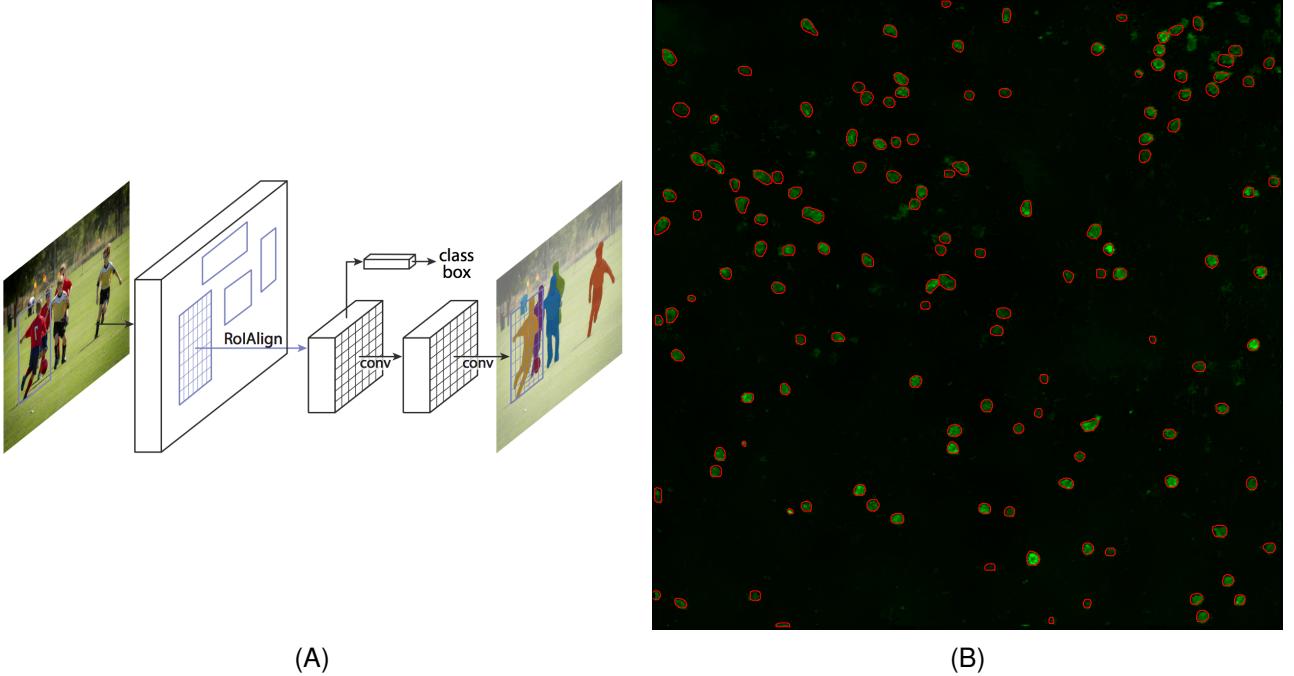


Figure 2.2: (A) Mask-RCNN framework. Source: <https://arxiv.org/abs/1703.06870> (B) Mask-RCNN instance segmentation result depicted by red contour on 1st frame of dataset 3.

whilst later layers extract high level features such cars and cells. Then the obtained feature map in the previous step is applied to the Region Proposal Network (RPN), which scans the image using a sliding window and predicts if an object can be identified in that region. In practice, there will be thousands of scanned regions with different sizes and aspect ratios to cover as much part of the image as possible. The output of RPN will be the class of that region (background or foreground) and a refined bounding box for the detected object. Regions obtained from RPN could have different sizes, thus a pooling layer is applied to all the regions to convert them to the same shape, this process is called RoIAlign and uses bi-linear interpolation. Only regions have an IoU (*Area of the intersection / Area of the Union*) with the ground truth greater than a certain threshold can be retained. After that RoI classifier will predict the actual bounding boxes and specific class label of the object (cells, cars, ...). Mask-RCNN also generates segmentation masks which are soft masks with 28×28 resolution. It contains more information than binary mask and can be scaled up in inference mode.

An example of Mask-RCNN instance segmentation result is provided in Fig. 2.2(B). For neural network based algorithms, the performance is highly depend on the quality and annotations of training set. It also requires parameter tuning therefore it is not as user-friendly as fastER. However, for

Mask-RCNN the format of input image can be RGB colour mode rather than a grey-scale image as fastER requires and the segmentation output has less merge errors (masks contains more than one centroid of ground truth). Which method to choose for optimal segmentation results should base on the purpose of segmentation and image acquisition type [17].

2.3 Hierarchical clustering

Since data are non-normally distributed, hierarchical clustering is chosen to identify sub-populations of neutrophils based on their migration behaviours. Hierarchical clustering is one of the unsupervised learning algorithms and thus could be applied on unlabelled data points. It has been implemented extensively in various fields, such as gene expression analysis and flow cytometry.

There are two types of hierarchical clustering: ‘Agglomerative’ uses bottom-up approach which treats each individual data as one cluster and form bigger cluster by joining two closest clusters based on certain distance metrics until one cluster includes all data points, whilst ‘Divisive’ uses top-down approach [29]. In this research, we will focus on agglomerative clustering and use dendograms to illustrate and determine optimal number of clusters.

To merge two clusters, Ward’s method (minimum variance criterion) is used to measure the merging cost between clusters A, B and the lowest pair of clusters will be merged,

$$\Delta(A, B) = \sum_{i \in A \cup B} \|\vec{x}_i - \vec{m}_{A \cup B}\|^2 - \sum_{i \in A} \|\vec{x}_i - \vec{m}_A\|^2 - \sum_{i \in B} \|\vec{x}_i - \vec{m}_B\|^2 = \frac{n_A n_B}{n_A + n_B} \|\vec{m}_A - \vec{m}_B\|^2 \quad (2.3)$$

where \vec{m}_i is centre of cluster i , \vec{x}_i is each data point, n_i is the number of points in it and Δ is merging cost to combine cluster A and B.

The advantage of dendrogram is that it provides clear relationship between each individual data point and their high-level clusters, however, this technique does not explicitly show the optimal number of clusters for the task. Therefore, based on the dissimilarities between clusters, a distance value is picked manually and that will yield an appropriate number of clusters.

3 Algorithm Development

A graphical representation of algorithm development pipeline of this study is shown in Fig. 3.1. This pipeline comprises three modules, cell extraction, cell segmentation and trajectory tracking.

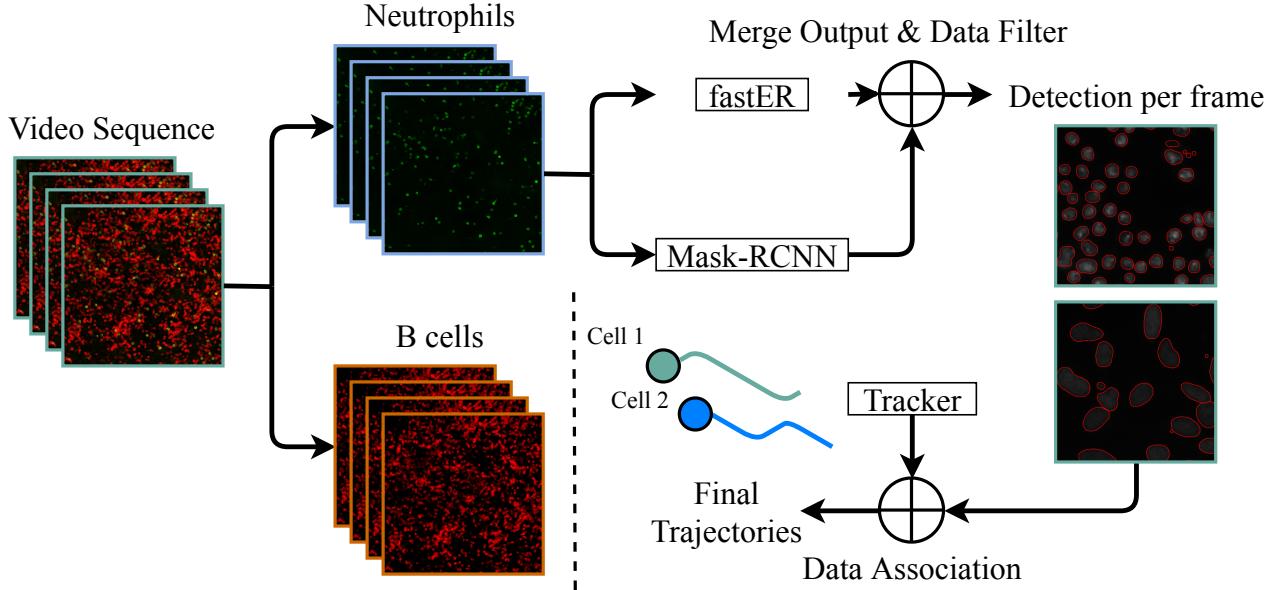


Figure 3.1: Cell extraction and tracking by detection paradigm. Firstly, frames only contain neutrophils or B cells are extracted from IVM video data based on different fluorescent colours. Secondly, two independent object detector fastER and Mask-RCNN are applied to all image frames and results are merged to boost segmentation performance. Then a tracker is run on cells with data association techniques embedded. Lastly, cell trajectories are created which enables downstream analysis of migration behaviour and cell interaction phenomenon.

3.1 Neutrophil segmentation

Cell segmentation results from fastER and Mask-RCNN are combined to increase true positive as well as true negative rate, as one objective of this research is to analyse migration behaviours based on trajectory of individual cells, the segmentation result should be as accurate as possible. Therefore, we do not consider the time constraints problem in this project and both fastER and Mask-RCNN are used to segment IVM data.

Original IVM video data has a duration of 6s, 12s and 14s. With parameter "*frame per second (fps)*" equals to 10, we extracted 60, 120 and 140 microscopy frames (each frame represents 1s in reality) for corresponding video containing only neutrophils based on fluorescent colour.

For fastER segmentation, 5 frames are randomly chosen from each dataset as training set, and the

entire dataset itself is the test set. In each frame of training set, 10 neutrophils are manually labelled as positive region and as we monitor the segmentation performance during neutrophil labelling stage, we choose to label parts of the background as a negative region if necessary. Thus, the number of negative labels can be higher or lower than the number of positive labels. Then the segmentation masks are generated automatically by batch processing.

For Mask-RCNN segmentation, the training model is essential as it is a deep learning method. Hyper-parameters in the model contain various features of images and objects. The trained model and original code we used is downloaded from https://github.com/mirzaevnom/data_science_bowl_2018. The model is trained by Inom Mirzaev on 2018 Data Science Bowl dataset. Parameters are fine tuned for our IVM dataset and used ensemble prediction in inference mode. As it achieved comparable results with Imaris, which shows the model's generalisation ability to new datasets, labelling and training a separate model based on our IVM data is not currently necessary.

Then the false positives (detected cells containing no centroids of the ground truth cells) and ambiguous cells (detected objects with a tiny volume or a disorganised contour) are filtered at this stage based on intensity, volume within contour and smoothness of contour. The union of two segmentation outputs is labelled as method 'FM' (fastER & Mask-RCNN combined) and used in analysis as the accuracy is higher than each individual method.

3.2 Multiple neutrophils' motion tracking

The second part of the pipeline is tracking by detection, which uses tracker to associate same objects among different frames, then the trajectory of each individual cell can be finalised. Based on SORT (Simple Online and Real-time Tracking) [30] which is a multiple object tracking (MOT) algorithm, we introduced the following modifications for our research purpose,

- Use distance matrix between centroids instead of IoU matrix between bounding boxes as small objects with high velocity will decrease the effect of IoU.
- Introduce a threshold in data association problem under assumption that distance of each cell travels between frames is unlikely to be extremely far.

- Allow the cells to have multiple IDs when merge-split problem occur, this will assist cell re-identification after they separate in further research.
- Features of each cell on individual frame are captured and stored as structured data format for downstream analysis and a tracking video is generated locally.

As there are no perfect configuration for every experiment, parameters need to be tuned accordingly.

We set *minimum hit frames* = 3 (only objects detected for at least 3 frames are considered as true cells) and *maximum disappeared frames* = 10 (cells are considered disappeared permanently if no detection occurred within 10 consecutive frames).

3.3 Migration feature extraction

Segmentation masks generated in Section 3.1 and motion tracking video in Section 3.2 can be used to extract features of cell. We extracted the following features (Table 3.1) which are used for downstream migration and interaction analysis,

Table 3.1: Six extracted features and their relevant definitions.

Feature	Description
Intensity	Average fluorescence intensity of each pixel inside the contour
Volume	Number of pixels inside the contour
Smooth	First derivative of distance between centroid and periphery pixels
Circumference	Length of the contour
Velocity	Velocity between consecutive frames
Acceleration	Acceleration between consecutive frames

Velocity and acceleration are direct indication of migration behaviours whilst volume, smooth, circumference and intensity are morphology features which might serve as indirect indication of migration behaviours or reflect different neutrophil sub-populations. Example of segmentation masks are illustrated in Fig. 3.2(A) - 3.2(D) from which first four features are extracted.

We measure smoothness under the assumption that cells with a perfectly round shape are most likely to stay still whilst a migrating cell is likely to have an oval or irregular shape. Thus, higher smoothness level means higher migration possibility and vice versa. To calculate smoothness level, distance between centroid and periphery pixels are measured (Fig. 3.2(E)) followed by first derivative

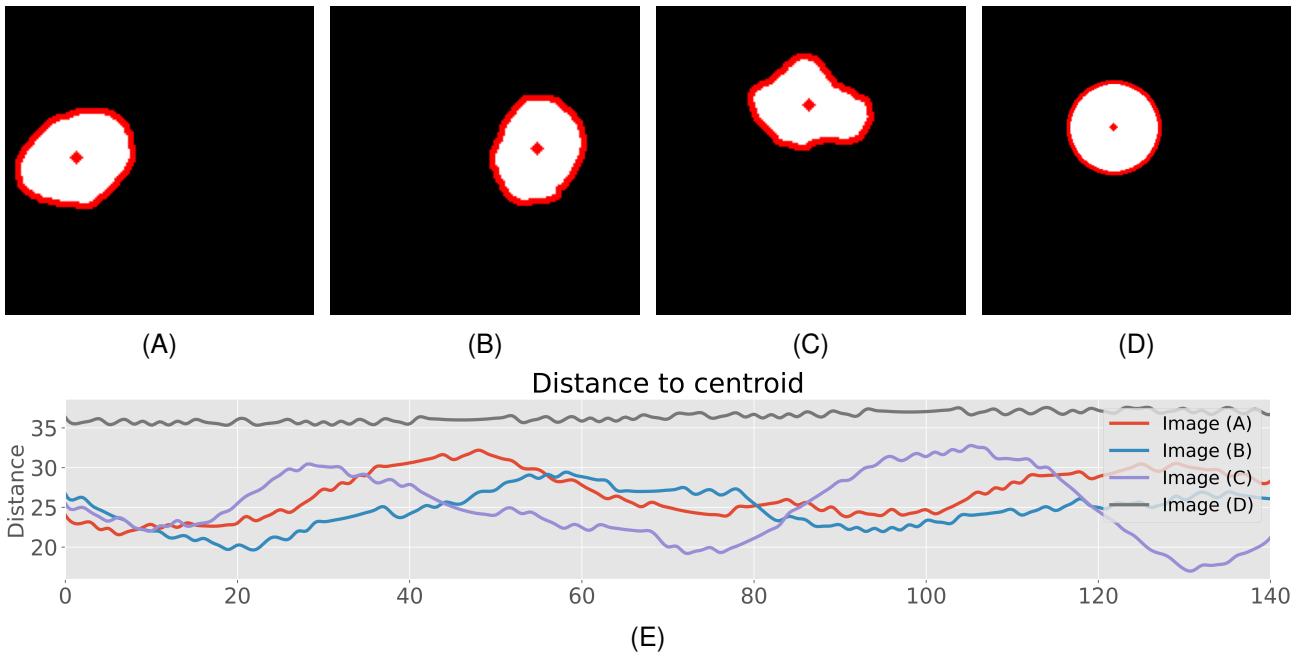


Figure 3.2: Example of cell segmentation masks, contours, centroids and distance of each peripheral point to centroid. (A) (B) An oval shape cell. (C) An irregular shape cell. (D) A theoretical round shape cell. (E) The variance of distance and its first derivative could be used to measure smoothness or roughness of cell and subsequently linked to migration behaviour. For instance, the above four cells have smoothness level of 0.311, 0.319, 0.446, 0.223 respectively.

of distance.

Velocity and acceleration of first frame are assumed to be 0 and that of subsequent frames are calculated using formula below,

$$V(n) = \sqrt{\frac{(x_n - x_m)^2 + (y_n - y_m)^2}{T_n - T_m}} \quad (3.1)$$

$$A(n) = \frac{V_n - V_m}{T_n - T_m} \quad (3.2)$$

where n, m are indexes of two frames, x, y are centroid of the cell and T is time point. All features are then scaled from pixel level to μm level with a scale factor 0.481 for interpretation.

3.4 Definition of interaction

A contact dependent interaction between neutrophils and B cells in this study is defined based on the proximity of two cells and their absolute velocity. The assumption is that such an interaction is most likely to occur when two or multiple cell contours have intersection with each other and their velocity are relatively lower than certain threshold.

In this respect, for each neutrophil at every frame, the intersection condition to every B cell in same frame is calculated. Compared with method using closest distance as metric which assumes one neutrophil can only interact with one B cell at the same time, this allows multiple interactions to happen simultaneously. The velocity of each neutrophil and B cell at every frame is also recorded and fitted by a Gamma distribution (Fig. 3.3). It was difficult to determine the precise threshold to classify

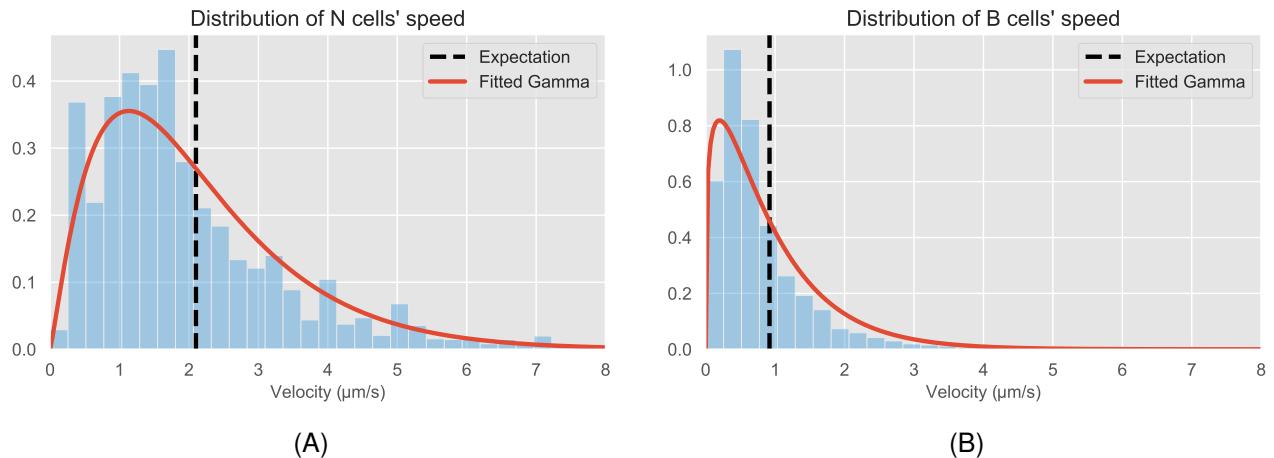


Figure 3.3: (A) Gamma distribution of neutrophils' speed, *Expectation* = 2.101. (B) Gamma distribution of B cells' speed, *Expectation* = 0.919.

cells into low-velocity group (possible for interaction to occur) and high-velocity group (activated cells tend to migrate). Expectation ($E[X] = \int_{-\infty}^{\infty} x \cdot f(x)dx$) of the fitted Gamma distribution is used here as threshold because expectation could divide a probability distribution evenly. Therefore, the maximum acceptable velocity for a neutrophil to interact is $2.101\mu\text{m/s}$ and $0.919\mu\text{m/s}$ for B cells.

This project mainly studies the interaction duration and frequency between neutrophil and B cell, specifically we attempt to answer the following questions,

- Is there any significant difference of interaction duration for a specific neutrophil across all frames before and after treatment?
- What is the distribution of interaction duration between a pair of neutrophil and B cell?
- Is there a difference in the number of interactions a neutrophil has depending on which class it derived from hierarchical clustering?

4 Results

4.1 Methods performance comparison

The first step in comparing the FM cell recognition with Imaris is to compare the average number of neutrophils detected per frame by each algorithm, shown in Table 4.1. 'Imaris-Only' and 'FM-Only' are neutrophils detected only by Imaris or FM, whilst 'Matched' are cells detected by both methods. Before directly comparing the performance of Imaris and FM, first cells assigned to 'Only' category need to be investigated whether they are false positives or not using the features extracted as metrics. The distributions of the features are compared with that of the 'Matched' case as shown in Fig. 4.1, since we do not have manual labelled ground truth (Imaris exports fixed morphological features, therefore only intensity can be compared here).

Table 4.1: Number of neutrophils detected on relevant dataset. FM: fastER and Mask-RCNN combined. Matched: detected by both Imaris and FM. Imaris-Only: only detected by Imaris. FM-Only: only detected by FM. A (B - C): median value across all frames (25% quantile - 75% quantile).

Dataset	FM	Imaris	Matched	Imaris-Only	FM-Only
1	163 (159 - 168)	174 (170 - 178)	140 (136 - 145)	34 (31 - 38)	22 (18 - 24)
2	137 (130 - 145)	123 (117 - 130)	108 (102 - 114)	15 (13 - 18)	29 (26 - 33)
5	240 (233 - 247)	145 (133 - 157)	138 (130 - 146)	8 (2 - 14)	102 (97 - 107)
6	473 (457 - 490)	247 (229 - 266)	224 (208 - 240)	25 (21 - 29)	249 (243 - 255)

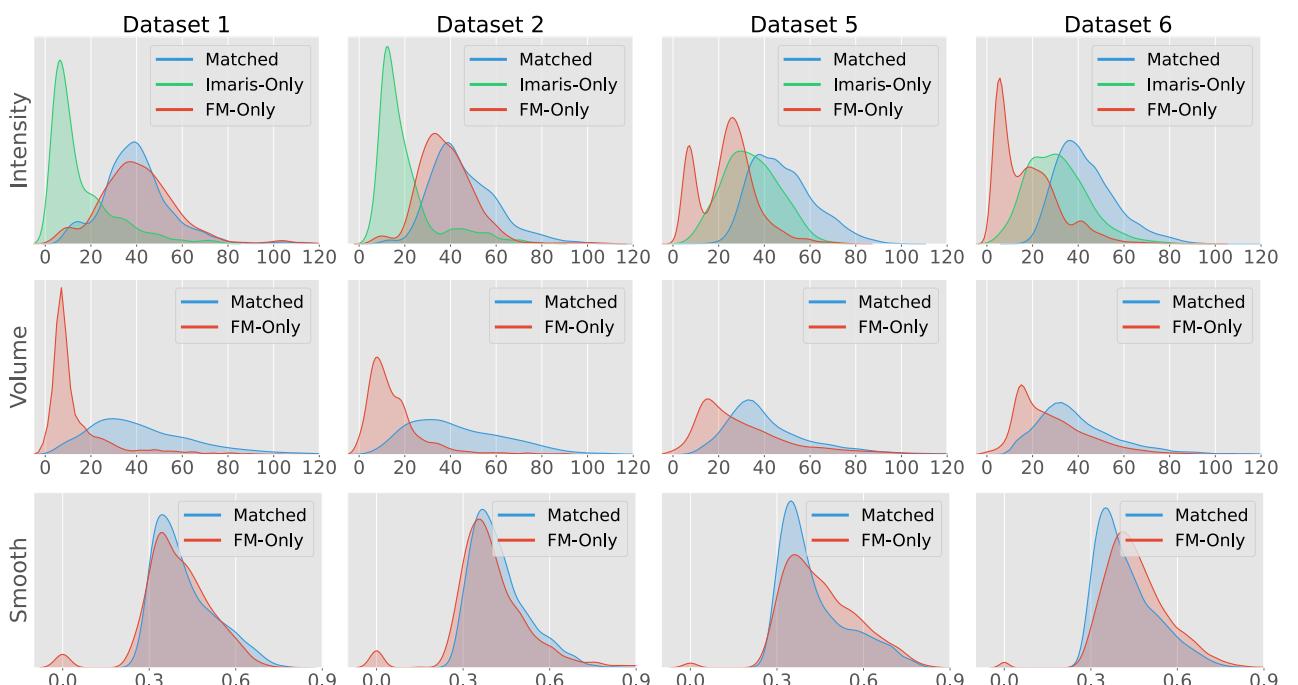


Figure 4.1: Distribution of features in different scenarios.

For dataset 1&2, the distribution of *FM-Only* is in accordance with *Matched* in terms of intensity and smoothness. The peak located at the leftmost of the volume plot might indicate *FM* has the potential to detect small volume cells as *Matched* also covers this range. *Imaris-Only* generally have much lower intensity compared with *Matched*, with no other features to be compared, we cannot conclude whether they are true positives or false positives. For dataset 5&6, generally the distribution of *FM-Only* and *Matched* in volume and smooth are homogeneous. In terms of intensity comparison, the distribution of both *Imaris-Only* and *FM-Only* are shifted left. However, with a large overlapping region on *Matched*, we tend to conclude they are true positives.

A McNemar's test shows that there is no statistical significance between the performance of Imaris and FM in dataset 1 for the number of neutrophils detected ($p = .140$) which means no method is better than the other one. However, for the other three datasets, there are significant differences ($p = .048, < .001, < .001$ for dataset 2, 5, and 6 respectively) where FM finds significantly more neutrophils than Imaris. Combined with the distribution of features in Fig. 4.1, it shows that generally ML&DL algorithms outperform commercial software Imaris with higher true positives and lower false positives regarding cell detection performance. To promote the accuracy of downstream migration and interaction analysis, detection results from FM are used instead of Imaris.

4.2 Neutrophil migration analysis

In order to investigate neutrophils' migration behaviour in spleen before and after intravenous injection of CD11a blocking antibody, dataset 1~4 are used during the experiment whilst 1&2 and 3&4 are conducted in different regions of spleen and used as self-control. Using pipeline described in Fig. 3.1, the trajectory of neutrophils during each experiment can be acquired and migration features described in Table 3.1 can be extracted afterwards. Here two MOT outputs are provided,

- Tracking result of splenic neutrophils' trajectory of dataset 3 (video) can be accessed [here](#).
- Tracking result of splenic neutrophils' trajectory of dataset 4 (video) can be accessed [here](#).

Each neutrophil is indicated by their associated labels and a red bounding box contains its geometry

features. Though a bounding box is used here to locate cells, all the calculations embedded within the pipeline are based on cell contours which are more accurate than bounding boxes in terms of morphological features, the pipeline can visualise either format at user's discretion.

As neutrophils migrate, the 'merge-split' phenomenon might occur. It refers to two or more neutrophils clustering together thus only one cell is recognised by detection algorithms and when those cells separate again, current data association methods based on centroids cannot achieve a satisfactory performance to re-assign correct labels accordingly. Therefore, cells with 'merge-split' circumstances took place were excluded from the input of downstream migration analysis (Table 4.2).

Table 4.2: Total number of neutrophils tracked and number of neutrophils used for migration analysis for each relevant dataset.

Dataset Index	1	2	3	4
Total # of tracked neutrophils	288	241	307	385
# of neutrophils used for analysis	197	181	239	225

Based on the definitions in Table 3.1, six features including intensity, volume, smooth, circumference, velocity and acceleration are extracted and their distribution is plotted accordingly. Here we use dataset 3&4 as example and the relationship between selected features is shown in supplementary Fig. S1. Then a linear regression is performed between velocity and acceleration, circumference and volume respectively. With a high R^2 indicating linearity, acceleration and circumference are shown in separate graphs (supplementary Fig. S2), as this association between features indicates high correlations (with non-parametric tests this becomes even more apparent).

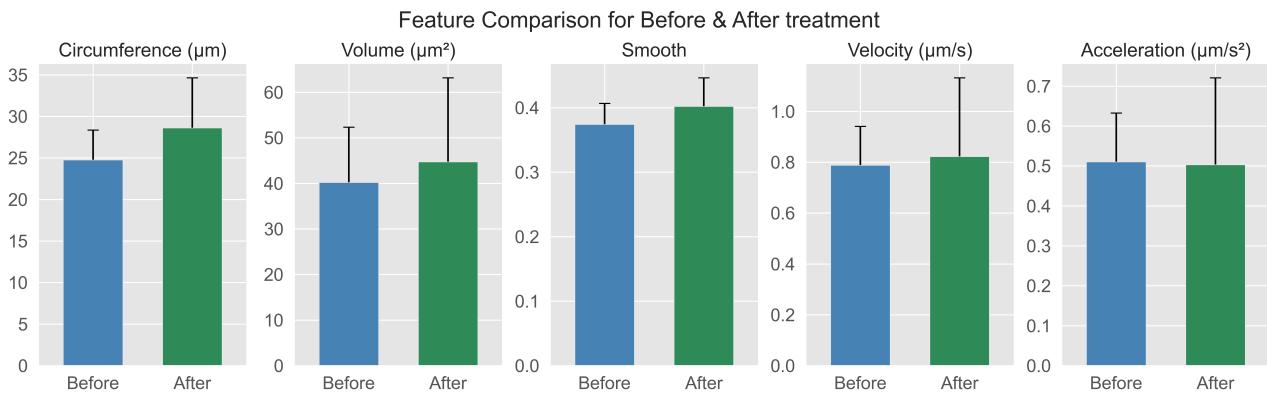
- $Acceleration = 0.584 \times Velocity + 0.057, R^2 = 0.681, F(1, 237) = 505.2, p < .001$
- $Circumference = 0.288 \times Volume + 12.722, R^2 = 0.886, F(1, 237) = 1845, p < .001$

A Shapiro-Wilk test was then applied on each feature both before and after CD11a blocking, and together with inspecting the Q-Q plots (supplementary Fig. S3), results indicate significant departures from normality for all of the scenarios (Table 4.3), which means that henceforth non-parametric tests will be used as they are distribution-free tests and do not require data to be normally distributed.

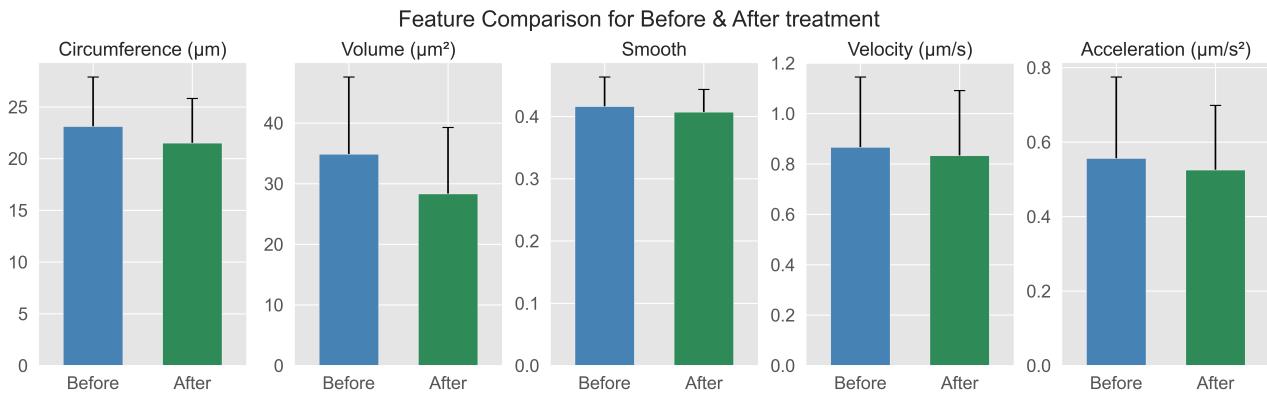
We choose to compare three morphology features and two motion features before and after treat-

Table 4.3: Shapiro-Wilke test shows departure from normality for almost all of the features thus non-parametric tests are used afterwards. df: Degrees of freedom.

Dataset	df	Intensity	Circumference	Volume	Smooth	Velocity	Acceleration
1	197	W = 0.92, p <.001	W = 0.98, p = .009	W = 0.97, p <.001	W = 0.97, p < .001	W = 0.94, p <.001	W = 0.83, p <.001
2	181	W = 0.94, p <.001	W = 0.97, p = .001	W = 0.94, p <.001	W = 0.99, p = .059	W = 0.87, p <.001	W = 0.64, p <.001
3	239	W = 0.96, p <.001	W = 0.98, p <.001	W = 0.94, p <.001	W = 0.83, p < .001	W = 0.87, p <.001	W = 0.95, p <.001
4	225	W = 0.96, p <.001	W = 0.98, p = .002	W = 0.95, p <.001	W = 0.87, p < .001	W = 0.90, p <.001	W = 0.83, p <.001



(A)



(B)

Figure 4.2: Median value with errorbar represents 75% of interquantile range. (A) For dataset 3&4, a Mann-Whitney U test indicate that circumference, volume, smooth and velocity in region 2 increased significantly after treatment, $p < .001$, $= .021$, $< .001$, $= .008$ whilst acceleration is not statistical significant, $p = .113$. (B) For dataset 1&2, a Mann-Whitney U test indicate that circumference, volume and smooth in region 1 decreased significantly after treatment, $p = .044$, $= .014$, $= .020$ whilst the change of velocity and acceleration is not significant, $p = .135$, $= .140$.

ment based on the assumption that the shape of a migrating cell might become more irregular than a still cell and different neutrophil sub-populations might possess distinctive behaviours. The median value of dataset 3&4 is presented in Fig. 4.2(A) whilst that of dataset 1&2 is presented in Fig. 4.2(B) and Mann Whitney U test is used to test whether two samples are likely to derive from same population. Both experiments confirms that the migration behaviour of neutrophils in the spleen remains intact in the absence of CD11a and neutrophils' velocity can actually increase in some part of the spleen with statistical significance which shows the behaviour at various parts of the spleen might

not be homogeneous. The experiments derived from two regions of spleen shown in Fig 4.2 indicates CD11a might not be the decisive factor of neutrophils' migration behaviour or neutrophil has alternative mechanism as substitution when CD11a is blocked.

Since neutrophil can switch phenotype and exist as various sub-populations with distinctive behaviours [27], hierarchical clustering is applied on all CD11a related treatments combined to identify sub-populations and how different sub-populations react in the absence of CD11a, then a clearer perspective of the role of CD11a could be provided.

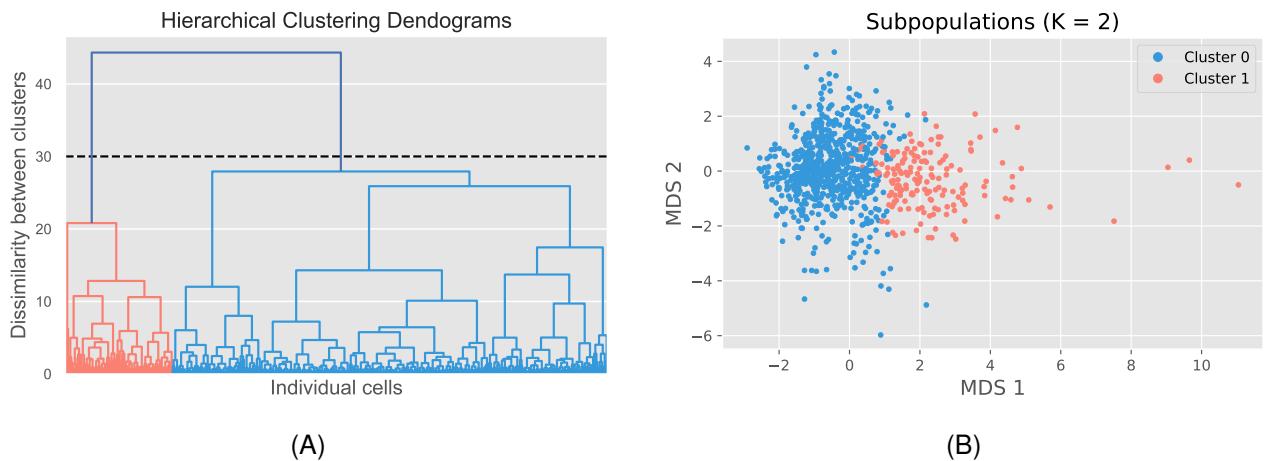


Figure 4.3: Figures are generated with dataset 1~4 combined. (A) Hierarchical clustering with dendograms shows the relationship between objects and number of clusters could be set to 2 according to dissimilarity between clusters. (B) Multi-Dimensional Scaling showing the motility relationship between the analysed neutrophils. This plot visualises the overall differences between 2 clusters based on 3 migration features (velocity, acceleration, smooth) and intensity.

Hierarchical clustering (Fig. 4.3(A)) and subsequent MDS (Fig. 4.3(B)) indicate that neutrophils can be divided into two sub-populations and these sub-populations are quite distinctive. Following this outcome, number of neutrophils and percentage accounted belong to corresponding clusters within different dataset are calculated and shown in Table 4.4.

Table 4.4: Number of neutrophils in each clusters and datasets. Intuitively, a decrease of cluster 0 and increase of cluster 1 can be observed in dataset 3&4, whilst that has not changed much in dataset 1&2. This phenomenon is confirmed by Chi-square test.

Dataset Index	Cluster 0	Cluster 1
1	148 (75.13%)	49 (24.87%)
2	146 (80.66%)	35 (19.34%)
3	213 (89.12%)	26 (10.88%)
4	169 (75.11%)	56 (24.89%)

A chi-square test of independence was performed to examine the relationship between treatment

and neutrophil types based on Table 4.4. The relationship between neutrophil subpopulations and CD11a blocking is insignificant for dataset 1&2, $X^2(1, 394) = 1.446$, $p = .229$ whilst that relationship is significant for dataset 3&4, $X^2(1, 478) = 15.033$, $p < .001$. Since cluster 1 generally have higher migration ability than cluster 0 (scatter plot example are shown in supplementary Fig. S4), thus, a prudential conclusion could be made that the injection of CD11a blocking antibody might indirectly trigger a mechanism which could activate some neutrophils, increasing their ability to migrate. However, this effect generated by CD11a might be region-dependent as the relationship in dataset 1&2 which experiment conducted at different part of the spleen, is not significant.

From the analysis above we suspect that CD11a does have the ability to affect neutrophils' migration behaviour and the behaviour of sub-populations, however, detailed mechanisms of how this effect occur and why the results are region-dependent needs further investigation.

4.3 Neutrophil-B cell interaction

A contact-dependent interaction between neutrophils and B cells should satisfy two requirements defined in Section 3.4, a low proximity and their velocities are smaller than the expectation of fitted distribution. For this research study, a pair of neutrophil and B cell are counted as one interaction, e.g. a neutrophil fulfils the above descriptions with 3 different B cells in a single frame, this is counted as 3 interactions instead of 1. In this respect, different interaction aspects are investigated.

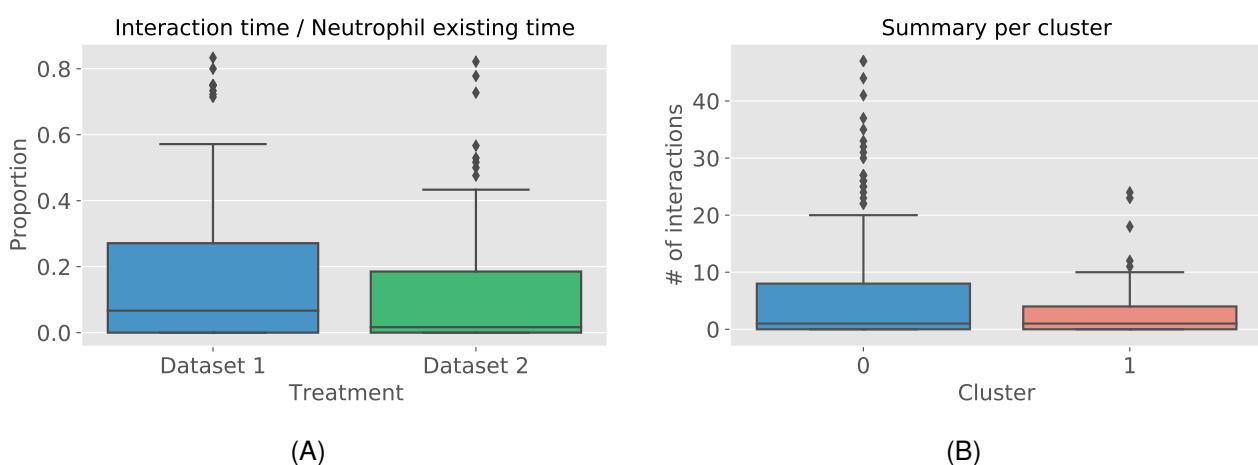


Figure 4.4: Figures are generated with dataset 1&2. (A) Interaction time is defined as number of frames neutrophil has interaction, existing time is total number of frames this neutrophil appeared in IVM. The proportion of these two variables for all neutrophils are shown. (B) Number of interactions per neutrophil of two sub-populations. A Mann-Whitney test suggests this is not significant ($p = .146$) due to distribution under 50% quantile although Cluster 0 have higher possibility of interacting more.

Fig. 4.4(A) shows the interaction time a single neutrophil experienced compared with total time this neutrophil appeared in IVM data. A Mann-Whitney test indicated that the proportion of interaction time and neutrophil existing time is significantly greater for dataset 1 ($Mdn = 0.067$) than that of dataset 2 ($Mdn = 0.017$), $U = 15617.5$, $p = .014$. This circumstance shows before the injection of CD11a blocking antibody, neutrophils tend to spend more time interacting with B cells compared with after injection which might illustrate the effect of drug on cell behaviour. During experiments, interactions between a neutrophil and multiple B cells simultaneously are also observed, with a maximum of 4 B cells. A chi-square test of independence was performed and the result is significant, $X^2(3, 4272) = 24.936$, $p < .001$, less simultaneous interactions occurred after treatment. In terms of neutrophil sub-populations, total number of interactions a neutrophil has with B cells is shown in Fig. 4.4(B). From the box plot we can confirm that Cluster 1, which have a higher average velocity than Cluster 0, interact with B cells less in terms of 75% quantile, maximum value and outliers.

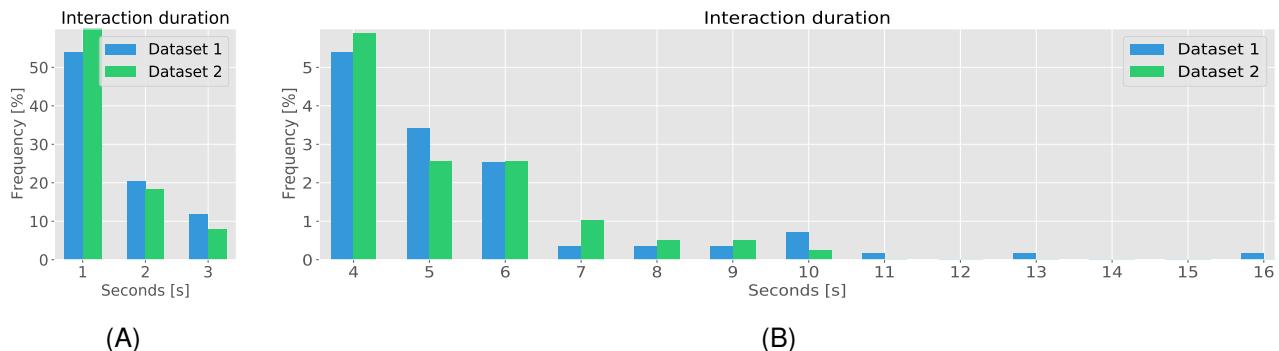


Figure 4.5: Interaction between a pair of neutrophil and B cell, $X^2(9, 941) = 8.999$, $p = .437$. (A) For both datasets, most of the interaction duration concentrates at a lower range. (B) The probability of longer interaction duration are low especially for dataset 2 for duration greater than 9 seconds.

Regarding interaction duration between a pair of neutrophil and B cell, Fig. 4.5(A) indicates that the majority of the interactions last less than 4 seconds, approximately 86% for both datasets. Whilst for longer interaction duration ($\geq 4s$), both datasets distributed seemingly equally except that dataset 1 comprises a few extreme lasting interactions ($\geq 10s$), as shown in Fig. 4.5(B). Combined with the above Fig.4.4, a conclusion could be drawn that although the interaction time for a single neutrophil decreased from dataset 1 to dataset 2, the interaction duration with same B cell has not been affected much. This could also demonstrate that more interactions happened in dataset 1 than that of dataset 2, which are 1164 and 768 respectively.

5 Discussion

Reliable and automated cell segmentation and detection in high-throughput imaging data such as IVM is essential for many biomedical research fields [31]. This affects downstream analyses such as the migration behaviour and interaction analyses as detailed in this project and therefore it is important that cell segmantation is as accurate as possible. Commercial software Imaris uses user-defined parameters (requires prior knowledge) to analyse microscopy images. It has a higher error rate compared to other approaches and is also more conservative on poorly detected cells [14].

We therefore implemented ML/DL methods to apply on IVM data for cell segmentation and achieved comparable or better performance than Imaris. Instead of assuming cells have constant features as in Imaris, for instance volume, ML/DL methods give the user more flexibility in defining the training set and thus could identify cells neglected by Imaris. Furthermore, the DL model Mask-RCNN is based on an external dataset in which most training data have more regular shapes and higher signal to noise ratio than IVM images. Although it already shows comparable results with Imaris in this project, the segmentation performance could be further increased if the model is trained from IVM image data. Since labelling target regions individually from raw image is a time-consuming process, for future prospective, the output of fastER could be feed into Mask-RCNN to train the model to acquire better segmentation results.

Beside focusing on the detection accuracy, MOT and data association techniques also are essential in the pipeline and the result should be more solid if cells that experienced merge-split phenomenon could be well monitored instead of excluded from analysis temporarily. To address the above limitation, several methods could be explored and embedded in the pipeline for future work. For instance, building an inference graph representing merge-split relations with generic object model based on spatial connections and coherent motion [32] could maintain identity across multiple interactions, a tracker utilises Markov Chain Monte Carlo based auxiliary variable particle filter [33] has shown feasible outcomes in some situations. A Bayesian multi-target tracking method which maximises the joint probability over all association events has been tested on microscopy image datasets and achieved

reasonable results [34]. Beside motion and morphology features, fluorescence intensity of each cell could also be integrated into the model to re-assign identity after merge-split process of two cells as this characteristic is unlikely to change significantly. We also introduced a parameter (*minimum hit frames* in Section 3.2) at data association stage which investigates a number of consecutive frames the object has been detected before we classify it as a true positive, this could reflect the dynamics of video and reduce false positive rate.

Using currently recognised cells, migration behaviour and interaction sequences are analysed. We observed the differences before and after the injection of CD11a blocking antibody on the neutrophil population as well as sub-clusters. For one experiment, neutrophils have a more active behaviour after treatment and number of cells belong to cluster that has higher average velocity also increased. Whilst for the other experiment, neutrophils' average migration behaviour and the behaviour of sub-populations does not exhibit a statistical significance in the absence of CD11a, which suggests that the effect of CD11a might be region-dependent or other mechanisms might also involved in immune modulation to offset this effect. Therefore, it is important to know which part of the spleen is used for future wet-lab experiments as number of neutrophils, B cells and their relative behaviours can change significantly in different regions.

Interaction between cells are investigated as this phenomenon could refer cellular communication or substances transport through membrane [35] whilst little work has emphasised on cell-cell interaction especially interaction duration to our best knowledge [11]. In this project only interactions between neutrophils and B cells are investigated and neutrophils derived in dataset 1 interact more often with B cells than that of dataset 2, the distribution of interaction duration between a pair of neutrophil and B cell, however, has not been affected much. Theoretically, interaction between neutrophils and neutrophils can be analysed as well under same procedure. Eventually we can finalise a sequence of interactions for individual cells and acquire in-depth insight on cell behaviour. This sequence of interactions is also important to find out whether the data sampling frequency is high enough. However, the accuracy depends on how well the merge-split problem can be addressed.

This project has not fully demonstrated the hypothesis that the migration or contact-dependent in-

teraction behaviour of neutrophils and sub-populations can be affected by the injection of CD11a blocking antibody, as results exhibit region heterogeneity and some thresholds are quite arbitrarily set. Nevertheless, and more importantly, this pipeline is a proof of concept that with the use of ML/DL algorithms, we can achieve equal or better performance in cell segmentation than Imaris and generate more detailed information regarding texture and morphology features. With sufficient information provided, downstream migration, interaction and other series of analysis can be conducted as well. Vivid data visualisation can be implemented as a supplement to static structured data, for instance, from raw IVM video data the multiple object tracking video in Section 4.2 and cell-cell interaction video (Result for dataset 1 can be accessed [here](#) and for dataset 2 can be accessed [here](#). Cells with white contour are currently interacting under the threshold defined in Section 3.4) can be produced. There are other possible aspects to be investigated further, e.g. a pathway that multiple cells travel one after another might have specific functions or a region where new cells pumping in could be a entry point of the spleen.

We have implemented this pipeline and analysis function into Python for versatility and user-friendly operating purpose. Codes and documentation are available via <https://github.com/MichaelGMZ/MRes-2020DS-Project2>. For future work, segmentation algorithms and object tracking techniques could be revised based on above discussion to increase the accuracy of the pipeline and subsequent trajectories. A more precise and flexible definition on migration and interaction could be given to analyse cell behaviours as this project is a proof of concept.

References

- [1] T. N. Mayadas, X. Cullere, and C. A. Lowell, "The multifaceted functions of neutrophils," *Annual Review of Pathology: Mechanisms of Disease*, vol. 9, pp. 181–218, 2014.
- [2] L. G. Ng, R. Ostuni, and A. Hidalgo, "Heterogeneity of neutrophils," *Nature Reviews Immunology*, vol. 19, no. 4, pp. 255–265, 2019.
- [3] E. Kolaczkowska and P. Kubes, "Neutrophil recruitment and function in health and inflammation," *Nature reviews immunology*, vol. 13, no. 3, pp. 159–175, 2013.
- [4] D. Luo, H. M. McGettrick, P. C. Stone, G. E. Rainger, and G. B. Nash, "The roles of integrins in function of human neutrophils after their migration through endothelium into interstitial matrix," *PLoS One*, vol. 10, no. 2, 2015.
- [5] B. L. Walling and M. Kim, "Lfa-1 in t cell migration and differentiation," *Frontiers in immunology*, vol. 9, p. 952, 2018.
- [6] J. Meerschaert and M. B. Furie, "The adhesion molecules used by monocytes for migration across endothelium include cd11a/cd18, cd11b/cd18, and vla-4 on monocytes and icam-1, vcam-1, and other ligands on endothelium..," *The Journal of Immunology*, vol. 154, no. 8, pp. 4099–4112, 1995.
- [7] C. Berlin-Rufenach, F. Otto, M. Mathies, J. Westermann, M. J. Owen, A. Hamann, and N. Hogg, "Lymphocyte migration in lymphocyte function-associated antigen (lfa)-1-deficient mice," *The Journal of experimental medicine*, vol. 189, no. 9, pp. 1467–1478, 1999.
- [8] R. E. Mebius and G. Kraal, "Structure and function of the spleen," *Nature reviews immunology*, vol. 5, no. 8, pp. 606–616, 2005.
- [9] W. Hoffman, F. G. Lakkis, and G. Chalasani, "B cells, antibodies, and more," *Clinical Journal of the American Society of Nephrology*, vol. 11, no. 1, pp. 137–154, 2016.
- [10] I. Puga, M. Cols, C. M. Barra, B. He, L. Cassis, M. Gentile, L. Comerma, A. Chorny, M. Shan, W. Xu, *et al.*, "B cell-helper neutrophils stimulate the diversification and production of immunoglobulin in the marginal zone of the spleen," *Nature immunology*, vol. 13, no. 2, pp. 170–180, 2012.
- [11] J. F. Deniset, B. G. Surewaard, W.-Y. Lee, and P. Kubes, "Splenic ly6ghigh mature and ly6gint immature neutrophils contribute to eradication of s. pneumoniae," *Journal of Experimental Medicine*, vol. 214, no. 5, pp. 1333–1350, 2017.
- [12] A. Masedunskas, O. Milberg, N. Porat-Shliom, M. Sramkova, T. Wigand, P. Amornphimoltham, and R. Weigert, "Intravital microscopy: a practical guide on imaging intracellular structures in live animals," *Bioarchitecture*, vol. 2, no. 5, pp. 143–157, 2012.
- [13] D. Duarte, E. D. Hawkins, O. Akinduro, H. Ang, K. De Filippo, I. Y. Kong, M. Haltalli, N. Ruivo, L. Straszkowski, S. J. Vervoort, *et al.*, "Inhibition of endosteal vascular niche remodeling rescues hematopoietic stem cell loss in aml," *Cell stem cell*, vol. 22, no. 1, pp. 64–77, 2018.
- [14] C. Mitchell, L. Caroff, A. Vigilante, J. A. Solis-Lemus, C. C. Reyes-Aldasoro, F. de Chaumont, A. Dufour, S. Dallongeville, J.-C. Olivo-Marin, and R. Knight, "Cell tracking profiler: a user-driven analysis framework for evaluating 4d live cell imaging data," *bioRxiv*, p. 859397, 2019.
- [15] S. Dasiopoulou, V. Mezaris, I. Kompatsiaris, V.-K. Papastathis, and M. G. Strintzis, "Knowledge-assisted semantic video object detection," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 10, pp. 1210–1224, 2005.
- [16] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [17] O. Hilsenbeck, M. Schwarzfischer, D. Loeffler, S. Dimopoulos, S. Hastreiter, C. Marr, F. J. Theis, and T. Schroeder, "faster: a user-friendly tool for ultrafast and robust cell segmentation in large-scale microscopy," *Bioinformatics*, vol. 33, no. 13, pp. 2020–2028, 2017.
- [18] Z. Ma, J. M. R. Tavares, R. N. Jorge, and T. Mascarenhas, "A review of algorithms for medical image segmentation and their applications to the female pelvic cavity," *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 13, no. 2, pp. 235–246, 2010.

- [19] Z.-Q. Zhao, P. Zheng, S.-t. Xu, and X. Wu, "Object detection with deep learning: A review," *IEEE transactions on neural networks and learning systems*, vol. 30, no. 11, pp. 3212–3232, 2019.
- [20] F. Lu, F. Wu, P. Hu, Z. Peng, and D. Kong, "Automatic 3d liver location and segmentation via convolutional neural network and graph cut," *International journal of computer assisted radiology and surgery*, vol. 12, no. 2, pp. 171–182, 2017.
- [21] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241, Springer, 2015.
- [22] T. Kanade, Z. Yin, R. Bise, S. Huh, S. Eom, M. F. Sandbothe, and M. Chen, "Cell image analysis: Algorithms, system and applications," in *2011 IEEE Workshop on Applications of Computer Vision (WACV)*, pp. 374–381, IEEE, 2011.
- [23] R. E. Kalman, "A new approach to linear filtering and prediction problems," 1960.
- [24] Y. Cheng, "Mean shift, mode seeking, and clustering," *IEEE transactions on pattern analysis and machine intelligence*, vol. 17, no. 8, pp. 790–799, 1995.
- [25] A. Issekutz and T. Issekutz, "The contribution of Ifa-1 (cd11a/cd18) and mac-1 (cd11b/cd18) to the in vivo migration of polymorphonuclear leucocytes to inflammatory reactions in the rat.," *Immunology*, vol. 76, no. 4, p. 655, 1992.
- [26] C.-W. Hong, "Current understanding in neutrophil differentiation and heterogeneity," *Immune network*, vol. 17, no. 5, pp. 298–306, 2017.
- [27] C. Rosales, "Neutrophil: a cell with many roles in inflammation or several cell types?," *Frontiers in physiology*, vol. 9, p. 113, 2018.
- [28] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *Proceedings of the IEEE international conference on computer vision*, pp. 2961–2969, 2017.
- [29] O. Maimon and L. Rokach, "Data mining and knowledge discovery handbook," pp. 197–264, 2005.
- [30] A. Bewley, Z. Ge, L. Ott, F. Ramos, and B. Upcroft, "Simple online and realtime tracking," in *2016 IEEE International Conference on Image Processing (ICIP)*, pp. 3464–3468, IEEE, 2016.
- [31] M. Etzrodt, M. Endele, and T. Schroeder, "Quantitative single-cell approaches to stem cell research," *Cell stem cell*, vol. 15, no. 5, pp. 546–558, 2014.
- [32] B. Bose, X. Wang, and E. Grimson, "Multi-class object tracking algorithm that handles fragmentation and grouping," in *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8, IEEE, 2007.
- [33] Z. Khan, T. Balch, and F. Dellaert, "Multitarget tracking with split and merged measurements," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 1, pp. 605–610, IEEE, 2005.
- [34] A. Genovesio and J.-C. Olivo-Marin, "Split and merge data association filter for dense multi-target tracking," in *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, vol. 4, pp. 677–680, IEEE, 2004.
- [35] A. Aplin, A. Howe, S. Alahari, and R. Juliano, "Signal transduction and signal modulation by cell adhesion receptors: the role of integrins, cadherins, immunoglobulin-cell adhesion molecules, and selectins," *Pharmacological reviews*, vol. 50, no. 2, pp. 197–264, 1998.

Supplementary Materials

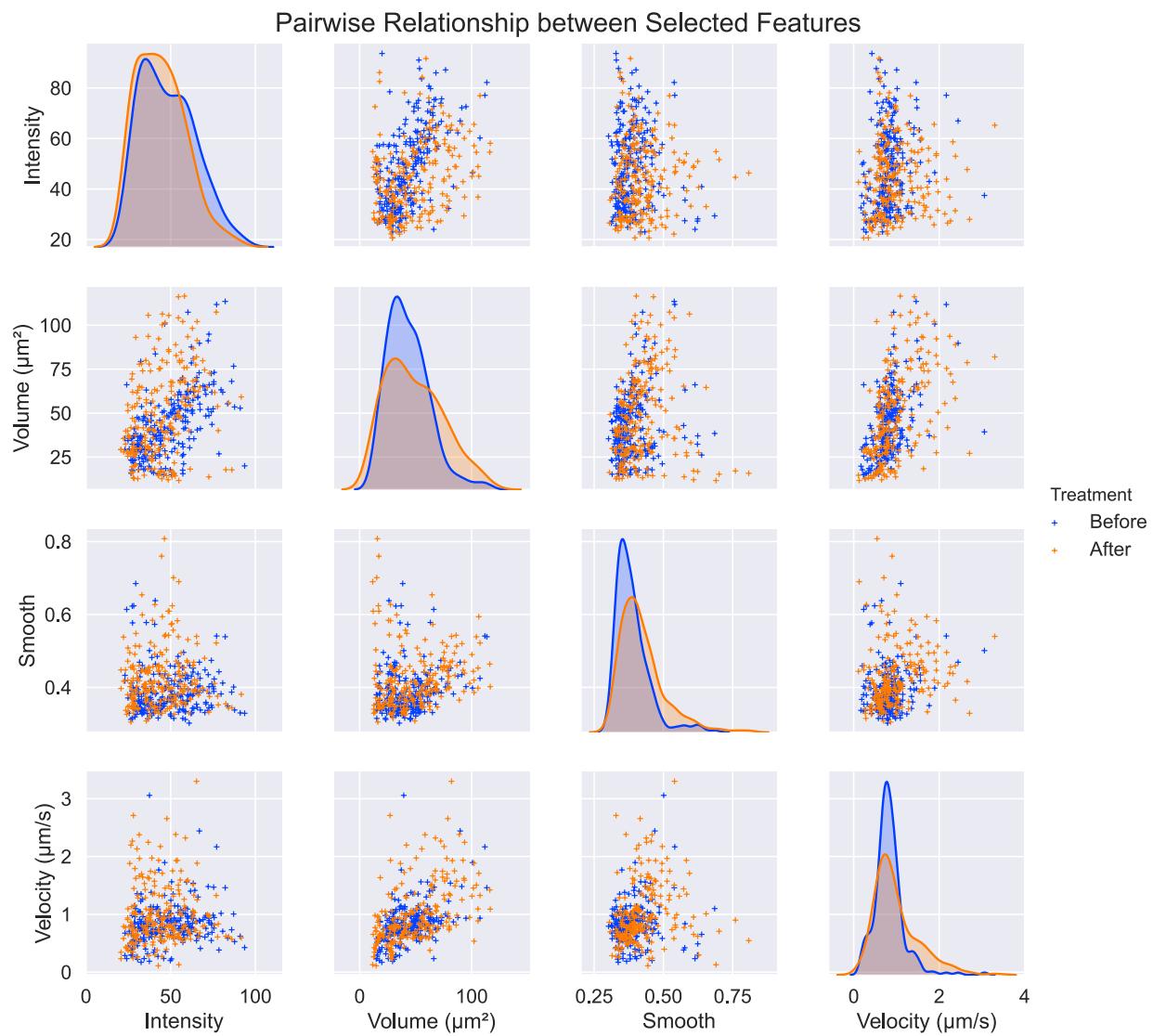


Figure S1: Relationship among features extracted from dataset 3 & 4 are shown as scatter plot and kernel density estimation.

Pairwise Relationship between Selected Features

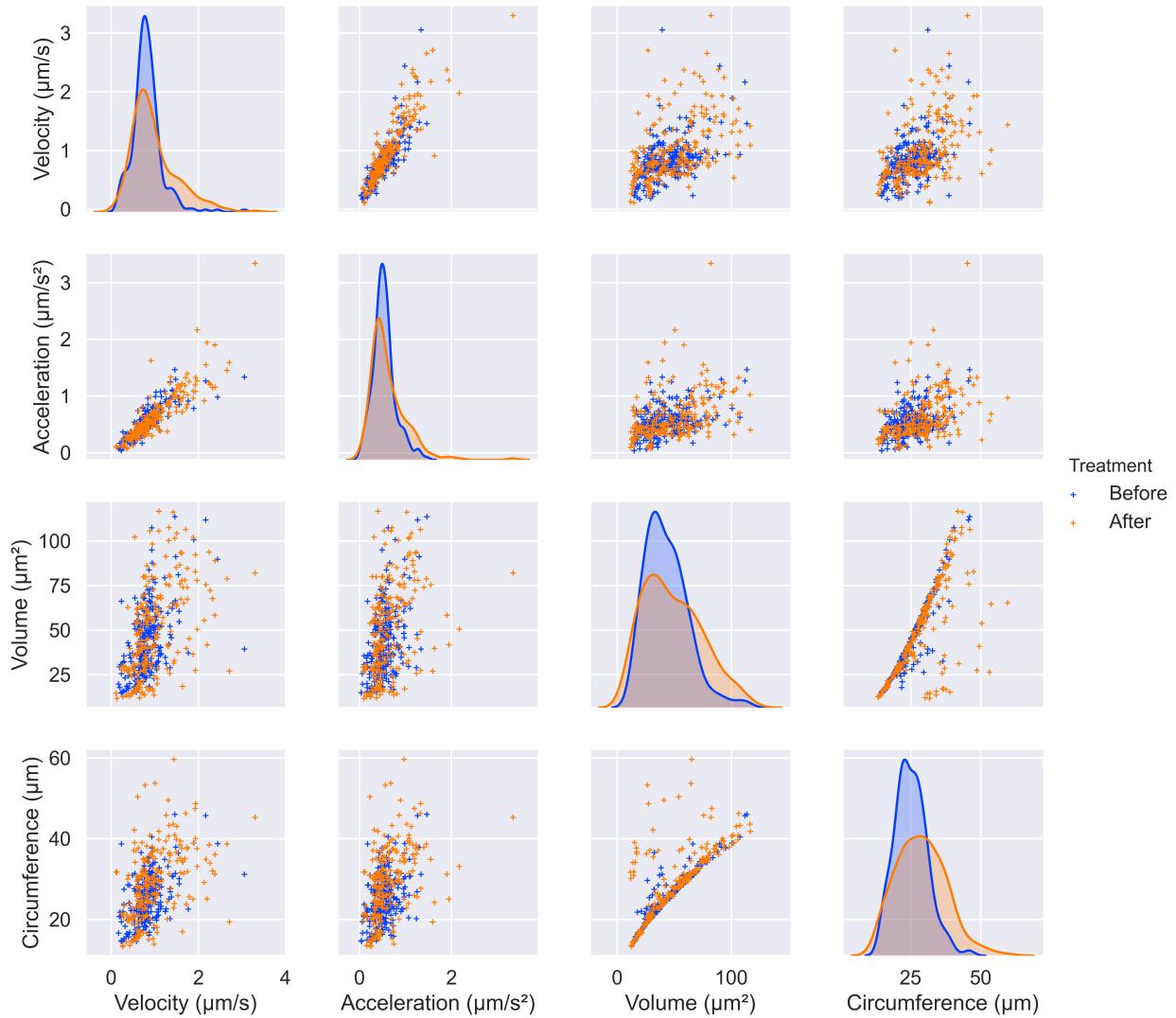


Figure S2: Four features are shown as scatter plot with kernel density estimation. High R^2 among velocity with acceleration and volume with circumference in dataset 3&4.

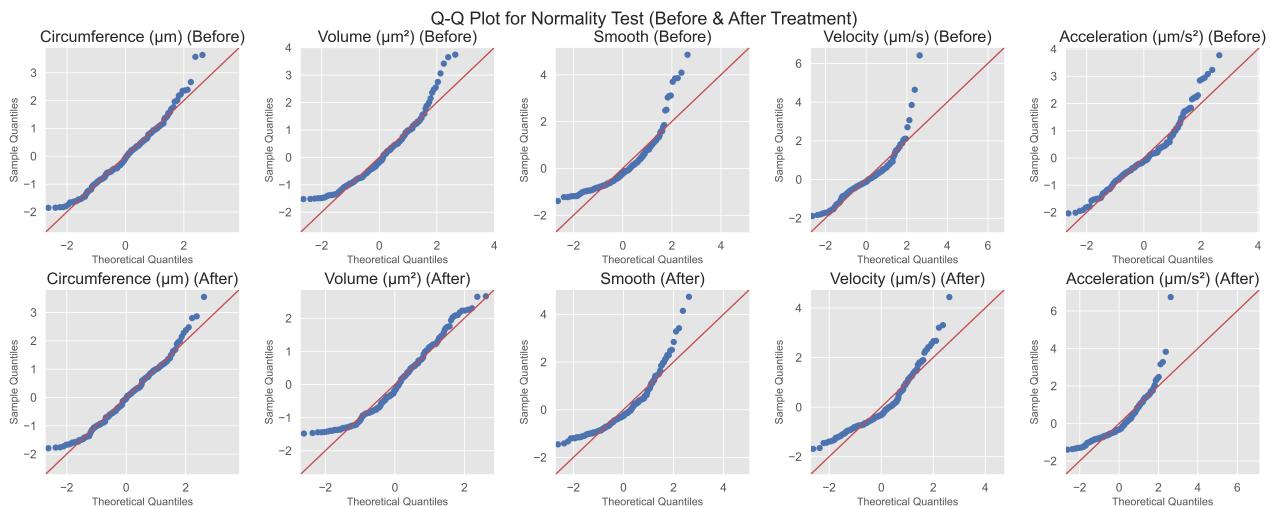
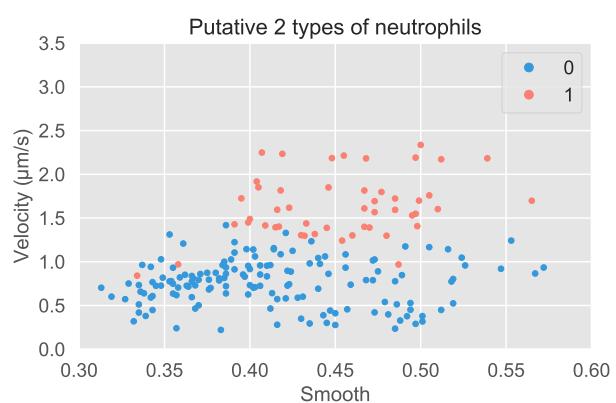
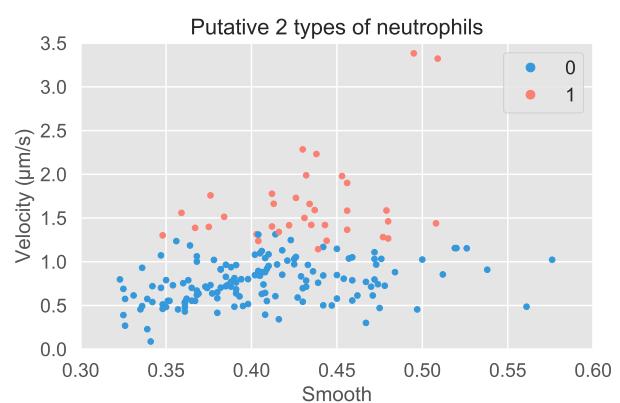


Figure S3: QQ-Plot shows the distribution of those features are not likely to derive from normal distribution. X-axis represents theoretical quantile whilst Y-axis means sample quantile.



(A)



(B)

Figure S4: (A) Scatter plot comparison of smooth and velocity for dataset 1. (B) Scatter plot comparison of smooth and velocity for dataset 2.