Title

Data-Driven Modeling and Analysis of Biological Systems’ Response Over Time

Author

Kesapragada, Manasa

Publication Date

2024

UNIVERSITY OF CALIFORNIA
SANTA CRUZ

DATA-DRIVEN MODELING AND ANALYSIS OF BIOLOGICAL SYSTEMS' RESPONSE OVER TIME

A dissertation submitted in partial satisfaction of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

in

APPLIED MATHEMATICS

by

Manasa Kesapragada

September 2024

The dissertation of Manasa Kesapragada
is approved:

Dr. Marcella M. Gomez, Chair

Dr. Hongyun Wang

Dr. Min Zhao

Dr. Yao-Hui Sun

Peter Biehl
Vice Provost and Dean of Graduate Studies

Contents

List of Figures

List of Tables

Abstract

Data-Driven Modeling and Analysis of Biological Systems' Response over Time

Manasa Kesapragada

Understanding the dynamic responses of cellular processes to stimuli is crucial for uncovering the regulatory mechanisms underlying complex biological phenomena. This work primarily focuses on modeling and analyzing biological system responses over time. In some cases, there is no stimulus, while in others, stimuli such as electric fields (EF), nutrient exposure, or initial wounding are present. This study underscores the significance of considering dynamics in biological systems and the need to develop methods for analyzing and modeling temporal dynamics.

In the first part of our research, we focus on macrophages, a type of immune cell, and their subtypes (M0, M1, and M2). We develop robust image processing methods for single-cell segmentation and tracking using single-cell time-lapse microscopy and label-free live-cell imaging. By mapping the morphological features to cell migratory behavior, we train a deep-learning model to classify macrophage subtypes. Our findings reveal distinct migratory behaviors for M1 and M2 macrophages, demonstrating that cell motility and morphology can effectively identify functionally diverse macrophage phenotypes. This has significant implications for developing cost-efficient, high-throughput screening methods targeting macrophage polarization. Building on this foundation, the second part of our work explores cell subtype classification in the context of galvanotaxis, where cells migrate in response to electric fields (EF). We extend our image processing and machine learning framework to control and analyze cell migration under EF, providing new insights into the mechanisms driving cellular responses to electrical stimuli. In the final segment of this thesis, we apply our single-cell microscopy image processing techniques to study bacterial spore germination, developing a predictive model for spore germination with various germinants administered at different intervals. This interdisciplinary approach enhances our understanding of spore biology and demonstrates the versatility and applicability of our image-processing methods across different biological systems.

This study presents a comprehensive approach that combines advanced image processing techniques, machine learning algorithms, and experimental methodologies to model and analyze biological system responses over time. Additionally, we highlight the importance of non-computationally expensive methods for processing images and quantifying behavior, enabling real-time control and analysis.

To My Family

Acknowledgements

I would like to thank my advisor, Marcella Gomez, for giving me this opportunity and for her continual guidance and support. My sincere gratitude to Ksenia Zlobina for being an exceptional mentor and for her invaluable assistance. I would like to thank Yao-Hui for sharing his extensive biological knowledge and for encouraging me to produce exciting research. I would also like to thank Min Zhao for the inspiring ideas and research opportunities provided to me. I am grateful to Mohammad Jafari for his support during the initial years. To all those I've had the opportunity to work with, past and present, on the DARPA project, I have learned much from each of you.

1 | Introduction

1.1 Background

Image processing techniques can be applied to quantify various aspects of cellular behavior, including morphology, motility, and migratory patterns, which are crucial for understanding the dynamic responses of biological systems. These quantified behaviors can then be integrated into machine-learning models to predict cell phenotypes and interactions in a variety of contexts. This work is centered on modeling and analyzing how biological systems respond over time, with an emphasis on dynamic changes that occur in response to different stimuli, whether intrinsic or extrinsic, such as electric fields, nutrient availability, or physical perturbations.

This research was initiated to address the broader challenge of classifying and predicting cellular behaviors in complex environments. The ability to precisely characterize and model cellular responses provides valuable insights into fundamental processes such as immune responses, tissue regeneration, and disease progression. For instance, understanding how cells, including but not limited to macrophages, dynamically alter their behavior in response to environmental cues can inform strategies for therapeutic intervention.

This study seeks to develop methodologies that can generalize across different cell types and conditions by utilizing a combination of image processing and machine learning. Whether tracking the migration of immune cells, the proliferation of stem cells, or the differentiation of neurons, the goal is to uncover patterns in cellular responses that are predictive of broader biological outcomes. These insights can contribute to developing more effective treatments and interventions by providing a deeper understanding of the temporal and spatial dynamics that govern cellular behavior.

1.2 Motivation

Cellular morphological changes [2] and motility behaviors are central to a wide range of biological processes, from immune responses to tissue development and disease progression. While cell shape and size have been utilized in characterizing various cell types, the role of motility properties in such characterizations still needs to be explored across different cell systems. Understanding how distinct migratory behaviors correlate with specific phenotypic states can provide deeper insights into cellular function and response to environmental cues.

Quantifying dynamic features such as lamellipodial protrusions and other motility-related attributes using image processing techniques can significantly enhance our ability to identify and classify cell behaviors. By leveraging these techniques in combination with machine learning algorithms, we can detect and predict cellular behavior and migratory patterns in various contexts, from immune cell recruitment to responses in other cellular systems.

Importantly, this analysis is an important step towards real-time control of biological systems, a goal that requires methods to be computationally efficient and based on information that can be measured in real time. This is why focusing on phase contrast images is crucial, as they allow for non-invasive, real-time analysis of cell dynamics. An example of such an application is detailed in Chapter 5, where the developed methods are applied to immune cells to control and guide the cells along a desired trajectory in response to varying electric fields.

1.3 Outline

The proposed contributions of this work is to lay a framework of methods that can be applied towards understanding cellular morphology, motility and migratory patterns. This work spans from the study of immune cells to bacterial spore dynamics, showcasing the adaptability of the developed methods. The outline of this work can be summarized as follows:

- **Image Processing and Cell Segmentation - Tracking**

  Computer vision analysis of non-fluorescent microscopy images, representatively phase-contrast microscopy images, promises to realize long-term monitoring of live cell behavior with minimal perturbation and human intervention. In this section, we propose a computer vision-based system that detects and tracks individual cells in order to realize automated long-term monitoring of cell behavior. This section detects cells to measure their shape and tracks cells to quantify cell migration.

- **Data-Driven Approach to Establishing Cell Motility Patterns as Predictors of Cell Subtypes and Their Relation to Cell Morphology**

  Using the image analysis section's cell segmentation and tracking results, we create a training data set without any pre-assigned labels or scores. We first self-discover any naturally occurring patterns in this training data set using unsupervised learning algorithms like clustering on the cell morphological parameters. We then find the behavioral parameters of the identified patterns. We also identify the trajectories of these patterns and analyze if this could help to distinguish the macrophage subtypes.

- **Deep Learning Classification for Macrophage Subtypes through Cell Migratory Pattern Analysis**

  To understand the pattern of cell migratory behavior, we propose a deep learning model to learn the representations from the cell trajectories and identify the cell patterns. This model helps classify the cell trajectories of the macrophage subtypes. We use the results from the clustering analysis as labeled data and train the model. The predictions from the model show that each macrophage subtype has a distinct migratory pattern. These conclusions can inform decisions when external cues like electric fields (EF) are applied to the cells to inform future wound intervention and management.

- **Controlling Cell Migratory Patterns and Classifying Cell Subtypes Under Electric Field**

  Feedback control can be used to direct cell migratory patterns by automatically regulating an external electric field, with potential medical applications such as in wound healing. In this chapter, we implement a machine learning-based controller to regulate cell migration in real-time. An image analyzer was developed to identify cells and measure their speed, directedness, and recruitment index within a closed-loop system. The controller was used *in vitro* to guide naive macrophage cells along a desired trajectory. One open area for future research is the consideration of changes in macrophage phenotype due to prolonged exposure to EF during control efforts. Preliminary work is presented towards classifying macrophages during galvanotaxis to identify changes in macrophage subtypes in real time.

- **Machine Learning based model to Predict Germination Status in Bacterial Spores**

Understanding bacterial spore germination can lead to a wide range of applications where spore reduction or encouragement of activity can be exploited, including medical, food safety, and agricultural applications. This chapter studies spore germination mechanisms through microscopy imaging to extract key physiological features and develop a model for predicting germination. This application highlights the versatility of our techniques, demonstrating their effectiveness across different cell types and with both phase contrast and fluorescence microscopy images. The results underscore the generalizability of our approach while focusing on the primary objective of controlling spore germination. Achieving effective real-time control depends on integrating real-time analysis, modeling, and prediction. Notably, this work introduces a predictive model for the first time, marking a significant advancement beyond previous efforts that were limited to classification.

- **Conclusion**

  Finally, this chapter summarizes the results and discussed potential directions for future research.

We intentionally made each chapter in this dissertation independent by including all relevant background information and method descriptions within each chapter. Readers are welcome to approach the chapters out of order.

2 | Image Processing and Cell Segmentation - Tracking

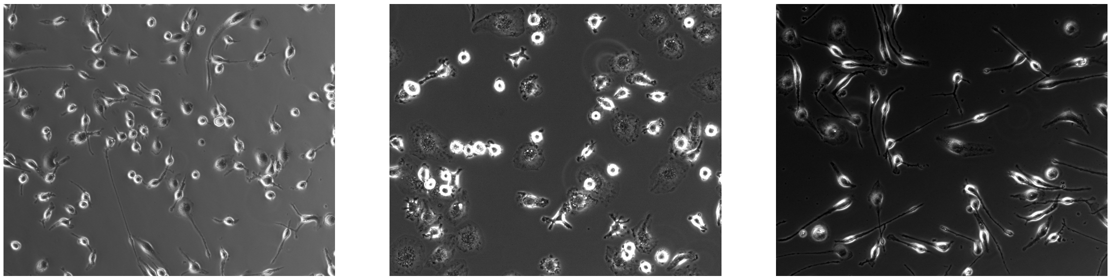
2.1 Types of Image Datasets

Cell images can vary significantly depending on the type of microscopy, staining methods, cell type, and cell density. One of the primary challenges in analyzing these images is extracting meaningful cell information. The data generated from imaging experiments are diverse, reflecting the variations in microscopy techniques, staining methods, and the inherent differences in cell types and densities [3]. For example, brightfield microscopy images show cells against a bright background, making the cells appear darker. Contrast microscopy images, on the other hand, present cells with boundaries that differ in contrast from the background, which may be similar in tone to the cells. Fluorescent microscopy images typically display cells as bright objects against a dark background, making them easier to segment. Additionally, as cell density increases across different images, the variety in cell shapes becomes more apparent, further complicating the process of identifying and segmenting cells. Each set of images poses unique challenges for segmentation.

Accurate cell segmentation is crucial because cell morphology is a critical phenotypic feature that reflects the physiological state of the cell. Moreover, defining the cell contour is often necessary for analyzing intracellular processes. Therefore, developing effective segmentation techniques is essential for extracting valuable insights from these varied cell images.

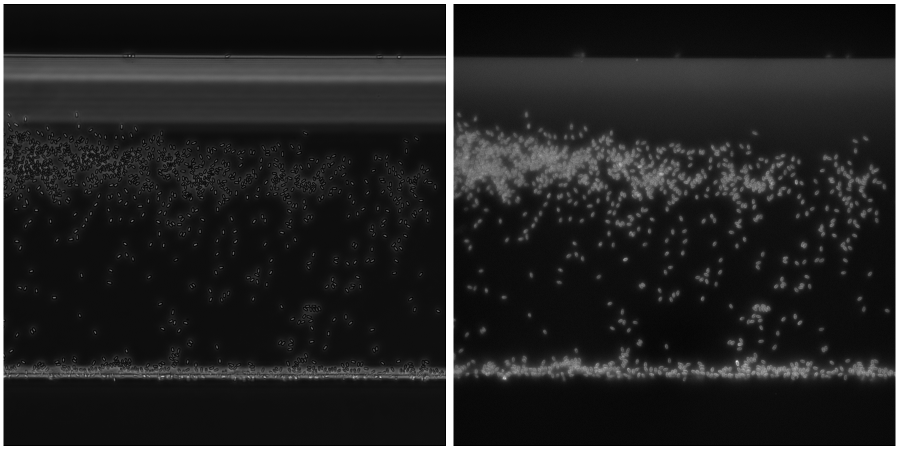
2.2 Image Analysis

In this study, most images were obtained using phase contrast microscopy [4] and fluorescence microscopy, as illustrated in Fig. 2.1 and Fig. 2.2.



(a) M0 image (b) M1 image (c) M2 image

Figure 2.1: Input phase contrast macrophage images for analysis



(a) Phase Contrast microscopy image (b) Fluoroscence microscopy image

Figure 2.2: Bacterial Spores

However, in phase contrast microscopy, a favorable high contrast at the cell boundary leads to problems, such as halo patterns, which can complicate cell segmentation [5]. Detection of the cell boundaries from the current images to quantify the protrusions of the cells, especially the cells like macrophages (Fig. 2.1), without any given annotated data, is a significant challenge. The existing cell segmentation methods for phase-contrast images [6, 7] are pre-trained on different cell types and would need annotated data to re-train their models. In order to identify the most prominent cells in the given images, manual segmentation and tracking would not be a considerate option. So, here I develop methods using existing scientific computing and image-processing libraries [8–12] in python, for cell counting (numbers), the identification of cell types or cell phases (shapes), and the quantification of cell migration (morphodynamics).

In the following sections, I demonstrate the pre-processing, segmentation, tracking, and post-processing steps using the example image M1, shown in Fig. 2.1 (b).

2.3 Image Pre processing

Pre-processing is the name given for operations on images at the lowest level of abstraction —both input and output are intensity images. Such images are usually of the same kind as the original data captured by the sensor, with an intensity image usually represented by a matrix or matrices of brightness values. The aim of pre-processing is an improvement of the image data that suppresses undesired distortions or enhances some image features important for further processing [13]. The following methods are performed on the input images consecutively for pre-processing in the current work:

1. **Grayscale image conversion**

   As a first step, we convert the images to grayscale. A grayscale image, or gray-level image, consists only of shades of gray, which simplifies the data by reducing it to a single intensity value per pixel. In contrast to full-color images, where each pixel requires three values (red, green, and blue) to represent its color, a grayscale image only requires one intensity value. This grayscale intensity is stored as an 8-bit integer, allowing for 256 possible shades of gray, ranging from black to white. The result is a two-dimensional array where the number of rows and columns corresponds to the pixel dimensions of the image. The grayscale conversion is done using the weighted average algorithm from OpenCV [9] as

$$\text{RGB to Gray: } Y \leftarrow 0.299 * R + 0.587 * G + 0.114 * B$$

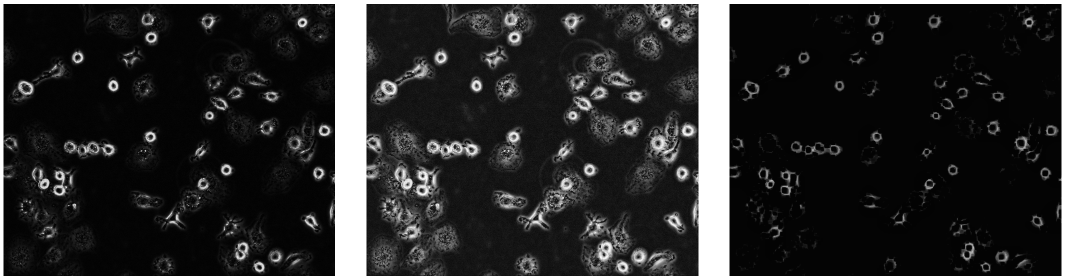
Although phase contrast images do not necessarily need to be converted to grayscale, this process is commonly applied to both fluorescence and phase contrast images to standardize formats for analysis. Phase contrast microscopy images may include color to enhance contrast and highlight different cellular components, but converting them to grayscale can facilitate uniform processing and analysis.

2. **Background Subtraction**

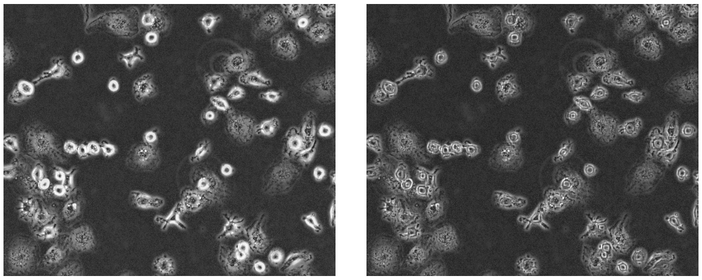
   We leverage the fact that images from similar environments often have comparable backgrounds, allowing us to isolate cells by subtracting the background. For this purpose, we used the rolling ball algorithm [14], a widely used method for background subtraction. This algorithm calculates a local background value for each pixel by averaging over a large spherical region around it. This background value is then subtracted from the original image, aiming to remove significant spatial variations in background intensity. The radius of the sphere is set to be at least as large as the size of the largest object that should not be considered part of the background.

3. **CLAHE - Contrast Limited Adaptive Histogram Equalization**

   Contrast enhancement is an important pre-processing step for the analysis of microscopy images. The main aim of this process is to increase the visibility of objects of interest. In the literature, most of the contrast enhancement techniques are based on modifying the image histogram to improve the contrast [15].



(a) Background Sub. (b) CLAHE (c) Erosion



(d) Dilation (e) Morph. Gradient

Figure 2.3: Results: Pre-processing methods applied in an order on M1 image (Fig. 2.1 (b)), with Fig. 2.2 (e) as the output

Early global Histogram Equalization (HE) techniques directly modify the histogram without considering the artifacts which may occur on the enhanced image. Moreover, these techniques fail to enhance the local details in the image. An efficient implementation, called contrast limited adaptive HE (CLAHE) [16], was proposed to overcome the limitations of HE approaches. The CLAHE technique limits the contrast and eliminates the artifacts caused by the mapping of two close gray-scale values to significantly different values.

4. **Morphological Gradient**

Morphology in images is known as the broad set of image processing operations that process images based on the shapes [17]. It is also known as a tool used for extracting image components that are useful in the representation and description of region shape. The basic morphological operators are erosion, dilation, opening, and closing. In grayscale morphology, images are functions mapping a Euclidean space or grid E into $\mathbb{R} \cup \{\infty, -\infty\}$, where $\mathbb{R}$ is the set of reals, ∞ is an element greater than any real number, and $-\infty$ is an element less than any real number. Grayscale structuring elements are also functions of the same format, called "structuring functions".

Let $f : E \mapsto R$ be a grayscale image, mapping points from a Euclidean space or discrete grid E (such as R2 or Z2) into the real line. Let $b(x)$ be a grayscale structuring element. Usually, b is symmetric and has short-support, e.g.,

$$b(x) = \begin{cases} 0, & |x| \leq 1, \\ -\infty, & \text{otherwise} \end{cases}.$$

Then, the morphological gradient of f is given by:

$$G(f) = f \oplus b - f \ominus b,$$

where $\oplus$ and $\ominus$ denote the dilation and the erosion, respectively. Dilation gives the original set plus an extra boundary, the size, and shape of the boundary depend on the shape and size of the structuring element. Erosion gives the points for which the structuring element is contained in the original set. The outer boundary of the original shape is removed by erosion. The morphological gradient is generated by subtracting an eroded image from its dilated version. The morphological gradient highlights sharp gray-level transitions in the input image.

In Fig. 2.3, we can see the results of the pre-processing methods applied in an order on the M1 image.

2.4 Segmentation

With the pre-processed images in hand, the next step is to individually identify the cells and segment them in every frame. This step helps to capture biologically relevant morphological information. The following methods are performed on the pre-processed images consecutively for the cell segmentation:

1. **Intensity Thresholding**

   Thresholding is a method that replaces each pixel in an image with a black pixel if the image intensity $I_{i,j}$ is less than some fixed constant T (that is, $I_{i,j} < T$), or a white pixel if the image intensity is greater than that constant (see Fig.2.4). To automatically detect this T value, Otsu's method is used which returns a single intensity threshold that separates pixels into two classes, foreground, and background. This threshold is determined by minimizing intra-class intensity variance, or equivalently, by maximizing inter-class variance [18]. This method is equivalent to a globally optimal k-means [19] performed on the intensity histogram.

2. **Binary Fill Holes**

   To fill the holes within a cell in a binary image, the algorithm used consists in invading the complementary of the shapes in input from the outer boundary of the image, using binary dilations [17]. Holes are not connected to the boundary and are therefore not invaded. The result is the complementary subset of the invaded region.

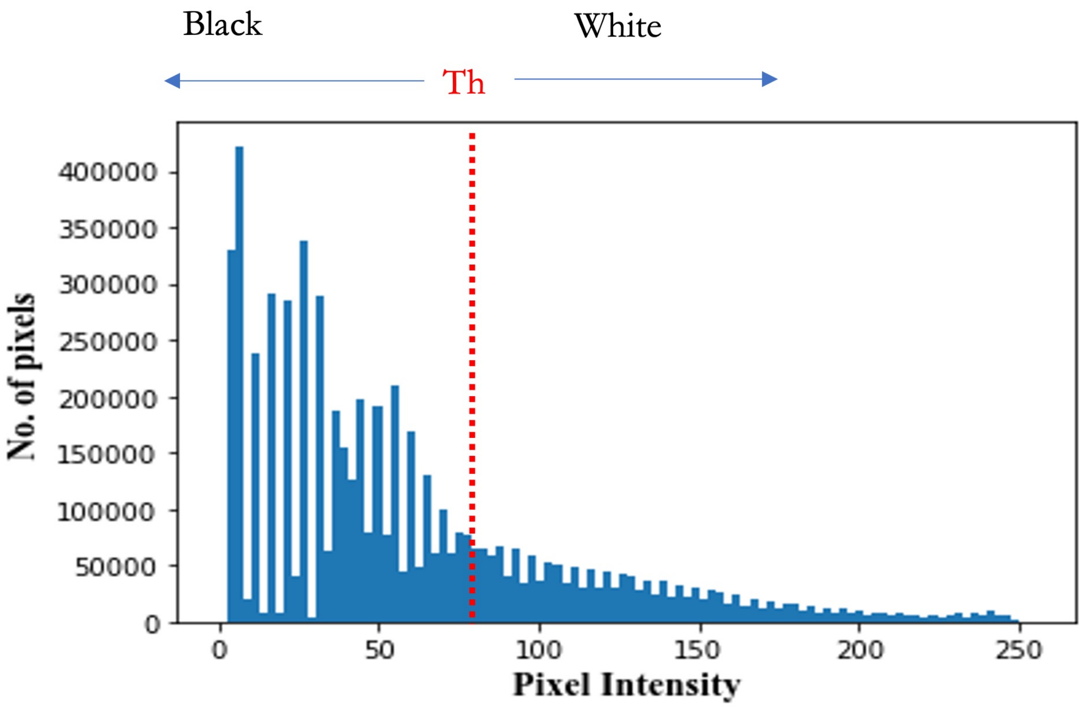


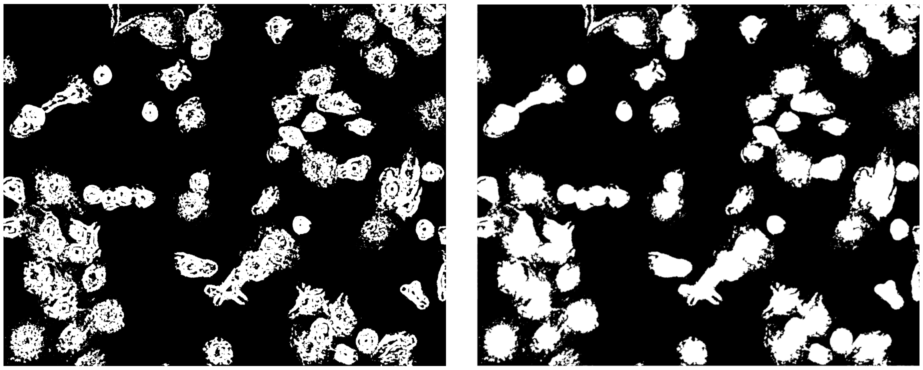
Figure 2.4: Intensity thresholding

3. **Euclidean distance transform and finding peaks**

   The Euclidean distance transform gives values of the Euclidean distance:

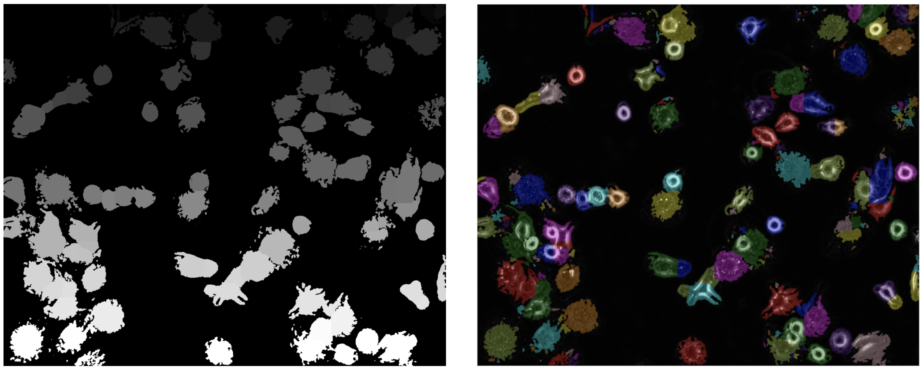
   $$y_i = \sqrt{\sum_{i}^{n} (x[i] - b[i])^2}$$

   where b[i] is the background point (value 0) with the smallest Euclidean distance to input points x[i], and n is the number of dimensions. We use this transform function to compute the exact Euclidean distance from every binary pixel to the nearest zero pixels. This converts a binary digital image, consisting of feature and non-feature pixels, into an image where all non-feature pixels have a value corresponding to the distance to the nearest feature pixel [20]. Now, we use this distance map and find the coordinates of local peaks (maxima) in an image. Peaks are the local maxima in a region of $2 *$ min-distance $+ 1$ (i.e. peaks are separated by at least minimum distance). If peaks are flat (i.e. multiple adjacent



(a) Thresholding

(b) Fill Holes



(c) Labels after Watershed

(d) Segmented Image

Figure 2.5: Results: Segmentation methods applied in an order on the pre-processed M1 image (Fig. 2.2 (e)), with Fig. 2.5 (d) as the output

pixels have identical intensities), the coordinates of all such pixels are returned.

4. **Connected component analysis and Watersheds**

   Now we perform the connected component analysis on the local peaks using 8-connectivity to detect connected regions. The connected component analysis is an algorithmic application of graph theory [21], where subsets of connected components are uniquely labeled based on a given heuristic. Once all the connected components are uniquely labeled, we use the watershed transformation that apportions pixels into marked basins. This transformation treats the image it operates upon like a topographic map, with the brightness of each point representing its height, and finds the lines that run along the tops of ridges [22].

The results of the segmentation methods on the M1 image are shown in Fig. 2.5

2.5 Tracking

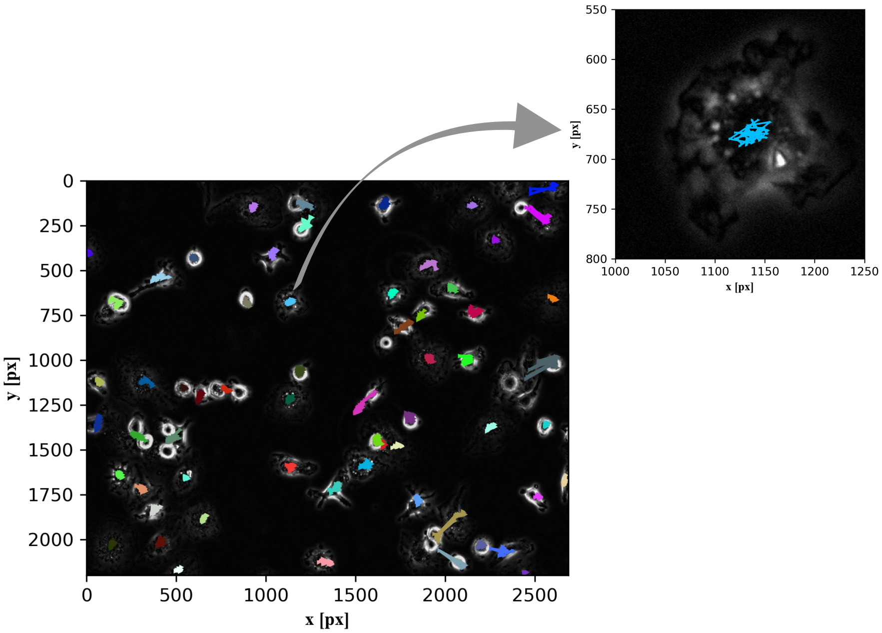


Figure 2.6: The first frame of M1 time-series image-set with trajectories that each cell passed in previous frames

To gain biological insights from time-lapse microscopy recordings of cell behavior, it is often necessary to identify individual cells and follow them over time. Extracting such quantitative information by tracking the cell movements over time, is crucial in understanding the dynamics of cell motion. Having the segmented images and their labels generated through the methods in previous sections, I have used the global linking of cell tracks using the Viterbi algorithm [23] for cell tracking. This algorithm adds tracks to the image sequence one at a time, in a way that uses information from the complete image sequence in every linking decision. This is achieved by finding the tracks which give the largest possible increases to a probabilistically motivated scoring function, using the Viterbi algorithm [24]. The trajectories of the tracked image with one cell highlighted to show in detail can be seen in Fig. 2.6

2.6 Post processing

This is an important step to consider only the cells of interest for analysis after segmentation. The M1 image after applying the post-processing methods is shown in Fig. 2.7, where the tiny particles as in the extreme right of the image, the cells on the boundaries are removed from the analysis.

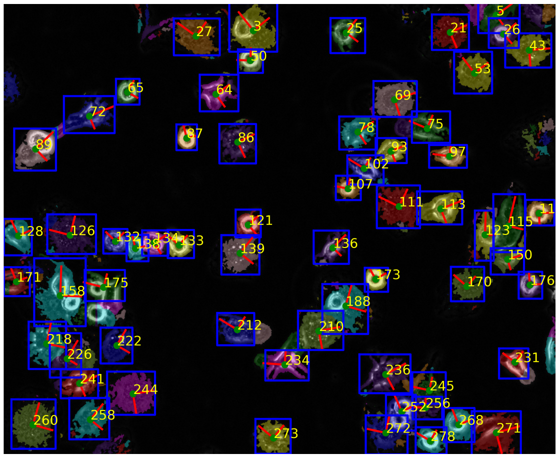


Figure 2.7: Post-processed Image

The post-processing filters to remove the cells which are not used for the analysis are:

- Cells partially present

  The cells which are partially present on the boundary of the images across the time-series are ignored for the analysis.

- Cells not present across all the time-frames

The cells which undergo mitosis and cell death in their time series are not considered for the analysis.

- Cell morphology filter

  As our goal is to classify the cells based on morphology and motility, focusing on cells that maintain consistent shape and size is essential. Therefore, we exclude cells that exhibit significant changes in morphology, specifically those where the change in size from the first frame to the last exceeds the average cell size.

- Cell size filter

  In some instances, small particles may be segmented that do not contribute meaningfully to the analysis. To ensure accuracy, such particles significantly smaller than the average cell size are removed from consideration.

3 | Data-Driven Approach to Establishing Cell Motility Patterns as Predictors of Cell Subtypes and Their Relation to Cell Morphology

The dissertation author was a co-first author of the paper [25], that makes up this chapter. The motility of macrophages in response to microenvironment stimuli is a hallmark of innate immunity, where macrophages play pro-inflammatory or pro-reparatory roles depending on their activation status during wound healing. Cell size and shape have been informative in defining macrophage subtypes. Studies show pro and anti-inflammatory macrophages exhibit distinct migratory behaviors, in vitro, in 3D and in vivo but this link has not been rigorously studied. We apply both morphology and motility-based image processing approaches to analyze live cell images consisting of macrophage phenotypes. Macrophage subtypes are differentiated from primary murine bone marrow derived macrophages using a potent lipopolysaccharide (LPS) or cytokine interleukin-4 (IL-4). We show that morphology is tightly linked to motility, which leads to our hypothesis that motility analysis could be used alone or in conjunction with morphological features for improved prediction of macrophage subtypes. We train a support vector machine (SVM) classifier to predict macrophage subtypes based on morphology alone, motility alone, and both morphology and motility combined. We show that motility has comparable predictive capabilities as morphology. However, using both measures can enhance predictive capabilities. While motility and morphological features can be individually ambiguous identifiers, together they provide significantly improved prediction accuracies (75%) from a training dataset of 1000 cells tracked over time using only phase contrast time-lapse microscopy. Thus, the approach combining cell motility and cell morphology information can lead to methods that accurately assess functionally diverse macrophage phenotypes quickly and efficiently. This can support the development of cost efficient and high through-put methods for screening biochemicals targeting macrophage polarization.

3.1 Introduction

Macrophages reside in or are recruited to all tissues in the body. They play an important role in tissue hemostasis and are an essential part of the human defense system. In wound healing, they are among the first cell types to traffic to the wound. They form an immune defense line, promote, and resolve inflammation, remove dead cells, and cell debris, and support cell proliferation and tissue restructure [26, 27]. There exists a continuum of macrophage functions [28, 29], which have been historically binned into two categories with primarily pro-inflammatory and anti-inflammatory functions, respectively. The two major groups of macrophages that display different functional phenotypes present in the wound bed are referred to as classically activated (pro-inflammatory) and alternatively activated (anti-inflammatory) macrophages. During the inflammation phase, several mediators such as bacterial products like lipopolysaccharide (LPS) and inflammatory cytokines like interferon (IFN)-γ can stimulate macrophages to differentiate into pro-inflammatory status [30,31]. Pro-inflammatory macrophages exhibit antimicrobial properties through the release of inflammatory mediators inducing tumor necrosis factor (TNF)-α, IL-1β, nitric oxide (NO) and interleukin (IL)-6 [32]. Prolonged persistence of pro-inflammatory macrophages without transitioning to anti-inflammatory phenotypes, however, is rather detrimental to wound tissues, which will likely stall, or delay wound healing [33, 34]. On the other hand, macrophages activated by IL-4 and IL-13 develop into so-called alternatively activated macrophages (anti-inflammatory) that suppress inflammatory reactions and adaptive immune responses and play a reparative role during the proliferation and maturation phases [35]. However, like pro-inflammatory macrophages, if anti-inflammatory macrophages persist for too long hypertrophic scars and keloids develop due to excessive collagen production [36, 37].

The involvement of macrophage phenotypes in the wound healing process indicates that detecting macrophage status could be diagnostically useful. However, advances in the field have been hampered, largely due to lack of definitive and optimal biomarkers for macrophage phenotypes. Traditionally, characterization of pro and anti-inflammatory subtypes is carried out by quantification of multiple cell surface markers, transcription factors and cytokine profiles. These approaches are time-consuming, require large numbers of cells/tissues and are resource intensive. Difficulty detecting pro and anti-inflammatory macrophage phenotypes in vivo is compounded by similar problems in in vitro-derived macrophages. For example, Arginase-1, which is considered a classic marker for anti-inflammatory macrophages, is also upregulated in pro-inflammatory macrophages [38,39] and protein expression of Arginase-1 or CD206 is too low for reliable flow cytometry detection [29, 30, 36, 40]. In conclusion, a more robust discriminating system is critically needed to improve the detection and understanding of macrophage phenotypes and their relation to their function.

Because of these distinct roles, macrophages undergo a precisely regulated dynamic transition in their functionalities in time and space. The ability to precisely regulate macrophage distribution, migration, and function polarization (i.e., pro vs. anti-inflammatory macrophages) that spatial-temporally corresponds to wound status offers a powerful approach to regulate wound healing. The immune response involves temporal tracking of the migration of different immune cell types in the wound. When tracking wound healing, we can use morphology and motility to track different cell types and train an algorithm to look for patterns in cellular responses characterized by identifiers, such as location (e.g., near blood vessels) or nearest cell type neighbors.

Macrophages undergo morphological changes when undergoing a change in phenotype [27]. Studies have shown that cell shape can help control macrophages — for instance, elongating them promotes behavior that enhances healing [26]. While cell size and shape have been used to characterize macrophage subtypes, how motility properties are contributing to such quantifications are not well studied, despite pro and anti-inflammatory macrophages exhibiting distinct migratory behaviors.

Machine learning methods have been successfully applied to automate the classification of different cell types from microscopy images [41, 42] and are viewed as a promising approach to high-through put screening. However, these methods are sensitive to the size of the data set, quality of the images, and sometimes facilitated by fluorescent dyes or reporters [43]. Previous work towards classifying macrophage subtype using machine learning leverages fluorescent dyes [44, 45]. Being able to identify cell types from phase contrast microscopy would simplify and accelerate macrophage classification. Furthermore, being able to identify cells based on their motility parameters can eliminate the need for high-quality imaging and result in methods robust to blurred images. We suspect that cell motility like cell morphology is linked to function. Cell tracking can be achieved with image processing techniques in real-time [46]. Motility patterns can be extracted from cell trajectories and machine learning used to classify cells based on their dynamic behavior. Finally, combining cell motility and morphological features can help elucidate the link between cell migratory behavior, morphology, and cellular phenotype.

Here we apply machine learning methods to classify macrophage subtypes and show a tight relationship between cell morphology and its motile behavior. We begin by first validating that our derived macrophage subtypes do present distinct morphological characteristics as has been shown in previous work. Morphological analysis of the label-free collective dataset suggests macrophage shapes are described by three principal shape modes. Clustering analysis is applied to identify the cells belonging to each of these three morphological groupings. Comparison with the labeled data confirms a correlation between macrophage subtype and morphology. Next, we show that independent of cell labels, there is a strong relationship between morphology and motile behavior. Motility features are computed within each cluster showing that distinct morphological features map to distinct cell motility parameters. We then apply discriminant analysis to confirm that this implies a relationship between the labeled cells with respective phenotypes and motility parameters, thereby, supporting the hypothesis that we can predict macrophage subtypes from motility behavior alone. Finally, to predict macrophage subtypes with machine learning, we propose a SVM classification model. We build multiple predictive models using morphological and motility parameters alone and in combination to predict the macrophage subtypes. A high-level schematic of the analysis carried out in this work is shown in Fig 3.1. We demonstrate predictions using motility are comparable to those using morphology and are equally good for the same subtype. This suggests that while they are tightly linked, they may in fact provide complementary information. We show that a combined use of motility and morphology shows an increased prediction accuracy in identifying macrophage subtypes. Hence, by combining both cell motility and morphological information we show that we can reliably detect non-activated naïve macrophages, or macrophages activated by either lipopolysaccharide (LPS) or interleukin 4 (IL-4) exposure at single-cell level.

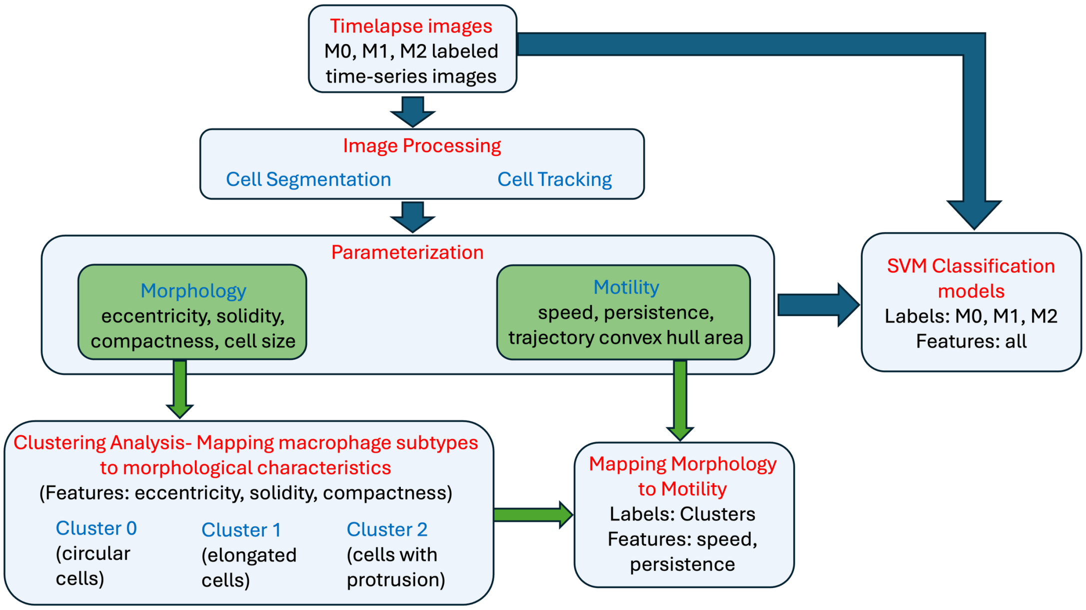


Figure 3.1: **Analysis Workflow.** A high-level schematic of the analysis carried out in this paper. The first step of the workflow consists of image processing methods to perform cell segmentation and cell tracking. Cell segmentation allows for parameterization of the cell morphology and cell tracking allows for parameterization of cell motility over time. Unsupervised clustering analysis on morphological features alone results in cell groupings consistent with parent image labels. Next, it is confirmed that cell motility characteristics are shared within each morphological-based cell grouping and distinct across clusters. Finally, various SVM models are presented using morphology and motility features independently and together to classify macrophage subtypes. The ground truth to the model is provided by the parent image labels.

3.2 Materials and methods

3.2.1 Isolation & culture of bone marrow-derived macrophages

The mouse strains used in this study were of the C57BL/6 background, and both male and female mice were included in the experiments. Mice were procured from Jackson Labs and housed under a strict 12-hour light cycle with a regular chow diet in a specific pathogen-free facility at the University of California (UC), Davis. All animal experiments adhered to regulatory guidelines and standards set by the Institutional Animal Care and Use Committee of UC Davis (Protocol # 21531). Bone marrow-derived macrophages (BMDMs) were generated using standard procedures as previously described [47]. Mice were euthanized using a CO2 chamber and then subjected to cervical dislocation to ensure minimal suffering. The femurs were aseptically dissected, and the bone marrow was flushed with cold Dulbecco's phosphate-buffered saline without calcium and magnesium (DPBS). Given the endpoint nature of these experiments, no anesthesia or analgesia was applied. Bone marrow from 3-6 mice was pooled per batch to account for biological variability. The harvested bone marrow cells were cultured in DMEM (Invitrogen) supplemented with 10% Fetal Bovine Serum (Invitrogen), 1× Antibiotic-Antimycotic solution (Invitrogen), and 20% L-929 conditioned medium for six days, with an additional feeding on day 3, followed by a one-day culture without the conditioned medium. Adherent macrophages were harvested by gently scraping with a "policeman" cell scraper and used in subsequent experiments as needed. Cell viability was determined using trypan blue staining and counting. This study utilized at least four batches of BMDMs, derived from a total of 20 mice of mixed gender, generated at different times.

3.2.2 Activation of bone marrow-derived macrophages

In each experiment, bone marrow-derived macrophages (BMDMs) were seeded into six tissue culture-treated well plates at varying densities and cultured in RPMI-1640 medium (Invitrogen) supplemented with 10% Fetal Bovine Serum (FBS) (Invitrogen) and 1× Antibiotic-Antimycotic solution (Invitrogen) overnight. For M1 activation, 100 ng/ml lipopolysaccharide (LPS) (Sigma, Cat number: L6143) was added to the culture medium, while for M2 activation, 20 ng/ml recombinant mouse interleukin-4 (IL-4) (R&D Systems, Cat number: 404-ML) was used [48]. Two days post-stimulation, activated M1 and M2 macrophages were employed for morphological and motility characterizations as well as functional studies. Macrophages that did not receive any stimulation served as M0 controls.

3.2.3 Time-lapse recording

In each experiment, freshly differentiated macrophages were seeded in 24-well tissue culture-treated plates, cultured, and activated following the procedures described above. Forty-eight hours after stimulation, cells were monitored using a Carl Zeiss Observer Z1 inverted microscope outfitted with a motorized stage and an incubation chamber (37°C and 5% CO2). Time-lapse contrast images were captured using the MetaMorph program (Molecular Devices) with a Retiga R6 (QImaging) scientific CCD camera (6 million, 4.54 μm pixels). Typically, in each experiment, four fields of each condition were selected from different wells under a long-distance 20× objective lens. Images were taken at 5-minute intervals for up to 3 hours unless stated otherwise. We note that in some experiments, certain wells were designated for control purposes or other specific experimental requirements, and thus, they were excluded from the imaging process. In practice, to ensure the quality of our data, we exclude out-of-focus images, retaining only the most suitable ones for subsequent morphological analysis.

3.2.4 Cell segmentation and tracking

The input images are segmented and tracked in a semi-autonomous fashion using the Baxter Algorithms software package [49, 50] written in Matlab. This program segments the cells and tracks them throughout each frame. It computes the 2D coordinates of a cell in each frame along with its region properties.

3.2.5 Computing speed and persistence

The speed and persistence are computed at each frame using the following formula: The speed of a cell from trajectory points (x_1, y_1) to (x_n, y_n) where n is the total length of time frames, could be any point from 2 to last time frame point, d is the total distance travelled, t is the time travelled given by:

$$\text{speed} = \frac{d}{t} = \sum_{i=1}^{n-1} \frac{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}{t_i} \tag{3.1}$$

Persistence is a measure to quantify the length of a cell's path along its entire time to know if a cell has remained closer to its initial time frame trajectory point or if it has moved farther away. It is defined as the ratio of the shortest distance traveled p between the initial point (x_1, y_1) and the nth time frame point (x_n, y_n) to the total distance traveled by the cell d given by:

$$\text{persistence} = \sum_{i=1}^{n-1} \frac{(x_n - x_i)^2 + (y_n - y_i)^2}{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2} \tag{3.2}$$

3.3 Results

3.3.1 Generation of primary murine macrophages

Bone marrow-derived macrophages (BMDMs) are primary macrophages that are differentiated from bone marrow cells [51, 52]. Murine BMDMs present an excellent model and are widely used as prototypical macrophages for investigation of mammalian macrophage functions in vitro. We isolate bone marrow from freshly cut femur bones of C57/BL6 mice. Four batches of bone marrow-derived macrophages were generated, each batch consisting of a mixture of bone marrow from 3-6 mice. In total, bone marrow from 20 mice was used for these experiments. In each experiment we grow bone marrow cells in the presence of murine macrophage colony-stimulating factor (M-CSF), as provided by supplementary L929-conditioned medium [53]. Within 6 days, the bone marrow monocyte progenitors proliferate and differentiate into a homogenous population of mature macrophages (M0). Subsequently, naïve M0 macrophages are further polarized into pro-inflammatory or anti-inflammatory macrophages by addition of bacterial lipopolysaccharide (LPS) or murine interleukin 4 (IL-4), respectively [48] (Appendix Fig. A.2). To simplify notation in the manuscript, we refer to pro-inflammatory macrophages as M1 and anti-inflammatory macrophages as M2. See the methods for details on the protocols used to differentiate macrophages into their respective phenotypes and the supplementary information for results of validation assays confirming three distinct phenotypes based on the literature. Validation assays include phagocytic capacity and quantitative PCR to determine the relative mRNA expression of a panel of selected target genes and transcription factors.

3.3.2 Generation of time-series data capturing spontaneous cell migratory patterns

Using time-lapse microscopy, we generate multiple sets of M0, M1, and M2 images, named after their most representative macrophage subtypes, taken at 5-minute intervals for up to 3 hours (Appendix Fig. A.5). These are the phase contrast images without ground truths or cell labels except the image names based on culturing conditions. Although the macrophages are traditionally binned into three subtypes, they exist on a continuous spectrum, i.e., some cells may be "in-between" identified subtypes, hence we note that the parent images may contain cells of the other morphological subtypes with respect to their labels. In the current analysis, we look mainly for the dominant representations. Cell segmentation and tracking details can be found in the methods section. After cell segmentation and tracking, we select the cells that remain in the image dataset's time-lapse for at least 1 hour. This action allows us to eliminate cells that exit the image area within less than 1 hour of their lifespan along the time series. For the characterization of cell subtypes, we compute cell morphology parameters using the first-time step and determine motility parameters based on time-lapse data. Following this, our time-lapse image datasets resulted in the identification of a total of 1018 distinct cells. These cells comprised 251 cells from M0 images, 369 cells from M1 images, and 398 cells from M2 images.

3.3.3 Three principal shape modes characterize observed macrophage morphologies

We extracted and aligned [54] label-free phase contrast images of a mixed population of live macrophages. To identify the existing principal shape modes, we consider each cell in every frame of the time-lapse datasets as individual data points (n = 2329 cells). We obtain the shape information for M0 (n = 965 cells), M1 (n = 698 cells) and M2 (n = 666 cells) (Appendix Fig. A.3). These shape modes provide a meaningful and concise quantitative description of macrophage morphology, providing significant insights that may underly macrophage phenotypes and functionalities depending on their activation status.

3.3.4 Morphological characterization of macrophage phenotypes

Macrophages are multifaceted, which is determined by the surrounding microenvironment, as well as multifunctional depending on their activation status, which assumes a variety of cell shapes. BMDMs when cultured in vitro and stimulated with cytokines to induce M1 or M2 polarization, displayed markedly different cell morphologies [55]. While unstimulated M0 macrophages show a round or slightly stretching appearance, addition of LPS, which stimulate M1 polarization, cause cell spreading with multiple protrusions within just 24 h of stimulation. In contrast, addition of IL-4, which stimulate M2 polarization, led to cell elongation (Appendix Fig. A.4(A)). Quantitative analysis using Celltool [56] confirmed these three orthogonal modes of shape variability, which account for 86% of the total variation in shapes observed. The existence of only a few meaningful modes in combined macrophage populations implies that the phase space which macrophages reside is in a relatively limited subregion of the space of all possible shapes. In the literature, researchers have used various metrics to describe cell shape [56, 57]. There are several cell shape metrics like the cell area, perimeter length, convex area, major/minor- axis length, and many other. The best choice of metrics depends on the class of cell shapes that are found in the dataset. Here we aim to characterize and quantify the three primary shape modes identified above —circular (mode 1), "with protrusions" shape (mode 2), and elongated (mode 3) through clearly defined morphological parameters. Different combinations and number of parameters were explored. We utilized the following three parameters to characterize the morphological features of macrophage phenotypes, particularly given their uneven boundaries: compactness - to identify the complex and irregular boundaries of the cells; eccentricity - to quantify cell elongation; and solidity - to determine cell density. The combination of compactness, eccentricity, and solidity provided valuable and independent information to describe and differentiate the three primary shape modes identified within the combined macrophage populations.

In Appendix Fig. A.4, we present examples of shape parameter values for the three major cell shapes, showcasing low, medium, and high parameter values for each shape. These parameter values effectively differentiate between the three primary

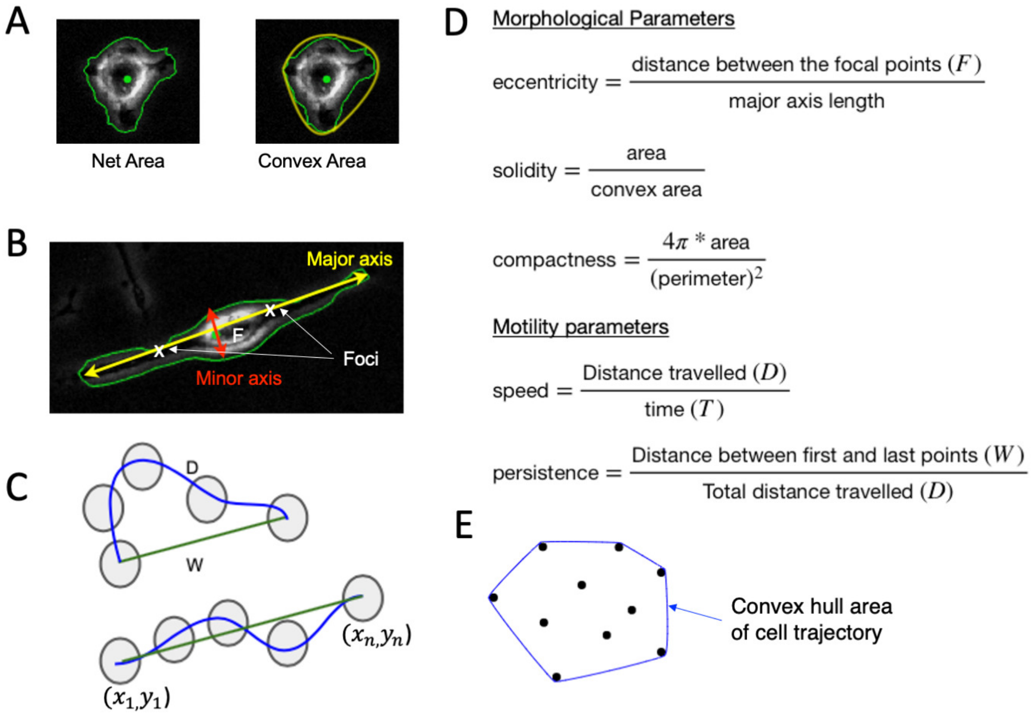


Figure 3.2: **Defining the morphological and motility parameters.** Morphology is defined by metrics related to the (A) area and perimeter of a cell and its convex hull, in addition to the (B) major and minor axis, while motility patterns are defined by (C) displacement metrics, (D) the corresponding formulas are used for morphological/motility parameters and (E) convex hull area of the cell trajectory points.

shape modes and demonstrate the sufficiency of their combination for shape identification. Fig. 3.2 describes and defines these parameters calculated based on the number of pixels in a cell, its boundaries, the distances between them, and the cell trajectories. For the morphological characterization of macrophage subtypes, we take the unsupervised learning approach where the cells are not labeled as M0, M1, and M2 from their image names. We perform k-means clustering [58] with an optimal k value of 4 derived using silhouette score [59] on the three morphological parameters compactness, eccentricity, and solidity to understand the shape of the macrophages. The silhouette coefficient score plot and the silhouette plots for various k-values are shown in (Appendix Fig. A.6). Fig. 3.3A shows the cluster assignment for the cells within the three-dimensional space of their morphological parameters' compactness, eccentricity, and solidity. The identified 1018 cells from parent images M0, M1 and M2 are distributed into four clusters. The following observations are made about each cluster where the data points are represented as swarm plots [60, 61] (Fig. 3.3B-D):

Cluster 0: The cluster of cells with the lowest eccentricity values, highest compactness values, and highest solidity values indicates that the cells in this group are mostly circular, have smooth boundaries with little or no irregular boundaries, and are dense (regular boundaries without any visible holes within the cell structure). Appendix Fig. A.4(C) shows an example cell from this cluster, along with their corresponding eccentricity, solidity, and compactness values.

Cluster 1: The cluster of cells with the lowest eccentricity values, lowest compactness values, and lowest solidity values suggests that the cells in this group are closer to being circular in shape, have highly irregular boundaries, and lack density, potentially due to their irregular boundaries or the presence of holes within the cell. Appendix Fig. A.4(B) shows an example cell from this cluster, along with their corresponding eccentricity, solidity, and compactness values.

Cluster 2: The cluster of cells with higher eccentricity values, above-average compactness values, and higher solidity values suggests that the cells in this group are closer to being elliptical in shape, have relatively smoother boundaries, and are dense (regular boundaries without any visible holes within the cell structure). Appendix Fig. A.4(A) shows an example cell from this cluster, along with their corresponding eccentricity, solidity, and compactness values.

Cluster 3: The cluster of cells with higher eccentricity values, lowest compactness values, and solidity values ranging from least to average suggests that the cells in this group are elliptical in shape, have highly irregular boundaries, and are not densely packed. In Appendix Fig. A.1, we present examples of cells from each cluster. Cluster

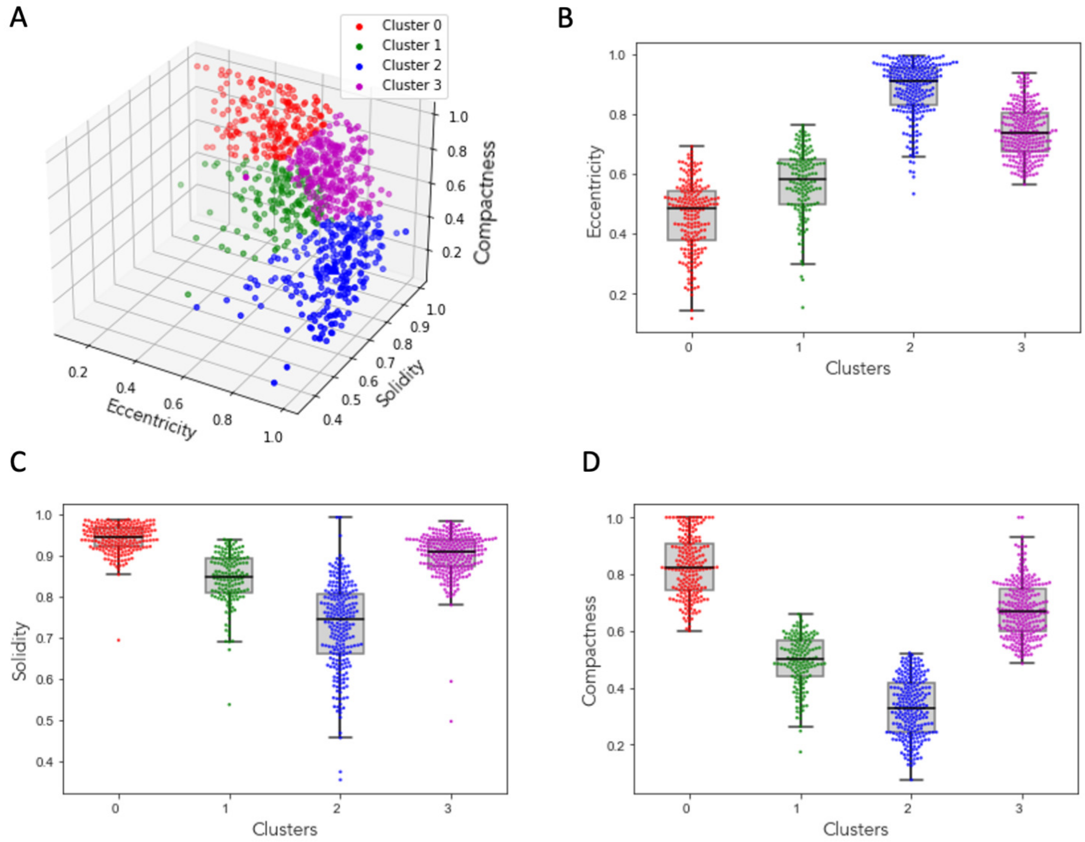


Figure 3.3: **Clustering macrophages based on morphological parameters.** (A) 3D plot of morphological parameters eccentricity, solidity and compactness of the cells showing four distinct clusters. Box plots with overlayed scatter points of eccentricity (B), solidity (C), and compactness (D).

3 contains 333 cells, with 28% from M0, 41% from M1, and 31% from M2 images. Our analysis revealed that this cluster includes indifferent groupings, cell overlapping, or transitional cells. The shape of these cells does not align with any of the primary modes identified in the principal component analysis.

From the above observations, we can describe cells in Cluster 0 as circular cells, cells in Cluster 1 as the cells with protrusions, and cells in Cluster 2 as the elongated cells.

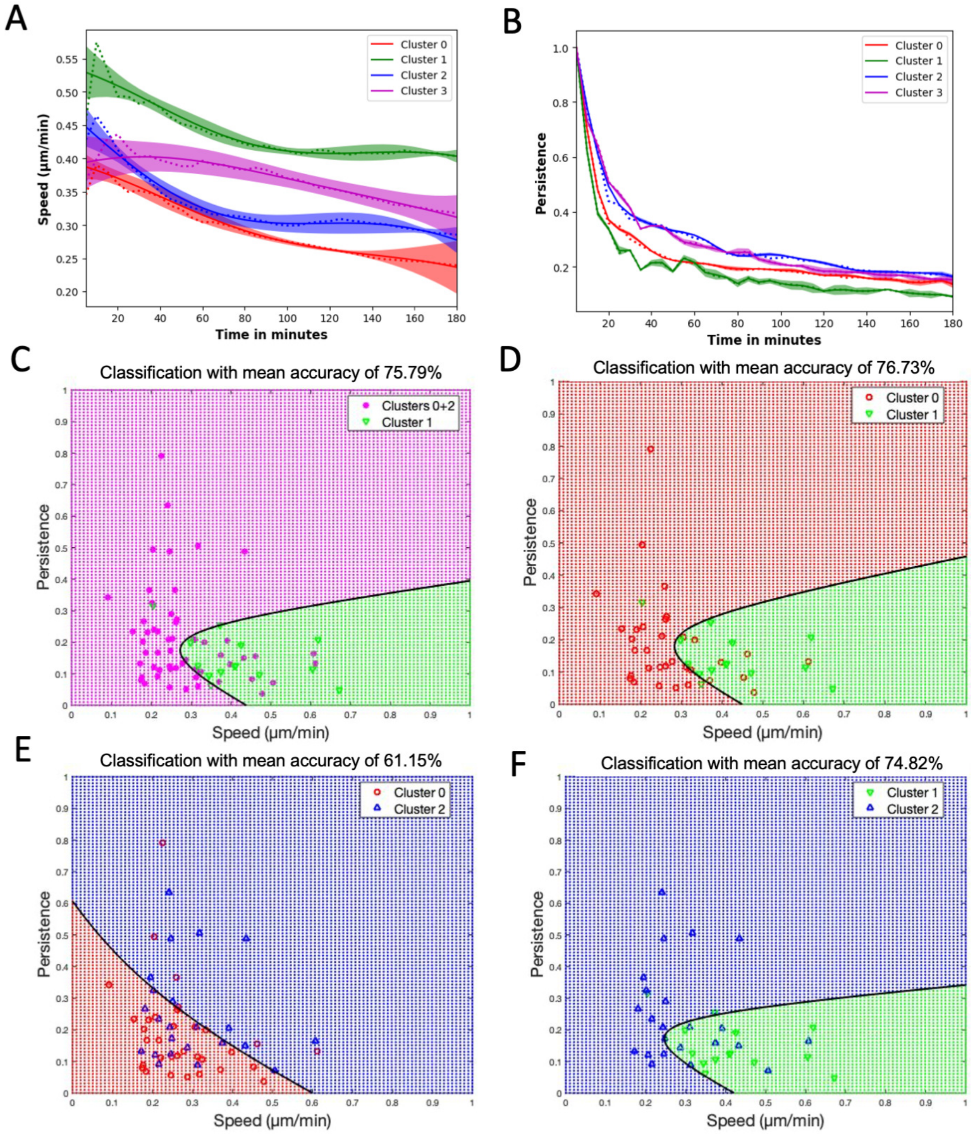


Figure 3.4: **Classification and discrimination of morphological groups using motility parameters.** (A) Speed and (B) Persistence, with dotted lines indicating the mean values of the clusters, solid lines representing the Gaussian Process Regressor (GPR) predictions, and color bands showing the 95% confidence intervals. (C) Quadratic classification of cells using Speed vs Persistence between Cluster 1 and the combined Clusters 0 and 2, achieving a 76% mean accuracy from 5-fold cross-validation. (D) Quadratic classification between Cluster 0 and Cluster 1, with a 77% mean accuracy. (E) Quadratic classification between Cluster 0 and Cluster 2, with a 61% mean accuracy. (F) Quadratic classification between Cluster 1 and Cluster 2, with a 75% mean accuracy. The black lines in C-F denote the decision boundaries between classes. Additionally, the mean classification accuracies are 55% for Clusters 1+2 vs Cluster 0 and 53% for Clusters 0+1 vs Cluster 2.

3.3.5 Motility patterns can be used to differentiate cell morphology-based groups

We would like to show that the chosen displacement metrics of speed and persistence can be used to differentiate cell groups based on morphology. To do this we apply discriminant analysis, a multivariate technique, to separate two morphology-based groups based on measured speed and persistence. This helps to identify the contribution of each variable in separating the groups. Here, we show that a cell's morphological features are in fact linked closely to its motile behavior. This relationship exists independent of the macrophage subtype. For this we leverage each of the unique cells identified above in clusters and quantify their corresponding motility patterns. To investigate this relationship, we identify a minimum set of motility parameters that capture the set of observed migratory behavioral patterns. We compute speed and persistence (a measure to find how far a cell has moved from its initial time step) for all the cells in clusters 0-2 (Fig. 3.4A).

We use the Gaussian process regressor (GPR) model [57, 62] to plot the speed and persistence of the cells. The GPR model utilizes a Gaussian process to model the distribution of the data and predict outcomes based on the observed data points. In Fig. 3.4, each cluster's results are depicted using a distinct color. The GPR model's mean is represented by a solid line, while the average value computed from the data at each time point is indicated by dashed lines. Additionally, the band in the plot represents a 95% confidence interval. Fig. 3.4A indicates that the speed of cells in Cluster 1 (protruded cells) is the highest, followed by the speed of cells in Cluster 2 (elongated cells), and the lowest speed is observed for cells in Cluster 0 (circular cells). Fig. 3.4B displays that the cells in Cluster 2 exhibit the highest persistence, followed by Cluster 0, with the lowest persistence observed for the cells in Cluster 1. Combining the motility parameters analysis with morphological clustering analysis, we observe that circular cells move slower and stay close to their initial time frame position; protruded cells move faster but still stay close to their initial time frame position; whereas elongated cells move slower but reach farther away from their initial time frame position. For Cluster 3 cells, which exhibited indifferent cell groupings or overlapping cells, we can observe that their speed falls between Cluster 1 and the remaining clusters. Also, these cells display a high persistence level, as observed in Cluster 2 cells.

To better visualize and quantify the relationship between cell morphology and motility we perform quadratic discriminant analysis on the motility parameters speed and persistence for the clusters from the morphological clustering analysis. This analysis provides a binary classification based on two input features (i.e., speed and persistence). With this method we can determine to what extent we can predict the cells cluster of origin (e.g., predict morphological characteristics) based on its computed speed and persistence (Fig. 3.4, showing the results for various pairs of classification labels). We perform 5-fold cross-validation and calculate the mean accuracy. Classification accuracy percentages confirm that the Cluster 1(protruded cells) can be distinguished from the Cluster 0 (circular cells) and Cluster 2 (elongated cells) combined or independently using the speed and persistence motility parameters (Fig. 3.4C, D, F). However, the classification accuracy percentage between Cluster 0 (circular cells) and Cluster 2 (elongated cells) is comparatively lower around 61% (Fig. 3.4E). The ability to predict morphological features based on cell motility imply a strong relationship between the two.

3.3.6 Motility characterization can be used to predict macrophage subtypes

We have shown that macrophage subtypes can be predominantly characterized by specific combinations of morphological features. Additionally, we have shown evidence to suggest that there is a relationship between motile behavior and morphology. Thus, we expect that macrophage subtypes should exhibit distinct motile behavior. To demonstrate that motility can be used as a substitute for morphological features, we quantify motile behavior for each cell from the cell tracking results by computing speed and persistence for each macrophage subtype given by the parent image label and taking a supervised learning approach. Based on this motility parameter analysis given in (Appendix Fig. A.7(A, B)), we find that M0 cells move slower and stay close to their initial time frame position; M1 cells move faster and but still stay close to their initial time frame position; whereas M2 cells move slower and reach farther away from their initial time frame position. These motility observations can act as an independent model in the identification of the macrophage subtypes without analyzing the morphology of the cells. Quadratic discriminant analysis is also performed (Appendix Fig. A.7(C-F)) to quantify the motility of the cells in M0, M1 and M2 images. We note that the mean accuracies are the same or worse when compared to those using cluster labels due to the heterogeneity observed for each parent label based on cell culturing protocols. Recall that the three orthogonal modes of shape variability account for 86% of the total variation in shapes observed. Furthermore, approximately 30% of the cells were placed into Cluster 3, which included cells with mixed or ambiguous features. Thus, a classification accuracy of around 70% should be considered formidable when using only morphology. Below we investigate accuracies achieved with motility information alone and combined information, respectively.

3.3.7 Increased prediction accuracy of macrophage subtype is obtained using both morphological and motility features

We note that in the previous analysis some macrophage subtypes are more easily distinguished by motility parameters and others by morphological features. Thus, we expect that combining both morphological and motile behavior can improve our ability to classify macrophage subtypes. The advantage of this would be an ability to classify macrophage subtypes using only phase contrast images. Additionally, we present a generalizable machine learning model to properly compare predictive capabilities based on both motility and morphology independently and combined. We compute eccentricity, compactness, and cell size to observe morphological features, and convex hull area of cell trajectory, speed, and persistence to extract motility features for all cells in M0, M1, and M2 images (Fig. 3.5). For our machine learning classifier, we consider a support vector machine (SVM) [63] model, which can deliver high precision and accuracy regardless of the number of attributes and data instances [60]. We have 1018 cells from all the image sets where 251 are from M0 images, 369 from M1 images and 398 from M2 images. We shuffle the data and use stratified 5-fold cross-validation. During each iteration, the dataset was divided into approximately [200, 296, 318] cells for M0, M1, and M2 classes in training and [73, 80, 51] cells for M0, M1, and M2 classes in testing sets. We now use Synthetic Minority Oversampling Technique (SMOTE) (35) to balance the minority classes M0 and M1. The final training dataset has 954 datapoints, with 318 datapoints from each of the M0, M1, and M2 image classes. We fit the training data using SVM classifier -radial basis function (RBF) kernel and predict the output on the test data where the output of the model is the macrophage subtype given by the parent image label. Fig. 3.5A shows the schematic of the machine learning model with the features, input labels and the expected output. To

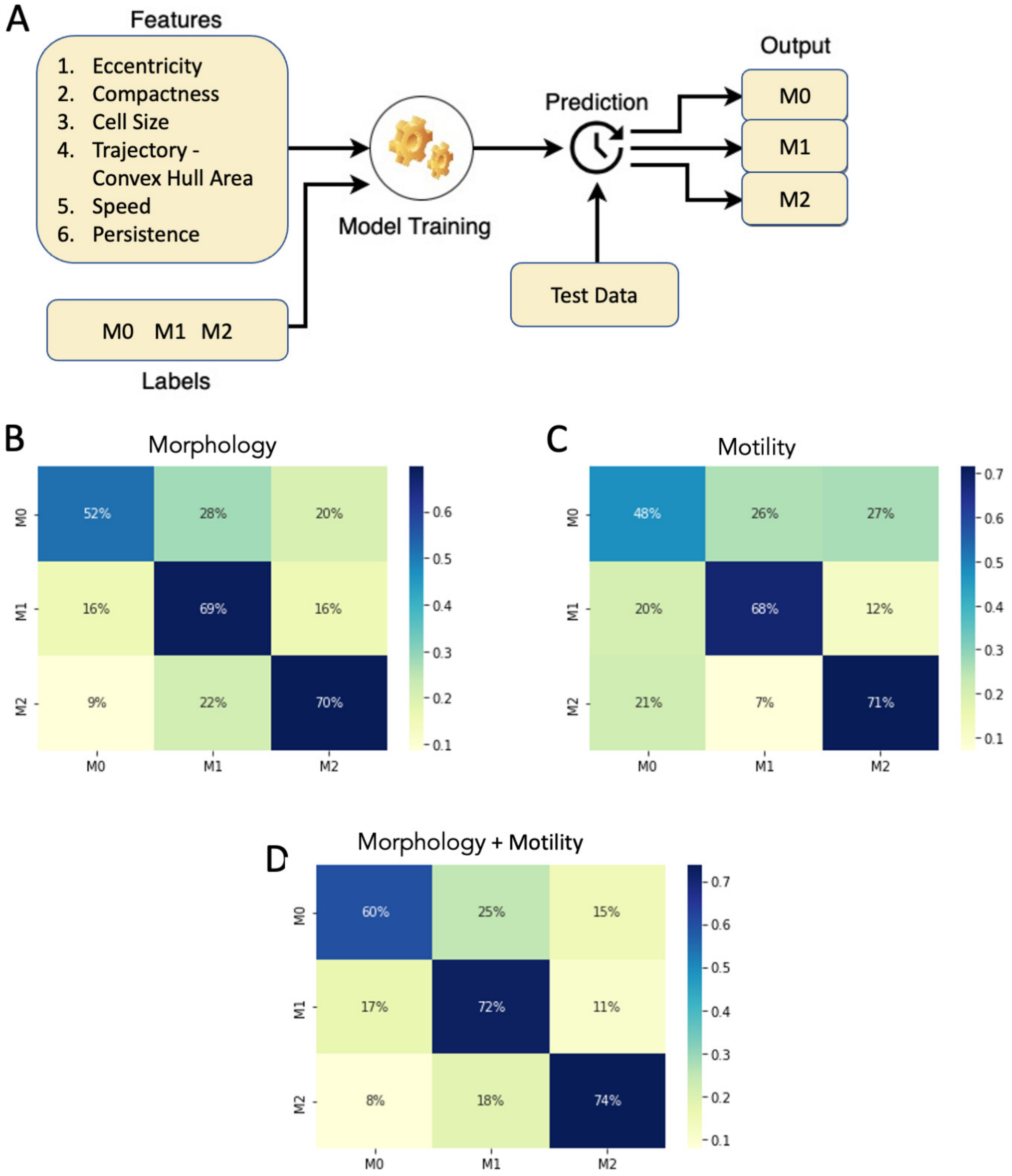


Figure 3.5: **Classification of macrophage subtypes using SVM on morphological and motility parameters - Stratified 5-fold cross-validation plots for the SVM model.** (A) Schematics of the input and output of the machine learning model. SVM classification results showing the percentage distribution of the cells in a confusion matrix with columns being the predicted labels and rows being the actual labels. (B) Input features with morphological parameters (Eccentricity, Compactness, Cell Size). (C) Input features with motility parameters (Convex Hull Area of the trajectory, Speed, Persistence). (D) Input features with both morphological and motility parameters.

demonstrate that the combination of morphology and motility enhances the classifier prediction, we develop three models. The first model is trained using only morphological features (eccentricity, compactness, cell size) as inputs. The second model is trained using only motility features (convex hull area of cell trajectory, speed, persistence) as inputs. The third model is trained using both morphological and motility features as inputs. The results are shown in Fig. 3.5B-D, where the prediction accuracy of the model on test data is plotted in a confusion matrix, where rows represent the actual labels and columns represent the predicted labels. The morphology trained model showed that the prediction accuracies for M0, M1 and M2 labelled cells are 52%, 69% and 70% respectively (Fig. 3.5B). While the motility trained model had the prediction accuracies for M0 (48%) and M1 (68%), M2 cells (71%) (Fig. 3.5C). Strikingly, the combined morphology and motility trained model significantly improved prediction accuracies for all the M0, M1 and M2 cells to 60%, 72% and 74% respectively (Fig. 3.5D). Hence, by using the combined morphological and motility parameters we can predict the macrophage subtypes with an improved accuracy compared to the models which use either morphology or motility alone (Appendix Table B.1). For reference, when analyzing individual biomarkers in flow cytometry, we note that achieving a prediction accuracy of 74% is considered high, owing to variability in biomarker expression influenced by experimental context [64]. Thus, the prediction accuracies obtained are meaningful within a biological context and support the claim that there exists a relationship between cell motility, cell morphology, and macrophage subtypes.

3.4 Discussion

It is well known that macrophages cannot be cleanly grouped into their subtypes, and in fact, various intermediate subtypes exist, yet their artificial classification is important to understanding gene regulatory pathways involved in the various functions carried out by the full spectrum of macrophage phenotypes [65]. Such understandings can help establish ways to target and control their activity to improve health outcomes in applications such as wound healing.

Image-based methods provide a fast and efficient way to classify macrophage subtypes. Previous work has focused on machine learning methods for predicting macrophage subtypes based on morphology [55, 66]. Machine learning effectively substitutes the manual classification of cell phenotype by researchers and clinical practitioners and offloads the decision-making to an algorithm. The reasons automated image processing, segmentation, and analysis have found such interest are that they offer huge possibilities to free up researchers' time and to ensure consistent interpretation of results free from any human bias. Prediction methods that rely on cell morphology require high-quality images to be able to accurately segment cells and characterize their morphology. In this work, we show that spontaneous cell motility patterns can also serve as unique signatures for macrophage subtypes. Furthermore, an impressive amount of cell tracking algorithms [67], benchmark techniques [68, 69] and software [38–40, 70, 71] have been developed and made available to the scientific community which are more ubiquitous and robust to blurry images. However, we note that it can be a complex task to extract the morphology of a cell if it were to be suspended rather than spontaneously moving on a substrate. We also recognize that the motility can be affected by the substrate itself [72]. This means that changing the substrate may require new characterization for predictions. However, this may imply that the different cell types move differently in different environments. We know the extracellular matrix changes over time in wound healing, which allows different cell types to carry out their work.

Independently, motility and morphological features can provide ambiguous discriminant information depending on the macrophage subtype. For example, the prediction accuracies derived from motility-based information closely match those obtained through morphology-based predictions. Thus, combining both cell motility and morphological information can reliably and independently detect non-activated naïve macrophages, or macrophages activated by either lipopolysaccharide (LPS) or interleukin 4 (IL-4) exposure at single-cell level. Furthermore, the numerous metrics applied to characterize the morphology and migration patterns are simplified representations (e.g. information is lost). One may consider exploring ML methods to extract patterns with nuanced information from the raw data.

We note that the images used for analysis required that cell densities remain low. We noted that some of the cells in cluster 3 could likely be clusters of cells that are identified as a single cell due to the segmentation regions formed when the cells are very close to each other. Further improvement of the segmentation algorithm to correctly identify the high-density cell regions and post-processing methods to remove the cells for analysis beyond a particular cell area, might lead some of these cells to get distributed in the above three clusters. Additionally, Cluster 3 also seemed to include cells that had combined features from two or more of the remaining clusters. Thus, Cluster 3 may include macrophages that are in a transient state between subtypes or may represent an ambiguous state not yet explored. The methods in this paper could potentially be used to further understand the full spectrum of macrophages and track potential shifts in macrophage subtypes over time. Furthermore, a future adaptation of this method can be used in bioengineering applications to characterize and quickly screen genetic variants such as in engineered CAR-Macrophages [73].

3.5 Conclusion

In this paper, we present that using image processing techniques, we can characterize cell behavior and migratory patterns that can be fed into our machine-learning model to predict macrophage subtypes. While we recognize that other methods such as flow-cytometry can offer improved accuracy in macrophage classification [64],our method has the advantage of being non-invasive. It does not require cell fixation or labeling, preserving cell viability for further functional assays. Additionally, this approach allows for real-time analysis. The ability to assess macrophage polarization through dynamic migration and morphology changes offers insights into the real-time behavior of cells in response to various stimuli. Our method also has the potential for high-throughput screening. Once fully optimized, it could facilitate high-throughput screening of biochemical effects on macrophage polarization, providing a valuable tool for both basic research and drug discovery. Analysis of these dynamics without any external cues given like electrical guidance and staining would give us insights that are critical to quantifying macrophage recruitment and activation in vivo. In future work, this analysis could also be applied to the images in galvanotaxis to observe the dynamics of the cell behavior with and without electric fields. Furthermore, a direct comparison between fixed tissue analysis and flow cytometry, could enhance understanding of the potential applications of our approach. While this comparison was not conducted in the current study due to its exploratory nature and focus on developing the machine learning model, this is a potential direction for future research. Finally, this method can be adapted to serve other cell types and applications but may require identifying more suitable morphological and displacement metrics where applicable.

4 | Deep Learning Classification for Cell Subtypes through Cell Migratory Pattern Analysis

This chapter is from as it appears in Frontiers Cell Biology. The dissertation author was a first author of this paper [74]. Macrophages can exhibit pro-inflammatory or pro-reparatory functions, contingent upon their specific activation state. This dynamic behavior empowers macrophages to engage in immune reactions and contribute to tissue homeostasis. Understanding the intricate interplay between macrophage motility and activation status provides valuable insights into the complex mechanisms that govern their diverse functions. In the chapter 3, we developed a classification method based on morphology, which demonstrated that movement characteristics, including speed and persistence, can serve as distinguishing factors for macrophage subtypes. In this chapter, we develop a deep learning model to explore the potential of classifying macrophage subtypes based solely on raw trajectory patterns. The classification model relies on the time series of x-y coordinates, as well as the distance traveled and net displacement. We begin by investigating the migratory patterns of macrophages to gain a deeper understanding of their behavior. Although this analysis does not directly inform the deep learning model, it serves to highlight the intricate and distinct dynamics exhibited by different macrophage subtypes, which cannot be easily captured by a finite set of motility metrics. Our study uses cell trajectories to classify three macrophage subtypes: M0, M1, and M2. This advancement holds promising implications for the future, as it suggests the possibility of identifying macrophage subtypes without relying on shape analysis. Consequently, it could potentially eliminate the necessity for high-quality imaging techniques and provide more robust methods for analyzing inherently blurry images.

4.1 Introduction

Macrophages are involved in the body's immune responses and tissue homeostasis. They play a critical role in infectious diseases [75], cancer progression [76], autoimmunity [77], wound healing [78], and many other diseases [79, 80]. Two main subsets of activated macrophages with different functional phenotypes, M1 (classically activated, pro-inflammatory) and M2 (alternatively activated, anti-inflammatory), have been identified [80, 81]. In tumor progression, M1 and M2 macrophages assume distinct roles. Specifically, the M2 subtype, M2d, demonstrates pro-neoplastic characteristics, while M1-like macrophages exert anti-tumor effects [82]. Recent studies have shown that besides M1 and M2 types, a continuum of macrophage subtypes exists [81, 83].

The study of macrophage subtypes plays a crucial role in identifying strategies for disease control [84–87]. Consequently, developing effective methods for detecting macrophage subtypes in vitro is essential.

The conventional method to identify M1 and M2 subtypes involves analyzing multiple cell surface markers, transcription factors, and cytokine profiles, which can be time-consuming and resource-intensive. Furthermore, uncertainty remains about how to identify macrophage subtypes confidently. This is in part due to the existing continuum of states. Recent studies of macrophage cultures led researchers to hypothesize that cell morphology could indicate macrophage activation status [88].

Previous research on the classification of macrophage subtypes using machine learning has been based on fluorescent dyes and cell shape parameters [88–90]. More recently, it was suggested that motility parameters like cell speed could be used to classify macrophage subtypes [91]. These publications show that, although a continuum of phenotypes exists, there indeed are three primary shape modes associated with three distinct phenotypes, respectively (inclusive of the so-called "naïve" macrophages). Furthermore, it was shown that these shape modes are closely linked with predetermined cell motility metrics.

This research aims to gain a deeper understanding of macrophages' migratory patterns and explore the potential of classifying macrophage subtypes based on raw trajectory patterns without relying on cell shape analysis. This could provide more robust methods to analyze blurry images.

We suggest that a holistic use of cell motility information, i.e. a time series of cell coordinates, could enhance the differentiation of macrophage subtypes. We develop a deep learning model that uses cell position over time as input and demonstrate that our model effectively distinguishes between M1 and M2 macrophage phenotypes. This classification method could potentially be used to understand the continuum of states further.

4.2 Macrophages migratory pattern analysis

In this study, we leverage labelled data published in our previous work [91]). In brief, bone marrow-derived macrophages (BMDMs) were isolated and cultured, resulting in an M0 macrophage culture. M0 macrophages were further activated into either M1 (with LPS) or M2 (with IL-4) (see Methods) in correspondence with existing protocols [92, 93]. Time-lapse recording of cell images, segmentation, and tracking of cell trajectories was performed in correspondence with [94]. This global track-linking algorithm links cell outlines generated by a segmentation algorithm into tracks. Tracks are incrementally added to the image sequence using information from the complete image sequence in every linking decision.

Overall, we obtained three videos of single cells and three videos of cell cultures.

Single-cell video sequences were captured for one M0 cell, one M1 cell, and one M2 cell, respectively. Each single-cell video comprises 240 images, with a frame interval of 1 minute, resulting in a total duration of 240 minutes.

Cell-culture video sequences were obtained from a non-activated M0 cell culture, an M1-activated cell culture, and an M2-activated cell culture, respectively. Each cell culture video comprises 37 phase contrast frames, each frame captured at a 5-minute interval, spanning a total duration of 180 minutes.

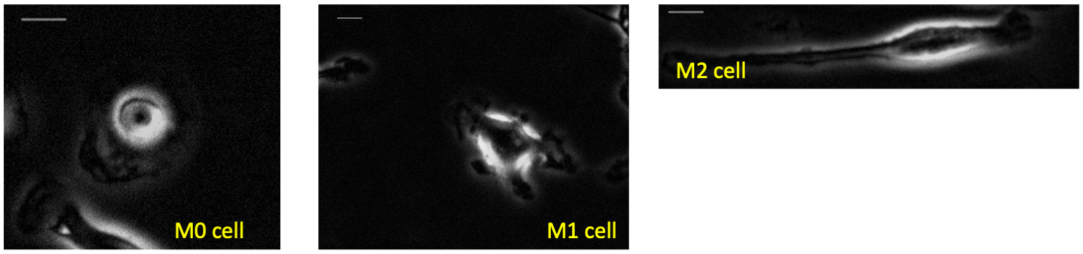
4.2.1 Single-Cell Macrophages images

Each of the three videos of a single macrophage consists of time-variant phase-contrast images. A frame from each video is shown in Fig.4.2a. The M0 macrophage is seen as a circular cell, the M1 cell contains protrusions, and the M2 is an elongated cell. Corresponding differences in cell shape were observed in [91]), and cell-shape-based clustering was created.

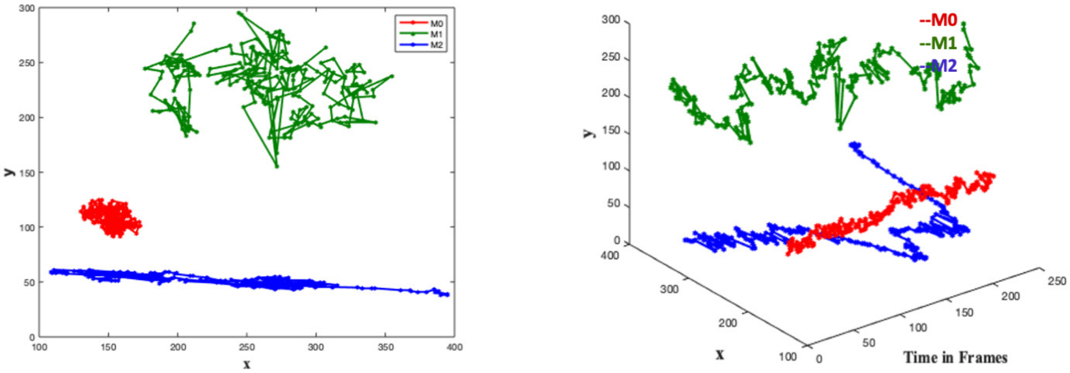
Single Macrophages trajectory analysis

We notice distinct trajectory patterns among the three cell types from the M0, M1, and M2 cell trajectories by plotting the x and y positions of their trajectories in 2D and by representing the trajectories as 3D objects in (x,y,t) space, see Fig.4.2b. It is seen that the M0 cell shows a spinning pattern in the 3D plot and stays close to its initial position in the 2D plot. The M1 cell makes large jumps in each frame in the 3D plot and wanders around its initial point in the 2D plot. The M2 cell moves a greater distance directionally farther away from the initial point in both 3D and 2D plots.

These variations in the cell paths are evident, suggesting the possibility of identifying cell types not only by their shape but also by their trajectory.



(a) Single cell macrophage images of M0, M1 and M2 with an interval of 1 min/frame for 4 hours - one frame from each single-cell video. Scale bar is 10 µm.



(b) Trajectories of M0 (red), M1 (green) and M2 (blue) cells in a single graph with the comparative representation of their respective area spreads in two-dimensional and three-dimensional plots.

Figure 4.2: The images depict single-cell macrophages images alongside their respective trajectories, revealing distinct patterns among the three cell types (M0, M1, and M2). Trajectory analysis involves plotting the x and y positions in 2D and representing trajectories as 3D objects in (x, y, t) space. In the 3D plot, the M0 cell exhibits a spinning pattern and remains close to its initial position in the 2D plot. The M1 cell demonstrates large jumps in each frame in the 3D plot, while meandering around its initial point in the 2D plot. Conversely, the M2 cell covers a greater directional distance away from the initial point in both the 3D and 2D plots.

4.2.2 Analysis of Macrophage Culture Images

In practical settings, cell cultures typically consist of multiple cells present in each frame of recorded videos. The image sets utilized in the present study are labeled according to their culture of predominant macrophage subtypes, i.e., how cell culture was activated (M0, M1, and M2), as shown in Fig.4.3a. Given the macrophages' high plasticity to transform into various functional phenotypes, it is likely that each macrophage can transform from one phenotype to another within each image set. Therefore, the challenge associated with this data set is that the labeled images may contain cells of different subtypes. For instance, M0 images may include cells of both M1 and M2 subtypes, and the same principle applies to M1 and M2 images. Therefore, we utilize the morphological clustering analysis technique described in our previous work [91] to categorize cells based on their shapes. In this analysis, circular cells are assigned to Cluster C, protruded cells (cells with uneven edges) are grouped in Cluster P, and elongated cells are classified in Cluster E. Most macrophages in non-activated cell culture (M0) belong to Cluster C, M1-activated cell culture is represented mostly by Cluster P cells, and M2-activated cell culture is mostly of Cluster E, as shown in Fig.4.3a (see [91] for details of shape-based cell clustering). In the present study, we examine the trajectory patterns of the cells of these shape clusters.

Trajectory analysis of Macrophages from their Shape Clusters

Consider the trajectory patterns of macrophages from the three shape clusters. The trajectories of three typical representatives of Cluster C, Cluster P, and Cluster E are shown within a single plot in Fig.4.3b in order to see spatial differences of cells' trajectories. It is seen that the main difference between Cluster E cell trajectory from the trajectories of the cells of two other clusters is in the more elongated shape of the trajectory. In contrast, trajectories of the Cluster C and P cells are similar in shape but different in the size of the space occupied by the trajectory. Cluster C cell stays closer to its initial location than Cluster P cell.

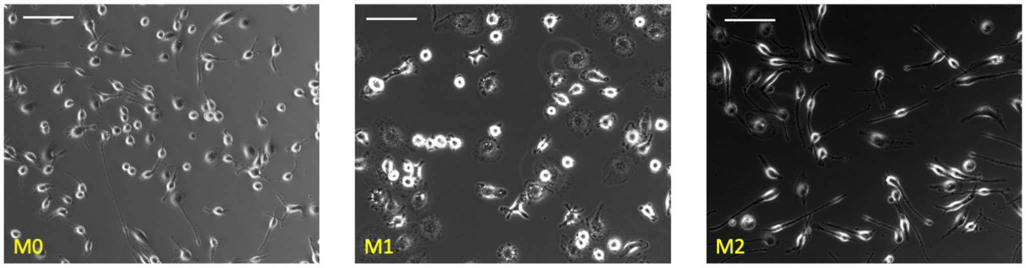
However, the difference in the overall space occupied by the trajectory is not the only unique characteristic. To gain insight into temporal differences of the trajectories, consider Fig.4.3c that shows the trajectory patterns of these representative cells with time encoded by a color gradient from original light (t=0) to dark color (t=180 min). Cluster C cell exhibits a spinning pattern that occupies a smaller area, Cluster P cell demonstrates large jumps and wanders around its initial point, and Cluster E cell tends to move a greater distance in a directional manner away from the initial point. Despite variations in experiment lengths and cell densities between single and multiple-cell videos, we consistently observe similar trajectory patterns among the cells of shape clusters' C, P, and E.

In order to quantify this behavior, we further use several mathematical measures of the trajectory pattern: convex hull perimeter and area of the trajectory, and distance measure: maximum pairwise distance between trajectory points.

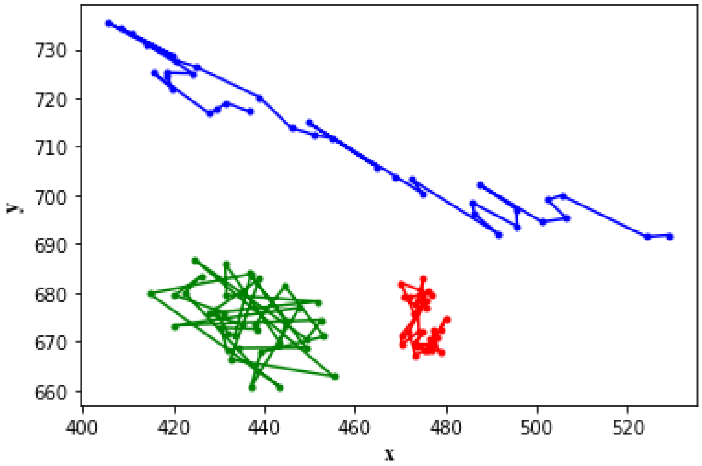
Quantitative measures of the trajectories

Convex Hull Perimeter and Area To understand the spatial differences in Fig.4.3b between the cell trajectory paths of the three cell clusters, we first find the perimeter and area of the convex hull. The convex hull of a set of points is the smallest convex polygon that encloses all the points in the set (Fig.S3). The convex hull area is the total area enclosed by the convex hull, and the convex hull perimeter is the total length of the boundary that encloses the points.

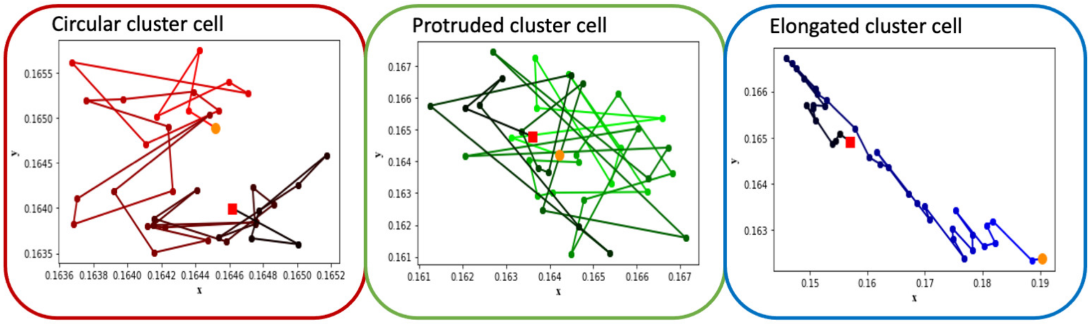
From Fig.4.4, we observe that the convex perimeter of Cluster E (elongated cells) is more significant, followed by Cluster P (protruded cells) and Cluster C (circular cells). The convex area plot clearly distinguishes that the elongated cells have a larger area, followed by protruded and circular cells. Although Cluster E cells can be easily distinguished from the other two clusters of cells using convex hull perimeter and area, the difference between these metrics of Cluster P and Cluster C cells is small, and distinguishing their trajectories is challenging. Hence, we need more specific



(a) Macrophage images of M0, M1 and M2 cell cultures - one frame from each video set. Scale bar is 100 µm.



(b) Trajectory patterns of the representative cells from shape Cluster C (red), Cluster P (green) and Cluster E (blue).



(c) Trajectories of the cells shown in Fig.2b with time frame encoded by color: t=0 corresponds to light color, t=180 corresponds to dark. Starting point is marked as an orange dot and the ending point is marked as a red dot.

Figure 4.3: The depicted images showcase macrophage culture alongside their corresponding trajectories, categorized based on shape-based clusters. Notably, the observed patterns align with those seen in the Fig. 1 images of individual cells representing M0, M1, and M2, which correspond to clusters labeled as circular, protruded, and elongated, respectively.

measures to distinguish the trajectories of the P and C clusters.

Maximum Pairwise Distance As shown in Fig.4.3c, Cluster P cells exhibit larger jumps between frames and show wandering behavior around the initial point. This unique characteristic of Cluster P cells can potentially be used as a distinguishing feature. To identify this characteristic, we use pairwise distance measures. We measure the pairwise distances between each pair of trajectory points and extract the maximum pairwise distance for each cell. From Fig.4.4, it is evident that Cluster P cells exhibit higher maximum pairwise distances than Cluster C cells. We can also observe here that the maximum pairwise distance of Cluster E cells is similar to that of Cluster P cells.

Observations

While the quantitative features mentioned above allow us to make observations about cell migratory patterns, they are not sufficient for reliable detection of macrophage type. For example, the limitations of the convex hull, such as its sensitivity to outliers and lack of consideration for temporal ordering in trajectories, weaken its suitability as a feature of the classification model.

Having observed that macrophage trajectory patterns are specific to each cluster but still are not sufficient for automatic classification, we have developed a deep-learning model in the hope that it can have better classification performance due to capturing some trajectory features that we were unable to detect. This model gets simple characteristics of cell movement as input features and does not require the preliminary calculation of complex metrics.

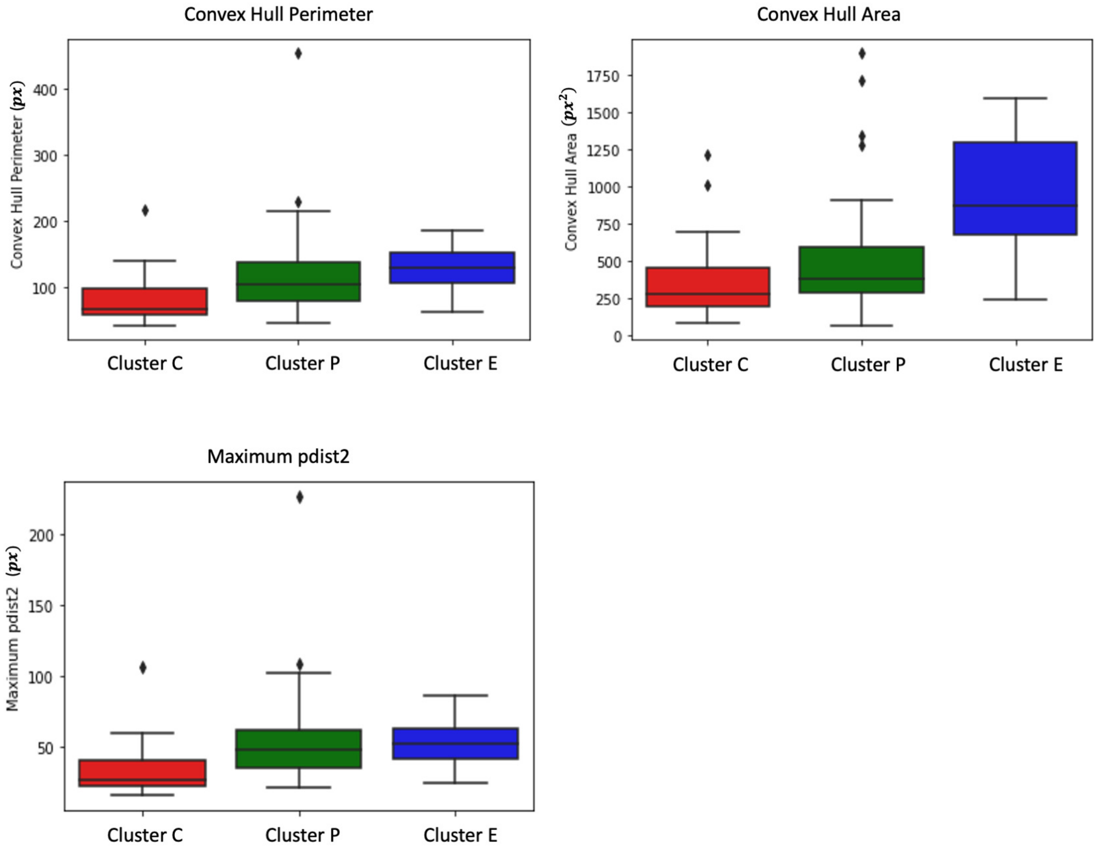


Figure 4.4: Box plots depict measures of convex hull perimeter, convex hull area, and maximum pairwise distance for trajectories from the shape-based clusters. The analysis reveals that Cluster E (elongated cells) exhibits a more substantial convex perimeter, followed by Cluster P (protruded cells) and Cluster C (circular cells). In terms of convex area, elongated cells have a larger area, followed by protruded and circular cells. While convex hull perimeter and area effectively differentiate Cluster E cells, distinguishing between Cluster P and Cluster C cells proves challenging due to the small differences in these metrics. Further examination shows that Cluster P cells display higher maximum pairwise distances compared to Cluster C cells. Interestingly, the maximum pairwise distance of Cluster E cells is similar to that of Cluster P cells.

4.3 Methods

4.3.1 Activation of bone marrow-derived macrophages

In each experiment, bone marrow-derived macrophages (BMDMs) were seeded into six tissue culture-treated well plates at varying densities and cultured in RPMI-1640 medium (Invitrogen) supplemented with 10% Fetal Bovine Serum (FBS) (Invitrogen) and 1× Antibiotic-Antimycotic solution (Invitrogen) overnight. For M1 activation, 100 ng/ml lipopolysaccharide (LPS) (Sigma, Cat number: L6143) was added to the culture medium, while for M2 activation, 20 ng/ml recombinant mouse interleukin-4 (IL-4) (R&D Systems, Cat number: 404-ML) was used [92]. Two days post-stimulation, activated M1 and M2 macrophages were employed for morphological and motility characterizations as well as functional studies. Macrophages that did not receive any stimulation served as M0 controls.

4.3.2 Deep Learning Model for Macrophage classification

We develop a multi-class, single-label classification deep learning model where inputs are: the position coordinates (x, y) of the trajectories in time frames, the distance traveled, and the displacement of the cell. The model's output is the classification of one of the three labels: M0, M1, or M2, assigned to their respective clusters: Cluster C (circular), cluster P (protruded), or Cluster E (elongated) (Fig.4.5).

Train/Test/Validation data:

We assign the label 'M0' to cells from Cluster C, 'M1' to cells from Cluster P, and 'M2' to cells from Cluster E in the training data. We only include cells with a trajectory path from the first frame to the last frame, which is a total of 71 cells. We augment

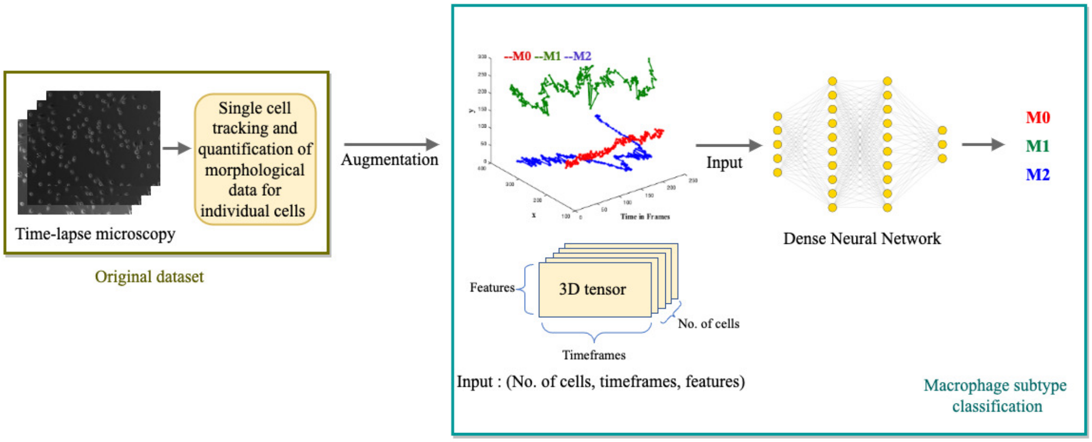


Figure 4.5: Illustration of the Deep Learning Model for macrophage subtype classification. The diagram depicts the process of extracting single-cell tracking and quantification from time-lapse images. The quantified parameters, including (x, y) positions over time, distance traveled, and cell displacement, are augmented and used as inputs for the deep learning model. The model outputs the classified macrophage subtype.

the data set to train the deep learning network. As the input to the model is (x, y) coordinates, we augment the data by inversion (y, x), translation (-x, y), (-x, -y), (x, -y) and inverse translation (-y, x), (-y, -x), (y, -x). These mathematical transformations generate a total of 560 cells. We shuffle the data and split 80% for training(448 cells) and 20% for validation(112 cells).

Architecture:

For the deep learning model, we utilized Keras [95], an open-source neural network library written in Python. Here, the neural networks work to separate three different classes (M0, M1, and M2). Since there are only a few labels to classify, a simple stack of eight fully connected (Dense) layers with Relu activations [96] is used. The hidden units that are passed to each Dense layer are 10, 20, 32, 64, 64, 32, 20, and 10, respectively. The network's final layer is a Dense layer with a size of 3. The network produces a 3-dimensional output vector for every input, where each dimension represents a distinct output class. The softmax activation function is used in this last layer, which generates a probability distribution over the three output classes. Categorical cross-entropy, the recommended loss function for a multi-class classification problem, is used to minimize the distance between predicted and true probability distributions. We used the RMSprop optimizer [97] with its default learning rate to minimize the loss. The evaluation metric used is accuracy, representing the proportion of correctly classified cells. The model is trained for 100 epochs to convergence (see Fig.4.6a).

4.4 Model Results

Given the complexity of the problem of classifying different trajectory patterns and the amount of data available, our results (Fig.4.6a) show that the network is not overfitting and learning effectively. We observe that the training loss decreases over time while the validation loss remains low. Similarly, the training accuracy increases while the validation accuracy remains high. This indicates that the model can generalize well to new data and is likely to perform well on new unseen data. The model's ability to accurately classify the macrophage subtypes is supported by its successful predictions on the validation data. Fig.4.6b represents a confusion matrix visualization where the percentage cell counts of each class are plotted on a heatmap. In this representation, the ground truth data labels "Cluster C", "Cluster P", and "Cluster E" are displayed in the rows, while the model predicted labels "M0", "M1", and "M2" are displayed in the columns. We can observe that the model has an accuracy of 91% in predicting Cluster E (elongated) cells as M2, 95% accuracy in predicting Cluster P (protruded) cells as M1, and 87% accuracy in predicting Cluster C (circular) cells as M0. We can see that the model has identified 12% of the Cluster C cells as M2 and 2% as M1. It is important to note that identifying M0 cells can be challenging, as they may exist in a continuum or may continuously transform into other subtypes.

4.5 Discussion

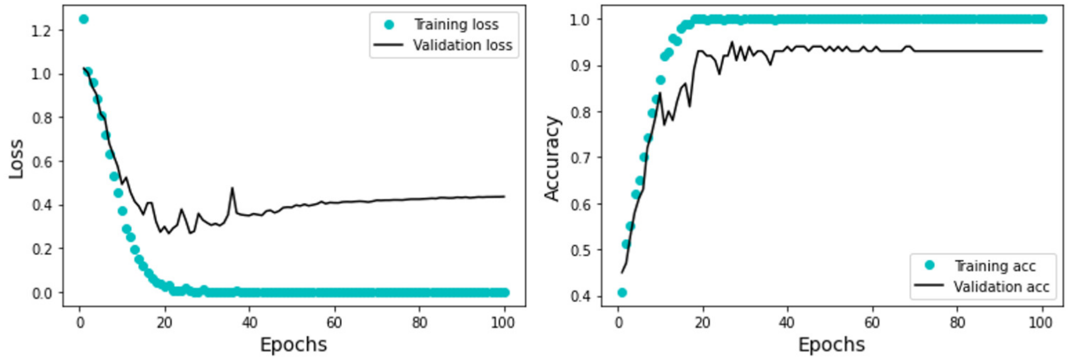
Cell trajectory based and cell shape based methods can potentially classify macrophage subtypes more accurately at the single-cell level. This differs from other methods, such as q-PCR, that give information about cell population subtypes.

As we observed in the cell morphology analysis [91], each cell culture (M0, M1, or M2) is dominantly represented by a corresponding morphology cluster: M0 cells are circular, M1 cells are protruded, M2 cells are elongated, but the shape-based clusters were nonetheless diluted.

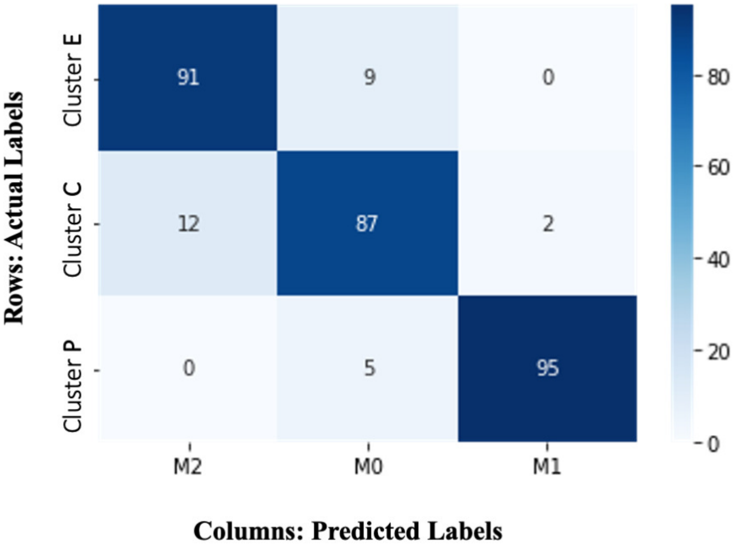
Current analysis of macrophage trajectories revealed typical cell movement patterns corresponding to cell shape clusters: circular cells spin at their initial position, protruded cells wander, and elongated cells move in a specific direction. However, trajectory-based clusters do not exactly correspond to shape-based clusters. This implies a variation in the resulting cell phenotype of macrophage culture activated into M1, M2, or non-activated (M0).

We also note that it is challenging to arrive at a set of metrics that holistically capture all the shape and movement features of the cell.

Several image-based models for macrophage type identification have been published previously. In [89], the authors created a model for noninvasive distinguishing between cell types employing activated and non-activated macrophages for testing. The classification model was linear, allowing greater biological interpretation. However, it required multiple sources of information as input, including phase microscopy images, Raman spectra, and autofluorescence microscopy. In [88], the authors used image-based machine learning approach to classify M1, M2, naive macrophages, and



(a) Training and validation loss and accuracy over 100 epochs.



(b) A heatmap confusion matrix, showing the percentage accuracy of the validation data for the M0, M1 and M2 macrophage subtypes.

Figure 4.6: The Deep Learning model results display the training and validation loss and accuracy over 100 epochs. The accompanying heatmap confusion matrix illustrates the percentage accuracy of the validation data for the M0, M1, and M2 macrophage subtypes. In this representation, the ground truth data labels "Cluster C," "Cluster P," and "Cluster E" are presented in the rows, while the model-predicted labels "M0," "M1," and "M2" are presented in the columns. Notably, the model demonstrates 91% accuracy in predicting Cluster E (elongated) cells as M2, 95% accuracy in predicting Cluster P (protruded) cells as M1, and 87% accuracy in predicting Cluster C (circular) cells as M0.

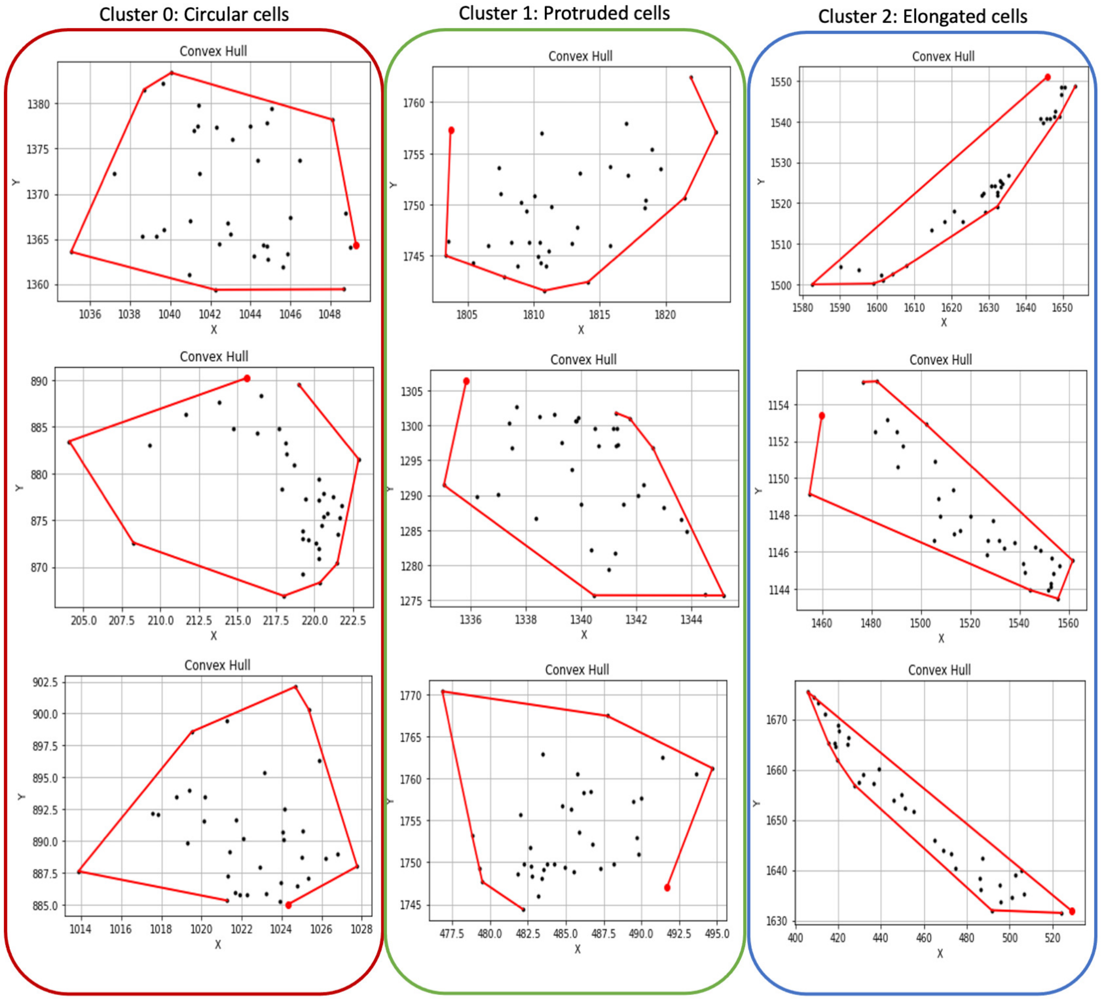


Figure 4.7: The red boundary in each plot represents the convex hull of a set of points, denoting the smallest convex polygon that encloses all the points in the set. These plots correspond to cells in Cluster C, P, and E, representing circular, protruded, and elongated cell types, respectively.

monocytes. Their algorithm relied on various cell metrics, including details related to the nucleus and cytoskeleton, and requires high-quality images.

Our deep learning model relies on the raw cell trajectory data in the x-y plane, as well as the cell displacement and traveling distance. It automatically captures the features of cell motion unseen by the human eye and demonstrates good accuracy.

4.6 Conclusion

Our study revealed unique migratory patterns and distinct morphology in the three subtypes of macrophages. M0, M1, and M2. By analyzing their trajectories and computing various quantitative measures, such as perimeter, area of the convex hull, and pairwise distances, we observed clear differences in the migratory patterns of these macrophage cell types. However, building a classification model with predetermined features is challenging in this context.

Therefore, we take a different approach by developing a deep learning model that incorporates the trajectory path and shape of cells, which proves to be more effective in accurately classifying macrophage subtypes. The correlation between cell shape and trajectory patterns can be highly valuable in future scenarios where obtaining precise cell morphology data is challenging. Additionally, identifying cells based on their migration patterns through phase-contrast microscopy has the potential to eliminate the requirement for high-quality imaging and provide more reliable methods for analyzing blurry images.

5 | Application of Image Processing Methods for Real-Time Feedback Control of Cell Migration and Classification of Cell Subtypes Under an Electric Field

Proper cell migration is vital for the health of multicellular organisms. As biomedical sciences advance, controlling cell migration may become a part of select medical procedures, which can be achieved by inducing an electric field with bioelectronic devices. Maximizing outcomes requires controlling biological systems, which is challenging due to innate uncertainties, nonlinearities, and stochasticity, making predicting their response difficult.

In the first section of this chapter, a feedback control algorithm based on microscopy images is developed to direct cell migration in real-time. This section highlights the advantages of efficient computational methods for image analysis in real-time control of biological systems under electric fields. Here, a version of the controller is demonstrated *in vitro*, where the recruitment of macrophages is controlled in an experimental setting by regulating the electric field. This is successfully set up by coupling the feedback control algorithm with image processing and cell tracking software using time lapse microscopy. In the second section of this chapter, we explore the effects of the electric field on the transformation of macrophage subtypes and present a preliminary study on classifying these subtypes. A deep learning model was developed for this classification under the influence of the electric field (EF). We also show evidence of potential changes in macrophage phenotypes when exposed to EF based on changes in morphological features. Thus, we explore methods to classify macrophage phenotypes in real-time, similar to Chapter 4, but during galvanotaxis.

The first section of this chapter is currently being prepared for resubmission for publication. The dissertation author was a co-author of this paper, contributing to the design and development of the image analyzer in a closed-loop system and the quantitative analysis of cell tracking properties. This image analyzer is an integral part of the experimental setup, where real-time images were read and analyzed, and the data was provided to the feedback controller.

5.1 Controlling Cell Migratory Patterns under Electric Field

5.1.1 Introduction

Regulation of cell migration plays a critical role in many biological processes essential to life. The role of cell migration is most prominent and studied in development, homeostasis, and disease [98]. Deregulation of cell migration can lead to autoimmune diseases and cancer [98]. Recent work suggests that such processes, when malfunctioning, can be externally regulated through feedback control facilitated by the integration of biological sensors and actuators [99,100]. Examples of sensors and actuators include bioelectronic devices, which have paved the way for humans to control the responses of biological systems in favorable ways [101, 102]. Bioelectronic devices allow for monitoring of blood sugar levels, controlling stem cell fate, applying electrical stimulation and delivering therapeutic drugs [103–106].

An effective signaling cue for cell migration that can be achieved and regulated automatically with current technology is the electric field (EF). Directed cell migration through electrical cues occurs naturally during wound healing [107, 108]. Smart bandages have successfully expedited wound healing in mice [109, 110] by artificially enhancing the EF through an additive external cue. In vitro studies further support the use of EF to control cell migration. The authors in [111] created a device platform to leverage the use of an EF to direct the movement of keratinocyte cells in a two-dimensional (2D) space using predetermined EF signals. The results were promising in controlling the cells' movement in an open-loop fashion. However, the presence of an EF can also result in undesired system responses, including phenotypic changes [112]. Thus, it is important to be able to regulate EF carefully to maximize target response while minimizing off-target effects. This careful regulation can be achieved through feedback control mechanisms integrated with image processing techniques, allowing for real-time monitoring and adjustment of cell migration.

5.1.2 Materials and methods

Experimental set up

Figure 5.1 shows the schematic of the experimental set up used in this work. Macrophages were seeded in a tissue culture dishes (Corning)-based electrotaxis chamber for the best in-vitro migration performance. The chamber was built with glass strips and sealed with high vacuum grease as previously described [113] A Direct Current (DC) electric field (EF) with a voltage of up to 3 V/cm across the chamber was applied through Ag/AgCl electrodes for inducing electrotactic cell movement. The electric current was monitored and controlled by the ML controller to achieve efficient electrotaxis at will. The power system consists of a Keithley current source. The current source transfers the current through Agar bridges and Steinberg's solution. This causes an electric field to be applied to the macrophages in the electric field chamber where a microscope takes an image of the cells at 5 minute intervals. The images are stored on a computer where MATLAB script is running that calls on the image analyzer to evaluate the cells recruitment index value. This value is feedback into the ML controller algorithm to decide the next current value to keep the recruitment index at the desired trajectory. The value is sent to the current source through a serial cable that connects it to the computer.

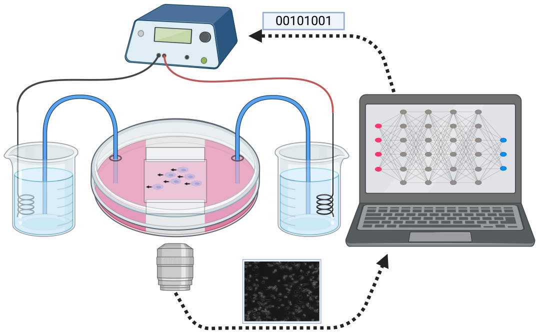


Figure 5.1: Schematic depicting the experimental set up.

Preparation of M0 macrophages

Bone marrow-derived macrophages (BMDMs) were generated and purified *in vitro* (see Fig.5.2) following standard procedures as previously described [114]. Briefly, bone marrow cells were isolated from the tibia or femur of C57BL/6 mice and cultured in (Dulbecco's Modified Eagle Medium) DMEM (Invitrogen) with 10% Fetal Bovine Serum (Invitrogen) and 1× Antibiotic-Antimycotic solution (Invitrogen), supplemented with 20% l-929 conditioned medium containing M-CSF for 6 days, followed by an extra 24-h culture without the conditioned medium. Adherent macrophages were then harvested by gentle scraping with a cell scraper and seeded in electrotaxis chambers for subsequent experiments. Cell viability was determined by trypan blue staining. C57BL/6 mice were purchased from Jackson labs and maintained under a strict 12-h light cycle and given a regular chow diet in a specific pathogen-free facility at the University of California, Davis (UCD). All animal experiments were performed in accordance with regulatory guidelines and standards set by the Institutional Animal Care and Use Committee of UCD.

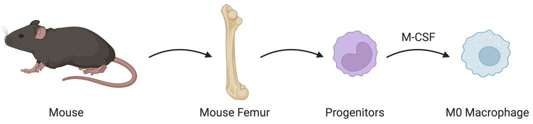


Figure 5.2: Schematic depicting how macrophages were prepared.

Image processing and cell tracking

The image analyzer (see Fig.5.3) is a Python script to evaluate the cells' recruitment index value called from the MATLAB script (as described in Section 5.1.2-5.1.2). The input to the image analyzer is microscope images acquired at 5-minute intervals. The cells in the images are identified and tracked over time using Trackpy [12], a particle-tracking Python package in 2D and higher dimensions. The resulting cell trajectories are analyzed to compute the recruitment index values. The computed recruitment index values are saved into a CSV (comma-separated values) file which is then used to provide feedback to the ML controller algorithm.

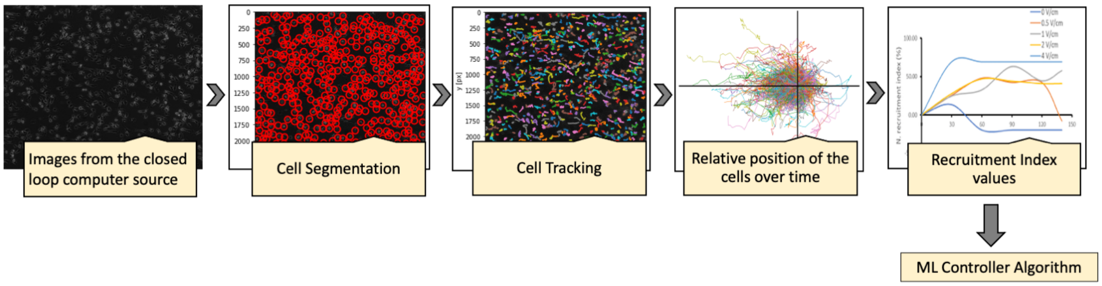


Figure 5.3: Schematic depicting the stages of Image Analyzer

Quantifying cellular response

The cells' relative $\hat{x}$ and $\hat{y}$ position values at every time step $\hat{t}$ are obtained from the cell trajectories and calculated based on the input image resolution. The migration speed of the cells is the ratio of accumulated distance to the time at every 5-minute, calculated as

$$\text{Migration speed} = \frac{\sum_i^n \sqrt{\hat{x}_i^2 + \hat{y}_i^2}}{i \times 5} \tag{5.1}$$

Cells that exhibit migration speeds below the 25th percentile are filtered out in order to eliminate immobile cells. The directedness of cell migration, or the angle of migration, is quantified by calculating the cosine of the angle between the electric field and the line connecting the centroid of the cell from its initial to its current location. Cell directedness is calculated for every 30 minutes, i.e., 6 frames ahead as

$$cos\theta_i = \frac{\hat{x}_{i+6} - \hat{x}_i}{\sqrt{(\hat{x}_{i+6} - \hat{x}_i)^2 + (\hat{y}_{i+6} - \hat{y}_i)^2}} \tag{5.2}$$

where $\hat{x}_i$ and $\hat{y}_i$ are the relative position values of a cell at time $\hat{t}_i$.

To quantify the percentage increase of the cells by each different EF, we define a measure called Recruitment Index (RI). The RI value is fed back to determine the error closing the loop calculated using the directedness values. The RI is normalized (subtracted) to the time point 0. If all the cells move to the anode, the index is 100%, whereas the index is -100% to the cathode. The Recruitment Index is calculated using

$$\text{RI} = \frac{\text{Cells to the Anode} - \text{Cells to the Cathode}}{\text{Total cell count}} \times 100 \tag{5.3}$$

where

- Cells to the anode are cells with directedness > 0.01
- Cells to the cathode are cells with directedness < -0.01

- Total cells are the sum of the cells in anode and cathode, including the cells with $-0.01 <$ directedness < 0.01

5.1.3 Results

Quantifying cellular response in an open loop system

The macrophage (M0) data, collected under electric fields ranging from 0V/cm to 4V/cm in an open-loop setup, reveal that macrophages exhibit high plasticity and unpredictable behavior, particularly in naive states. In our study, the EF levels were based on measurements from human and mouse skin wounds, which typically generate electric fields of several hundred mV/mm [115]. The maximum applied strength of 3 V/cm corresponds to approximately 6 volts across a 2 cm electrotactic channel, or up to 300 mV/mm. We ensured accuracy by measuring the voltage drop with a voltmeter before and after each experiment. In our setup, the relationship between voltage and current adheres to Ohm's law.

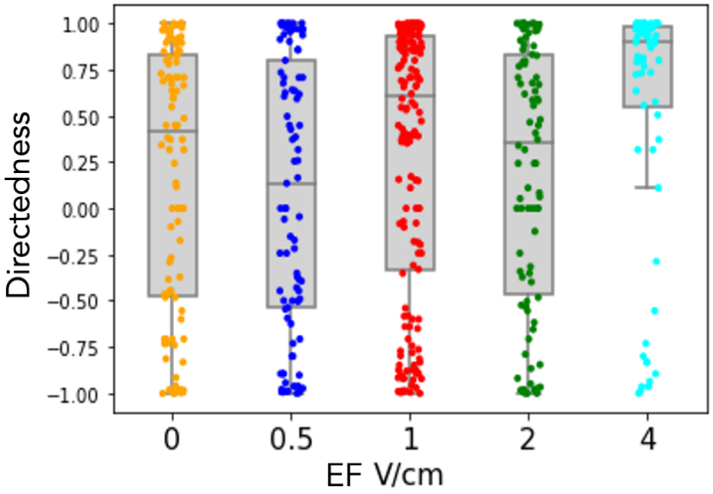


Figure 5.4: Directedness of the macrophages under varying EF.

The current during our experiments ranges from several hundred microamperes to

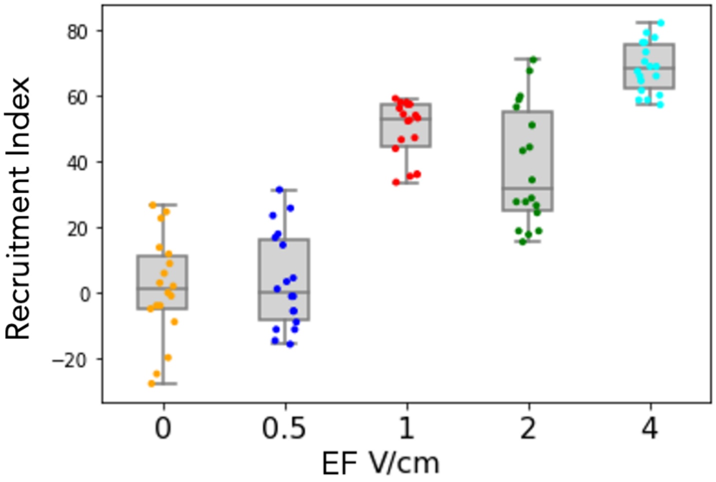


Figure 5.5: Recruitment Index of the macrophages under varying EF.

several milliamperes, depending on the conductivity of the medium and the specific chamber used. We regularly measure or monitor the current, and while there may be some variations, particularly during extended experiments, both voltage and current generally remain stable. Despite observable biases in the distribution, the cells' movement varies in both directions, as shown in the directedness plot in Fig. 5.4. Directedness is not a controllable parameter in this context. However, the recruitment index appears more controllable, displaying a nearly monotonic response to the applied electric field (EF) as shown in Fig. 5.5.

Feedback control algorithm in *in vitro*

Here, we demonstrate an *in vitro* implementation. We use a qualitative stochastic model of directedness for the *in vitro* analysis to demonstrate the advantages of using the feedback control algorithm with projection [116]. This algorithm is tested in an experimental setup and compared to a PID controller. The PID controller was chosen for comparison given that it is an industry standard. Figure 5.6 shows experimental results of feedback control on the recruitment index of macrophage M0 cells using

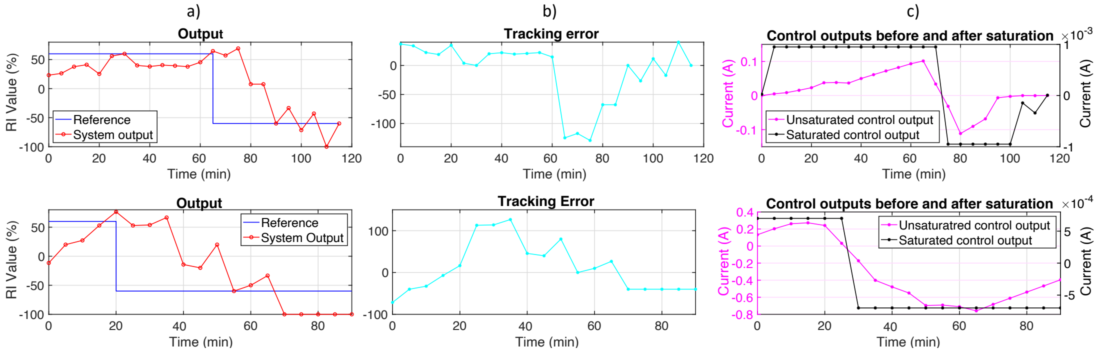


Figure 5.6: The plots in the top row are the results for the proposed ML algorithm while the bottom plots are results using a PID algorithm. Plots in column (a) show the experimental results for feedback control on the recruitment index of macrophage M0 cells using the purposed ML algorithm and a PID controller. Once the initial positive reference value of 60% RI, in blue, is surpassed, the reference changes to -60% RI, and the goal is to track the reference value from there on out. The red is the measured recruitment index value of the M0 cells during the experimental run. Plots in column (b) show the tracking error in cyan. The plots in column (c) show the saturated control output applied to the cells, in black, and the control output without the saturation limits, in magenta.

the purposed ML algorithm (top three figures) and a PID controller (bottom three figures). The initial goal was to have the recruitment index exceed the reference value of 60% RI. Once this was achieved, the reference value changed to -60% RI, with the goal being to track the reference value. In the plot titled output, the blue line is the reference value, and the red is the measured recruitment index value of the M0 cells during the experimental run (Figure 5.6-a). The plot titled tracking error shows the error in cyan (Figure 5.6-b). The plot titled control outputs before and after saturation shows the saturated control output applied to the cells and the control signal derived by the controller before the saturation bounds were applied (Figure 5.6-c).

Once the initial goal was met, however, the controller quickly changes signals to achieve the changed reference value. When saturation values are reached, the control output starts to get pushed back toward the bounded limits of the experimental setup. We see that the final three values applied by the saturated control output are the same as the values decided on by the controller. Further details of the controller algorithm can be found in [116].

5.1.4 Discussion

The methods developed in this work have significant implications for wound healing. M1 macrophages are abundant during the early stages of inflammation, aiding in wound cleaning, while M2 macrophages are more prevalent during the later stages, supporting tissue regeneration [117]. By leveraging image processing techniques, it is possible to monitor and control cell migration in real-time. For instance, an external electric field could be modulated to recruit M1 or M2 cells at specific stages, potentially accelerating the wound healing process. This approach makes it possible to monitor cell migration *in vivo* and in real-time by tracking biomarkers or chemical compounds in the wound bed as proxies for M1 and M2 abundance and activity.

These image-processing methods can also be applied across various biological systems, enabling real-time feedback control of cell migration. Self-regulation is a critical aspect of system health, as seen in examples like the body's response during exercise, where increased oxygen demand triggers vasodilation to enhance blood flow to muscles [118]. Similarly, in wound healing, different cells are deployed in a coordinated manner to facilitate recovery [119]. Integrating image processing with feedback control algorithms allows for precise monitoring and adjustment of cell migration, enhancing the effectiveness of targeted therapies and wound healing.

5.1.5 Conclusion

We developed an image analyzer within a closed-loop system to control cell migration. This analyzer measures cell speed, directionality, and the percentage of cell movement in response to varying electric fields. The controller uses this data to monitor and adjust the electric current, applying an electric field to macrophages in the chamber. The controller was used *in vitro* to guide macrophage M0 cells along a desired trajectory. A microscope captures images of the cells every 5 minutes, which are processed by the image analyzer, creating a real-time closed-loop system. This application demonstrates the importance of analyzing phase contrast images, allowing real-time cell dynamics analysis and control.

5.2 Deep Learning to Classify Cell Subtypes under Electric Field

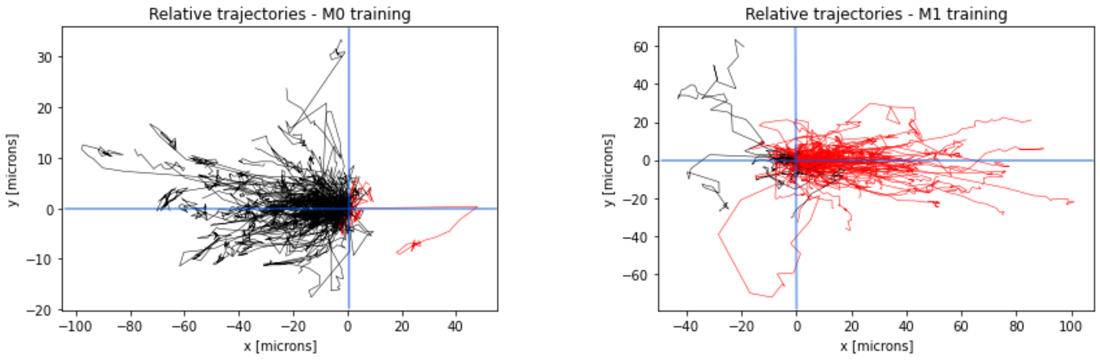
A key question is whether the electric field affects the transformation of macrophage subtypes. Since the electric field is a simple, non-invasive treatment, understanding its impact could make it a valuable tool for controlling processes involving specific macrophage subtypes. Classifying macrophage subtypes under the influence of electric fields represents an under explored area requiring substantial research. This section aims to conduct a preliminary investigation into categorizing macrophages subjected to electric fields.

5.2.1 Introduction

We acquired time-series images of macrophages in their M0, M1, and M2 states exposed to electric fields for durations ranging from 2 to 5 hours. By classifying these macrophages under electric field exposure, we seek to determine whether electric fields induce changes in macrophage phenotypes. Understanding these effects could provide insights into controlling macrophage behavior during wound healing, potentially leading to strategies that accelerate wound healing.

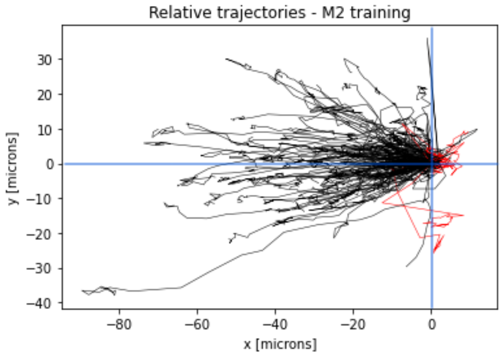
5.2.2 Materials and methods

The bone marrow-derived macrophages (BMDMs) were obtained similarly to the previous experiments - M1 (LPS and IFNgamma for 24 hours) and M2 (IL4 for 24 hours), details provided in chapter 3. The time-lapse movies were recorded in an interval of 5 min per frame for about two to five hours. An electric field strength of 2V/cm is applied to the macrophages.



(a) M0

(b) M1



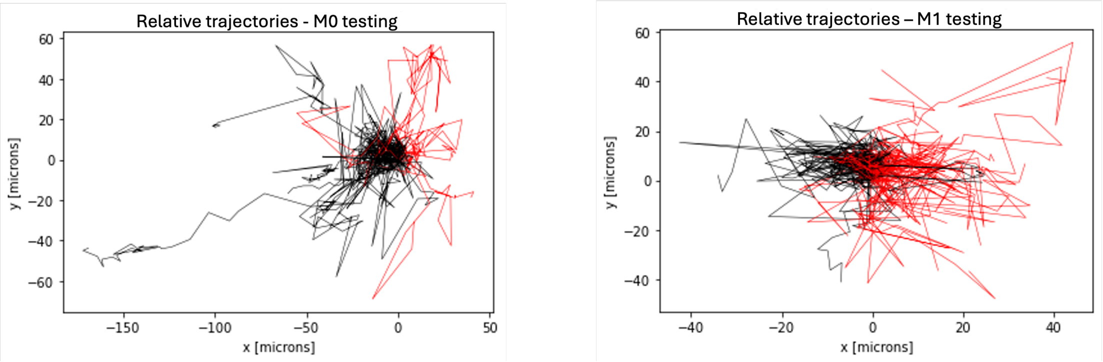
(c) M2

Figure 5.7: Relative trajectories of the training dataset of M0, M1 and M2 macrophage images under an EF strength of 2V/cm. The trajectories shown in black represent the cells moving along the negative x-axis, while the trajectories in red indicate the cells moving along the positive x-axis.

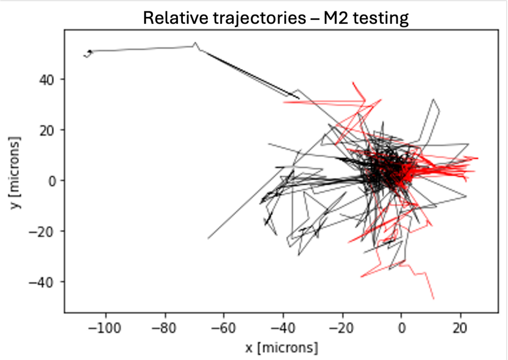
5.2.3 Results

Data

The single cells in the time-series images were segmented and tracked over time using Trackpy [12], a particle-tracking Python package capable of handling 2D and higher-dimensional data. The cells' relative $\hat{x}$ and $\hat{y}$ position values at each time step $\hat{t}$ were obtained from the cell trajectories and calculated based on the input image resolution.



(a) M0 (b) M1



(c) M2

Figure 5.8: Relative trajectories of the testing dataset of M0, M1 and M2 macrophage images under an EF strength of 2V/cm. The trajectories shown in black represent the cells moving along the negative x-axis, while the trajectories in red indicate the cells moving along the positive x-axis.

For this analysis, we used two batches of datasets derived from separate groups of mice—one for training the model and the other for testing. The training data included two sets of M0, M1, and M2 images: in total, 425 cells were tracked: 58 cells from M0 set 1, 71 cells from M0 set 2, 69 cells from M1 set 1, 83 cells from M1 set 2, 78 cells from M2 set 1, and 66 cells from M2 set 2. The combined trajectories of training dataset for M0, M1, and M2 cells are illustrated in Fig.5.7. No additional normalization or data augmentation was performed. From the testing dataset 33 cells from M0 image, 46 cells from M1 and 36 cells from M2 images were tracked. The trajectories of testing dataset for M0, M1, and M2 cells are illustrated in Fig.5.8.

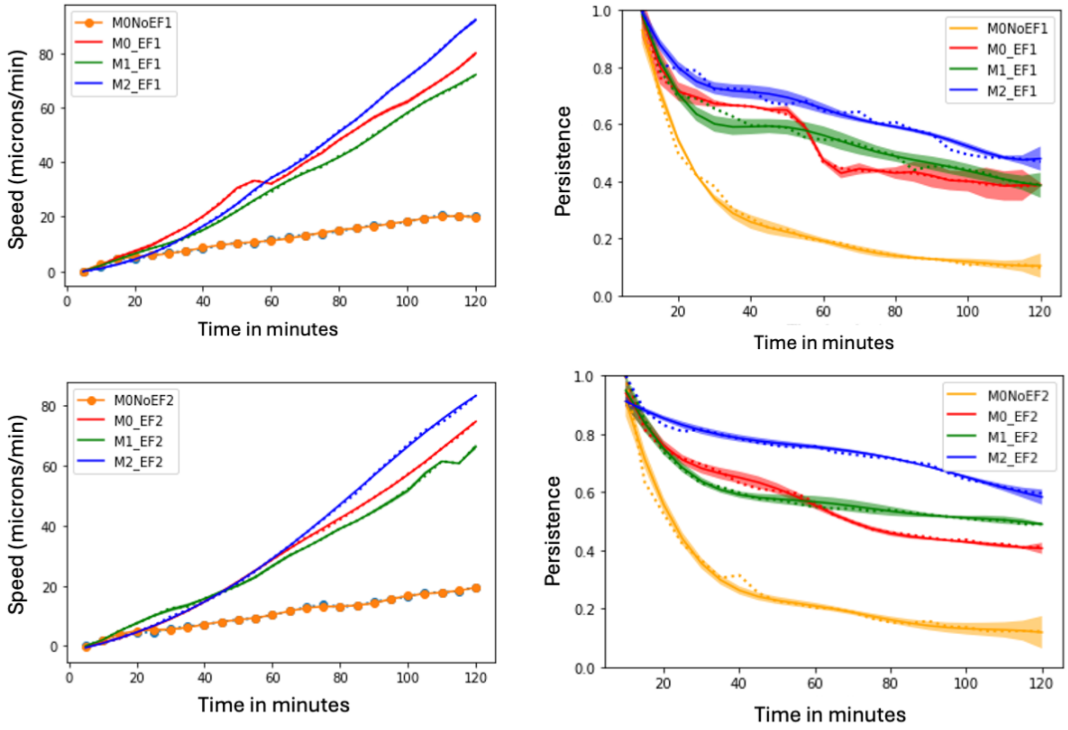


Figure 5.9: In each row, the two columns display speed and persistence plots. In these plots, dotted lines represent the mean values for the corresponding macrophage images, solid lines denote the predictions made by the Gaussian Process Regressor (GPR) model, and the color bands illustrate the 95% confidence intervals.

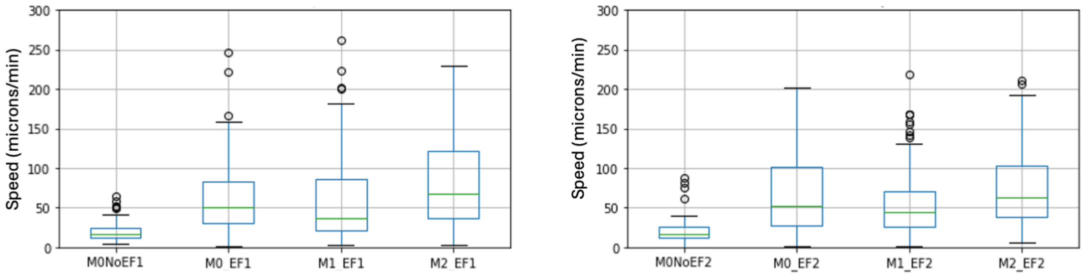


Figure 5.10: This figure features box plots comparing the mean speed of cells from the two training sets to M0 cells without EF, as well as M0, M1, and M2 cells subjected to an electric field of 2V/cm.

Motility analysis of the cells

We quantify the motile behavior of each cell by computing speed and persistence for each macrophage subtype as indicated by their parent image labels. Speed and persistence (a measure of how far a cell has moved from its initial time step) are calculated for all M0, M1, and M2 cells from training and testing images. To visualize the data, we use a Gaussian Process Regressor (GPR) model, which uses a Gaussian process to model the data distribution and predict outcomes based on observed data points. In Fig.5.9, the results for each image are depicted with distinct colors. Solid lines represent the GPR model's mean, while dashed lines show the average value computed from the data at each time point. The plot's band indicates a 95% confidence interval.

Fig.5.9 and Fig.5.10 show that the speed of the cells under an electric field (EF) is higher than without EF, as expected. There is no significant difference in speed among the three macrophage types. However, the persistence of M2 cells appears higher compared to M1 and M0 cells under EF in the training sets. Whereas in Fig. 5.11, no notable difference between M1 and M2 cells is observed in both speed and persistence parameters.

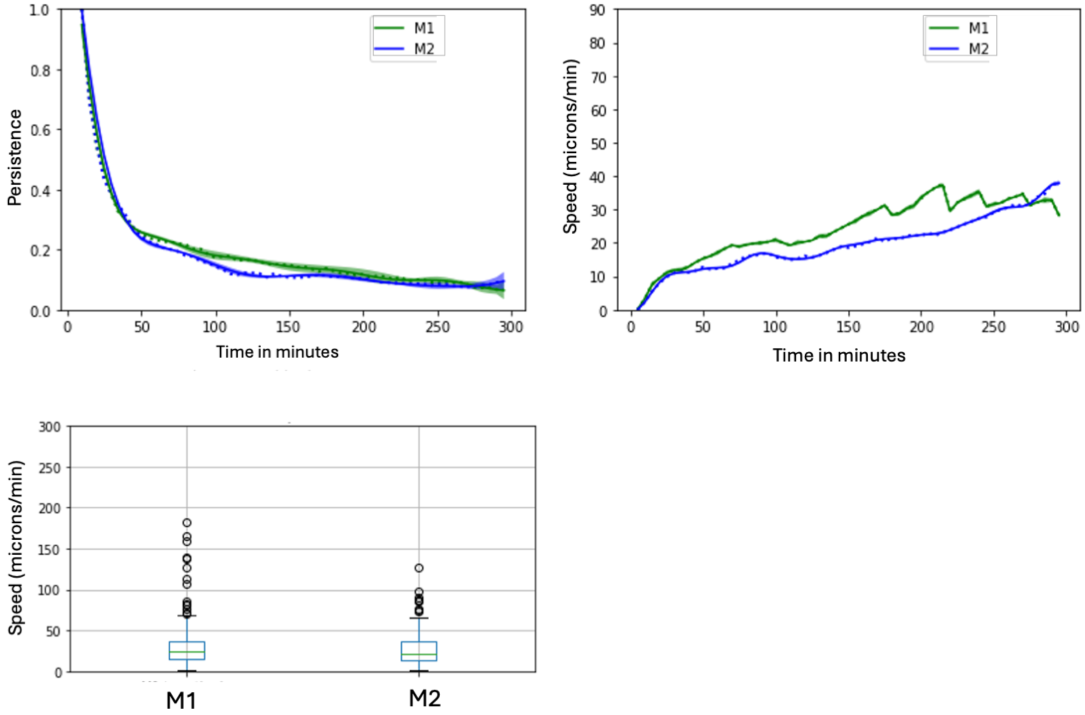


Figure 5.11: The first row displays speed and persistence plots. In these plots, the dotted lines represent the mean values for the corresponding macrophage images, the solid lines denote the predictions made by the Gaussian Process Regressor (GPR) model, and the color bands illustrate the 95% confidence intervals. The second row features a box plot comparing the mean speed of cells in the test set of M1 and M2 cells exposed to an electric field of 2V/cm.

Model

An LSTM (Long Short-Term Memory) model is created to categorize different types of macrophages using trajectory data in the form of (x, y) coordinates over a specific period. The model's architecture consists of:

- Two LSTM layers, each with 64 units.
- One bidirectional LSTM layer with 4 units.
- A flatten layer to prepare the data for the dense layers.
- A dropout layer with a dropout rate of 0.5 to prevent overfitting.

- A dense layer with 128 units and ReLU activation.
- An output-dense layer with softmax activation for classification.

The model is compiled using categorical cross-entropy as the loss function and the adam optimizer. Accuracy is used as the performance metric. The model is trained for 20 epochs with a batch size of 32. For training and validation, the training dataset is split into 70% training and 20% validation using StratifiedShuffleSplit to ensure that the relative class frequencies are approximately preserved in each fold.

As shown in Fig.5.12, the model successfully predicted the M0, M1, and M2 classes in the validation set with accuracies of 82%, 90%, and 89%, respectively.

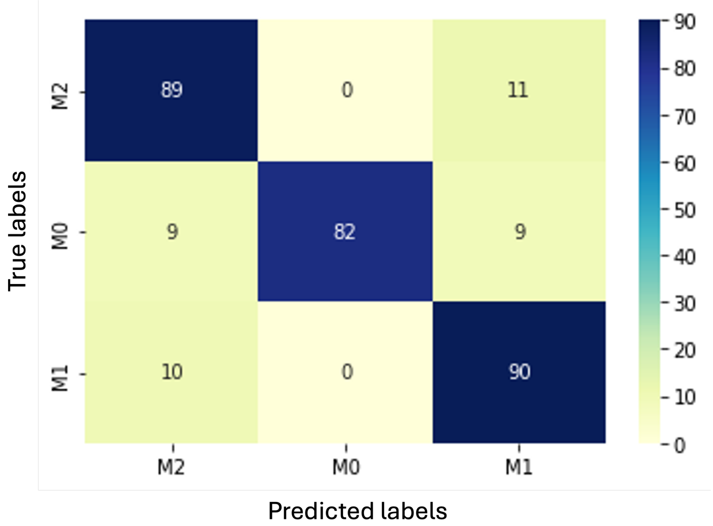


Figure 5.12: Model predictions on the validation data. A heatmap confusion matrix, showing the percentage accuracy of the validation data for the M0, M1 and M2 macrophage subtypes.

When tested on a different batch of mice data, the model showed a 61% misclassification rate for M0 cells as M2, a 74% accuracy for M1 cells, and a distributed accuracy for M2 cells, with 44% classified as M1 and 42% as M2 (Fig. 5.13). Although the study demonstrated higher accuracy for cells from the same batch, the test set revealed inconsistencies in the accuracy for M0 and M2 cells.

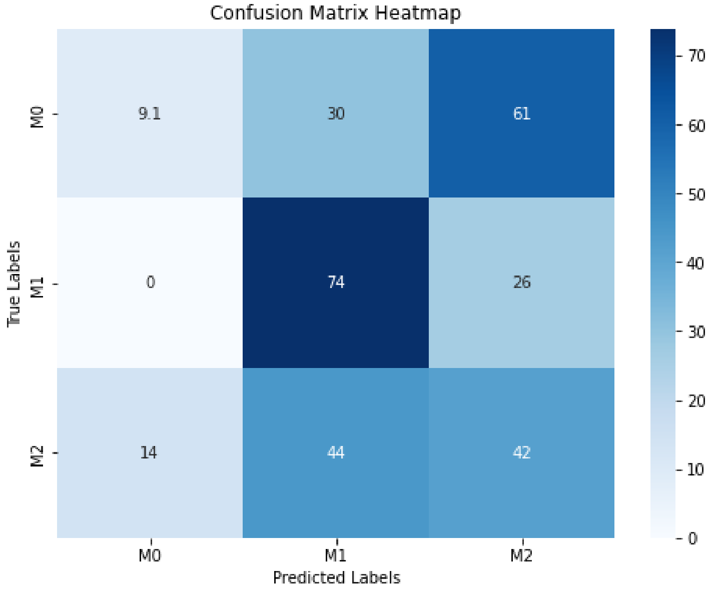


Figure 5.13: Model predictions on the test data. A heatmap confusion matrix, showing the percentage accuracy of the test data for the M0, M1 and M2 macrophage subtypes.

5.2.4 Discussion

Fig.5.8 illustrates that cells in the test set do not consistently follow the same directional patterns as those in the training set (Fig.5.7). Additionally, the similar trajectories of M0 and M2 cells in the test sets may pose a challenging problem for the model in identifying the difference between the two, given the amount of training data. This discrepancy may be attributed to the sticky nature of macrophages, which varies across different batches and experimental conditions.

Integrating morphological parameters could potentially help distinguish between M0 and M2 cells, but this requires high-quality images with low cell density to accurately assess cell morphology under electric fields. Given the plasticity of these cells and their tendency to change shape along the time series, quantifying morphological parameters presents significant challenges. Any definitive conclusions using only the motility parameters and trajectory patterns of the macrophages can only be drawn after training and testing the model with multiple batches of data under electric fields.

6 | Machine Learning based model to Predict Germination Status in Bacterial Spores

This chapter is currently under preparation for publication. The dissertation author is the co-first author of this paper. This work highlights the generalizability of the methods while emphasizing our primary goal of controlling spore germination. To ensure effective real-time control, it is essential to leverage real-time analysis, modeling, and prediction. Significantly, this is the first time a predictive model is being introduced, representing a substantial advancement beyond previous work that was solely focused on classification. All the data used in this chapter is generated by UC San Diego Süel Lab.

Understanding spore germination can lead to a wide range of applications where spore reduction or encouragement of activity can be exploited, including medical, food safety, and agricultural applications. To survive suboptimal conditions, bacterial spores (Bacillus Subtilis) develop a protective coat to endure their environment, and in some cases, they do so for years. Despite lacking metabolic and biological activity, spores can determine germination if given suitable conditions for survival. It has been shown that spores respond to changing environment signals through changes in their electrochemical potential and, thus, don't require cellular energy to determine optimal germination. While electrochemical potential plays a role in understanding its extracellular state, how individual spores determine a germination threshold is unknown. To understand the mechanisms of germination, we use microscopy imaging to extract physiological spore features and produce a model that can determine germination.

6.1 Introduction

Determining optimal germination time is critical to the survival of bacterial spores (Bacillus Subtilis), if they germinate prematurely they risk death due to the loss of their protective coat present during dormancy, and if they germinate too late they fail to maximize their time in optimal conditions. Bacterial spores have failed to show any biological mechanisms involved in determining the state of their environment, but it has been shown that a passive method utilizing the integration of electrochemical potential may allow the spore to determine whether their conditions are appropriate for germination [120]. To understand the criteria for germination, we have built a model to predict germination time of individual spores using measures of physiological features over time extracted from phase contrast and fluorescence imaging techniques. Understanding these criteria can influence industries where bacterial promotion or eradication is desirable. Such industries may include medical, food safety, or agricultural industries.

In the medical industry, bacterial eradication would be desirable for numerous treatments, such as infection control, wound healing, and environmental decontamination, which drastically affect an individual's state of health [121]. Less trivially, bacterial promotion may be beneficial in treatments such as vaccines where the controlled introduction of spores results in an immune response [122]. In addition to direct medical applications, understanding bacterial germination would benefit food safety industries through improved preservation methods and reduce food-borne diseases through bacterial eradication [123]. Agricultural industries would be able to prevent bacterial diseases affecting crops, increasing soil health through the promotion of beneficial bacteria or reduction of harmful bacteria [124], and development of bacterial-based pest control [125]. While many other industries and applications exist where control of spore germination may lead to substantial advancements, even just the few mentioned would greatly improve the quality of life for the general public.

6.2 Materials and methods

6.2.1 Isolation and culture of bacterial spores

Microfluidic Protocol

This protocol involves an initial 1-hour exposure at 0.75 psi RM (resuspension media), followed by ten cycles of the following sequence: 3 minutes at 2.5 psi RM with 10mM L-alanine, 2 minutes at 2.5 psi RM, and 1 hour and 55 minutes at 0.75 psi RM. The process concludes with a final exposure at 0.75 psi RM.

RM contains ThT (electrochemical reporter). The frames where L-alanine is added can vary by one frame between movies. The microscope and the microfluidic system are independent and started manually, resulting in a consistent shift throughout the movie. L-alanine addition occurs at high pressure (2.5 psi instead of 0.75). Following the L-alanine addition, there is a 2-minute high-pressure exposure to clean L-alanine from the chamber. Pressure changes during media switching can cause focus drifts. The distance in pixels is 0.065 µm.

ThT levels increase each time the spore encounters the germinant, with a significant increase in the ThT signal during germination. Phase contrast remains constant until germination occurs (white to black), and spores increase in size due to water uptake during germination.

6.2.2 Time-lapse recording

Spores are exposed to L-alanine at three minute pulses at one hour then every two hours afterwards to induce germination. Phase-contrast (PhC) and fluorescence (ThT) images are taken simultaneously at five minute intervals to observe time of germination and size metrics respectively.

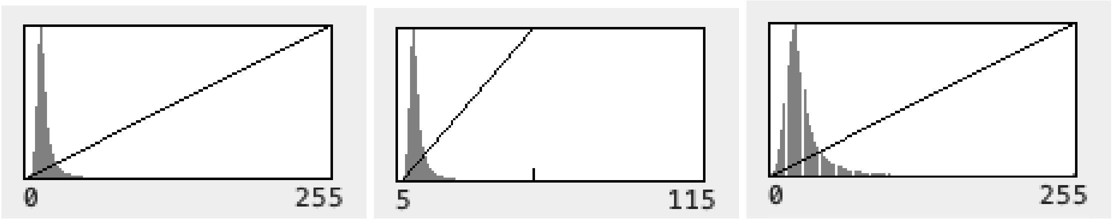
6.2.3 Generation of time-series data capturing bacterial spores germination

We use two sets of time-series images, M5871_s1 and M4576_s2, consisting of 289 and 276 frames, respectively, captured at 5-minute intervals per frame. We first preprocess the images to ensure accurate feature representation in each frame. Next, we use the TrackMate package [71] with Fiji software to track spore features over time for each microscopy technique. Fig. 6.2 shows the preprocessing pipeline and tracking results. Subsequently, we utilize the (x, y) position coordinates within the frame to map the extracted features from each imaging technique and gather comprehensive data for each spore.

Preprocessing

To enhance feature extraction, we crop our images to exclude densely populated spore areas that lead to improper identification. This misidentification occurs due to the high local volume of spores, where intersecting contrast or fluorescence can be mistaken for individual spores because of similar pixel color value, size, and circularity. To address this, we manually inspect the initial image to identify dense areas and then crop these regions out, consistently applying this selected subsection throughout the movie.

We then convert the cropped image from 16-bit to 8-bit, reducing the grayscale color range from 65,536 to 256 possible values for each pixel. We enhance the movie's contrast by adjusting the color range of the existing pixel values using Fiji's auto-rescaling function. This algorithm automatically stretches the pixel range histogram to maximize the range between all possible pixel values within the 8-bit image. Fig. 6.1 shows the adjusted color range histogram of a sample image.



(a) Original histogram (b) Original histogram with adjustment bounds (c) Rescaled histogram

Figure 6.1: **Color range histograms after maximization**
Figure showing the color histogram prior to and subsequent to the application of auto adjustment on ThT microscopy images. The pixel values are plotted along the x-axis, while the y-axis represents the count of pixels for each color value. The resulting histogram maps the image's minimum and maximum ranges to 0 and 255, ensuring optimal contrast without the loss of data.

We use a rolling ball algorithm to correct background intensity variations and isolate spore fluorescence. This algorithm calculates a local background value for each pixel by averaging the values within a specified radius around it. The average value is then subtracted from the corresponding pixel, reducing background intensity. Additionally, we disable smoothing to ensure the algorithm does not further alter the image data.

Segmentation and tracking

To extract time-series of spore features, we utilize the TrackMate package with the threshold detector to process the images. This involves implementing segmenting,

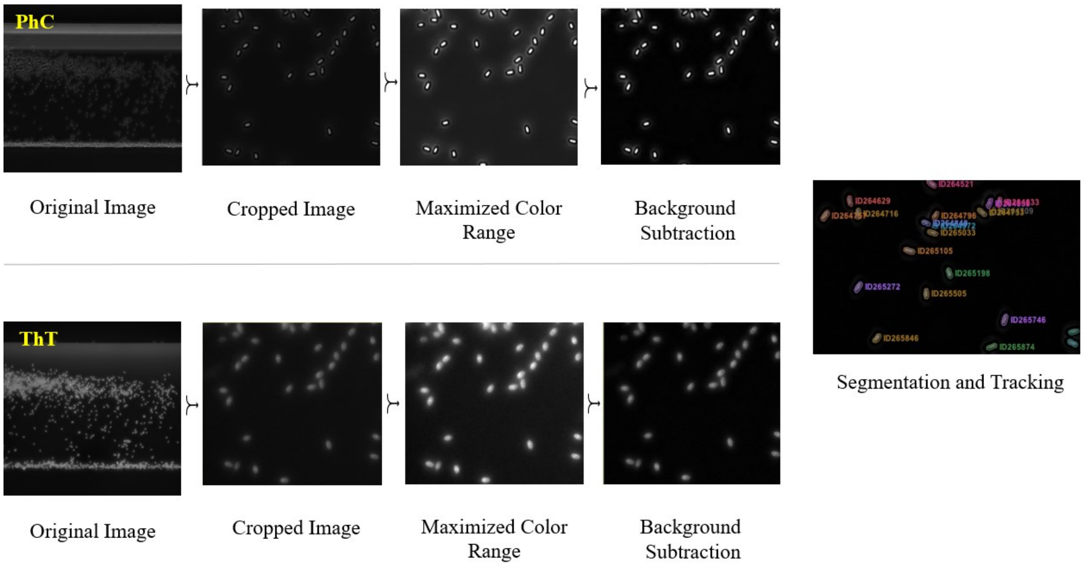


Figure 6.2: **Image processing pipeline for microscopy imaging**
Schematic illustrating the image processing pipeline for PhC and ThT images, involving preprocessing, segmentation, and tracking results. Initially, the original image is cropped to exclude densely populated areas of spores, selecting only the sparsely populated regions suitable for processing. Then, we maximize the color range to achieve optimal contrast by adjusting the minimum and maximum color values. The final preprocessing step involves applying the rolling ball background subtraction method to even out the background intensity variations. Lastly, we display each spore independently identified and tracked in the segmentation and tracking results.

filtering, and particle-linking on the images.

First we segment the spores at each timestep by specifying a grey value threshold to categorize the image into foreground and background pixels. This threshold is chosen through the auto function built into TrackMate, which detects the optimal fluorescence to maximize spot quality. Spot quality measures the object area in pixels in connected regions.

To link spores across frames, we use the overlap tracker provided by TrackMate. This tracker groups connected regions of pixels by identifying overlapping areas in consecutive frames. We chose this tracker because the stationary nature of our spores ensures their regions consistently overlap across frames. We parameterize the overlap tracker by specifying the Intersection over Union (IoU):

$$\text{IoU} = \frac{\text{Area of Intersection}}{\text{Area over Union}}$$

We measure the area of each region using TrackMate's precise calculation method rather than its fast calculation. The precise method calculates the area based on the actual shape of the spore, while the fast method simply uses the bounding box around the spore. Because the spores do not move, either method would technically work for our purposes. We have chosen a minimum Intersection over Union (IoU) threshold of 0.5 to account for potential variability in spore detection across different images, which might arise from changes in image quality. However, since the spores remain stationary, we anticipate that the actual IoU value between spores in subsequent frames will typically be much higher than 0.5. Furthermore, we apply a scale factor of 1 in our analysis, indicating that we do not adjust the size of interconnected regions before calculating their IoU.

6.2.4 Computing quantitative features from the images

We leverage temporal data from each microscopy technique to refine individual spore data through subsequent processing steps. Phase contrast (PhC) images provide germination timing, while fluorescence (ThT) images yield physiological metrics for each frame. By correlating tracks across both techniques, we compile comprehensive datasets for each spore.

PhC postprocessing

We refine our spore selection by analyzing data at each time point and tracking spore behavior throughout the movie. Initially, we filter data at each frame by excluding regions with low circularity and sizes that are either too small or too large for typical spores. We also evaluate track qualities, eliminating tracks that indicate germination before the initial germinant exposure and that lack data for any frame. The time of germination is determined by the duration of each track, which corresponds to a drop below a certain predefined intensity threshold.

ThT postprocessing

Similar to the postprocessing for PhC images, we apply filters based on inappropriate circularity and size characteristics to our ThT data. We refined our ThT tracks by excluding those that lacked data for any frame throughout the experiment. No additional processing is made to extract physiological metrics.

Extracting time-series data from microscopy images

We analyze each spore's (x, y) coordinates in every frame of the movie. Since the spores remain stationary, we compute the average coordinates over each time series to establish a single set of coordinates for each track. To synchronize spores between PhC and ThT imaging, we compare their coordinates and consider them a match if the variance in both x and y values is within a 5-pixel tolerance.

Now, we have obtained comprehensive behavior for each spore over the entirety of the experiment. The temporal data gathered from each spore includes germination status, average intensity (representing electrochemical potential), ellipse major, ellipse minor, ellipse aspect ratio, area (representing size), perimeter, and circularity. Each physiological metric is calculated using the following formulas:

$$\text{Average Intensity} = \frac{\sum \text{Pixel Grey Values}}{\text{Number of Pixels}}$$

$$\text{Area} = \text{Total Number of Pixels}$$

$$\text{Ellipse Lengths} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_2)^2}$$

$$\text{Perimeter} = \sum_{i-1}^{N-1} |x_i - x_{i+1}|$$

$$\text{Circularity} = \frac{4\pi * \text{Size}}{\text{Perimeter}^2}$$

$$\text{Ellipse Aspect Ratio} = \frac{\text{Ellipse Major Length}}{\text{Ellipse Minor Length}}$$

6.2.5 Dataset creation

Through segmentation and tracking, we identified 76 spores from time-series images M5871_s1 and 54 spores from M4576_s2. These sequences consist of 289 and 276 frames, respectively, with each frame captured at 5-minute intervals. We used the set of 76 spores for training and validation, and the set of 54 spores for testing. These sets originated from separate batches of experiments, making them suitable for distinct training and testing phases. The dataset division is detailed in the table below.

The model's training and testing datasets were constructed using the spores' time series features. These features include spore germination status, intensity, size, perimeter, and germinant exposure time.

Given that the Long Short-Term Memory (LSTM) network uses a 3-timestep lookback, each data instance within the dataset consists of the features from the three preceding timesteps. Specifically, for each instance, the model uses the values of spore germination, intensity, size, perimeter, and germinant exposure time from these three previous timesteps. The target variable for each instance is the spore germination status at the subsequent timestep, which the model aims to predict (Fig. 6.3).

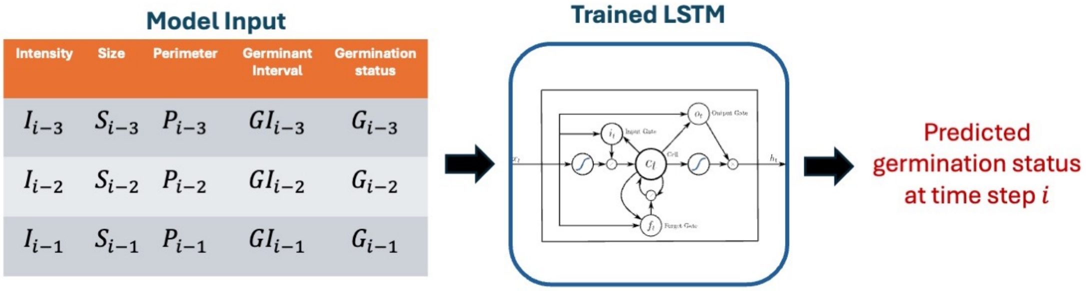


Figure 6.3: The trained LSTM model takes in spore features such as intensity, size, perimeter, and germinant exposure time from the past 3 steps and outputs the spore germination status at the current time step.

Data used for Training, Validation, and Testing				
Experiment Set	Spores	**Training**	**Validation**	**Testing**
M4576_s2	76	66	10	
M4581_s1	54			54

6.2.6 LSTM model architecture

The model used in this study is an LSTM (Long Short-Term Memory) network specifically designed for processing sequential data. The architecture begins with an LSTM layer comprising 80 units, which processes input sequences by maintaining temporal dependencies across the 3-timestep lookback period. This LSTM layer is followed by a Dense layer with a single neuron and a sigmoid activation function, which outputs a probability score for binary classification, indicating the likelihood of spore germination at the next time step. A Dropout layer with a dropout rate of 0.01 to mitigate overfitting is incorporated after the Dense layer. The model is compiled with the binary cross-entropy loss function and the Adam optimizer, which is well-suited for binary classification tasks. The network is trained for a maximum of 50 epochs with a batch size of 1. It utilizes early stopping to prevent overfitting, with the best weights restored upon encountering no improvement in validation loss over a specified patience period. After training, predictions were made on the training, validation, and test datasets using the trained model. We note that different lookback sizes were considered, and a lookback of 3 provided the best performance.

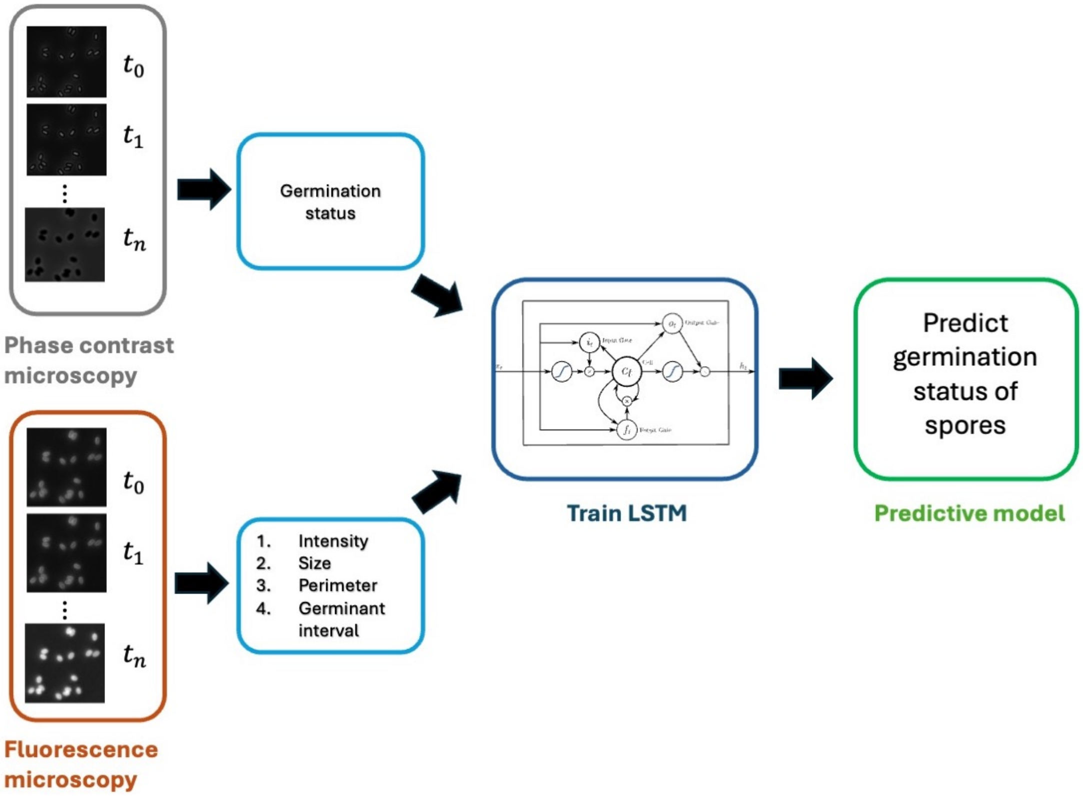


Figure 6.4: The image processing and spore germination pipeline. Phase contrast microscopy images are used to track the prediction status of the spores, and Fluorescence microscopy images are used to track features like intensity, size, perimeter, and the germination interval of the spores. The tracking data is used to create a time series of the five features, which are used as inputs to our LSTM model to make predictions about the next germination status of the spores.

6.3 Results

6.3.1 Germination events over time

We track spore germination events of 130 spores using phase-contrast imaging on experiments M4576_s2 and M4581_s1 6.1. Our approach involves visualization and analysis of germination events coupled with population dormancy at each time point and during each germinant exposure, as shown in Fig. 6.5. During sequential germinant exposures, significant portions of the population germinate, with the second germinant pulse having the most substantial effect. Specifically, 8.46%, 55.38%, 30.00%, and 5.38% of the population germinate immediately following these exposures. After the second germinant pulse, only 35.15% of the population remains dormant, indicating that a large portion of spores reach their germination threshold after two exposures. By the fourth pulse, only 0.77% of spores remain dormant.

Germination					
germinant exposures	1	2	3	4	8
germination events	11	72	39	7	1
germination after exposure (%)	8.46	55.38	30.00	5.38	0.77
population dormancy (%)	91.54	36.15	6.15	0.77	0.00
total germination events	130				

Table 6.1: Germination metrics of the experiments sets M4576_s2 and M4581_s1

6.3.2 Physiological trends

Throughout the experiment, we measured the electrochemical potential and size of spores, as well as the time it took for them to germinate. Our findings show that spores germinated after exposure to one of the first four germinant pulses, as illustrated in Fig. 1. As spores approached germination, their electrochemical potential and size slightly increased, with a significant increase after each germinant exposure. The most considerable electrochemical potential and size increase occurred at the germinant pulse, which led to germination. We display the four main trends in Fig. 6.6, using randomly selected spores representative of each trend.

For these four trends, we calculated the average feature values for electrochemical

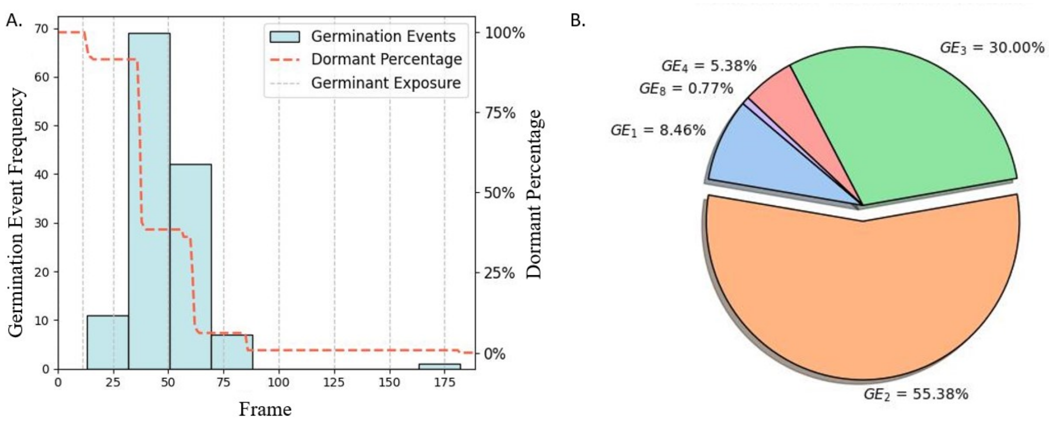


Figure 6.5: **Dormancy and germination metrics**
A) Population dormancy and germination events over time
Here we show germination events by frame, as well as population dormancy percentage for combined experiments M458_s1 and M4576_s2 with a total of 130 spores. Germinant pulses are shown in grey. The optimal bin width is determined using the Sturges method [1]. This is calculated using the following formula: bin width = $1 + 3.322\log(\text{Number of Spores})$. As a result, we organize germination events into nine bins. **B) Germination events induced by germinant exposures**
GE_I denotes population percentage germinated after exposure I.

potential and size based on the germinant exposure group and the change in feature values between exposures. These metrics are presented in Table 6.2. We observed that each group experienced a minimal change in electrochemical potential, ranging from 0.02 to 0.05 before germination. However, at each group's germinant exposure inducing germination, there was a notable difference in average electrochemical potential, with a minimum change of 0.15, three times greater than the highest change before germination. Similarly, for size, as shown in Table 6.3, we observed a change in size ranging from 0.03 to 0.05 between germinant exposures before germination. At the time of germination, there was a much more significant change in size, ranging from 0.15 to 0.19 across each group - at least three times as large as the changes between exposures preceding germination.

We quantify the relationship between all size metrics by considering the correla-

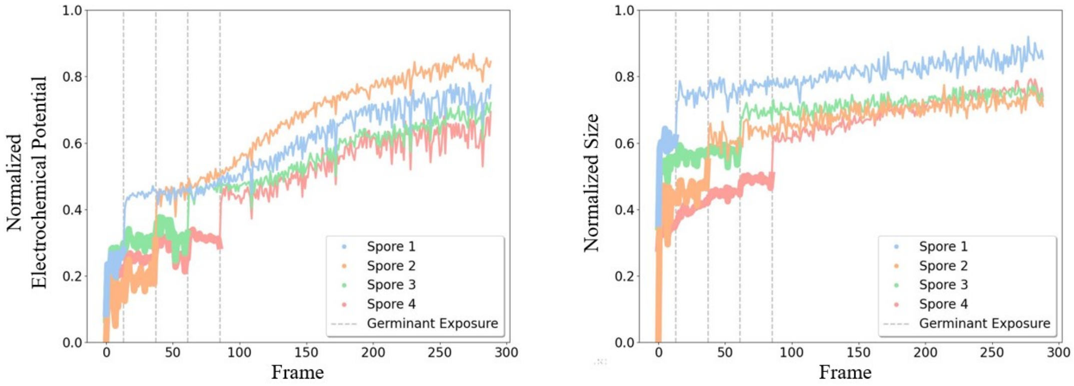


Figure 6.6: **Major trends in electrochemical potential and size**
We show the four major trends over time shown in the combined experiments. Spore 1 depicts a spore germinating after the first exposure, and similarly, Spores 2, 3, and 4 germinating after the second, third, and fourth exposure. The bold line depicts the feature trend before germination, and the thin line shows the trend while germinating. The vertical grey lines show when the collective spores were exposed to the germinant. We see that spores increased in electrochemical potential and size with each exposure, but the most significant change occurs at germination.

Average electrochemical potential by exposure group					
Germinant exposures	0	1	2	3	4
Group 1	0.29	0.48	0.48	0.52	0.57
change between exposures	-	**0.19**	0	0.04	0.04
Group 2	0.26	0.31	0.46	0.49	0.59
change between exposures	-	0.05	**0.15**	0.02	0.02
Group 3	0.25	0.3	0.33	0.49	0.5
change between exposures	-	0.05	0.03	**0.16**	0.01
Group 4	0.21	0.25	0.27	0.32	0.5
change between exposures	-	0.05	0.02	0.05	**0.18**

Table 6.2: **Average electrochemical potential by exposure group**
We grouped the spores based on the number of germinant exposures required for germination. Group 1 germinates after one exposure, Group 2 after two exposures, Group 3 after three, and Group 4 after four exposures. We present the average electrochemical potential for each group at various germinant pulses and highlight the changes in these values. The most significant change, shown in bold, occurs at the germinant exposure that triggers their germination.

tion of physiological features over time as shown in Fig. 6.7.

Average size by exposure group					
Germinant exposures	0	1	2	3	4
Group 1	0.45	0.58	0.61	0.63	0.66
change between exposures	-	**0.13**	0.03	0.02	0.03
Group 2	0.41	0.44	0.59	0.63	0.65
change between exposures	-	0.02	**0.15**	0.01	0.04
Group 3	0.43	0.47	0.48	0.63	0.66
change between exposures	-	0.03	0.01	**0.15**	0.03
Group 4	0.4	0.43	0.45	0.47	0.62
change between exposures	-	0.04	0.01	0.02	**0.15**

Table 6.3: **Average size by exposure group**
Similarly to Table 6.2, we group spores that germinate with different quantities of germinant. We compute the average size for each group at sequential germinant pulses, as well as the change in these values between pulses. We highlight the biggest change in bold, which occurs at the germinant exposure that induces germination.

6.3.3 Feature correlation

We conduct correlation analysis by considering the strength of the linear relationship between pairs of physiological features throughout the experiment. We utilize the Pearson correlation coefficient ($r_{\mathrm{F}_1,\mathrm{F}_2}$) [126] to quantify this relationship:

$$r_{\mathrm{F}_1,\mathrm{F}_2} = \frac{cov(\mathrm{F}_1, \mathrm{F}_2)}{\sigma_{\mathrm{F}_1}\sigma_{\mathrm{F}_2}}$$

where F_1 and F_2 are the features considered, $cov(\mathrm{F}_1, \mathrm{F}_2)$ measures the covariance between these features over time, and $\sigma_{\mathrm{F}_1}\sigma_{\mathrm{F}_2}$ measures the standard deviation of each feature. We consider $0.65 < |r_{\mathrm{F}_1,\mathrm{F}_2}| < 1$ to signify a strong relationship, and intermediate values to show a more moderate relationship.

This correlation analysis was conducted between germination status, intensity, size, ellipse minor, ellipse major, perimeter, circularity, and ellipse aspect ratio of the data extracted from our combined experiments. We visualize our results using a heatmap depiction of the correlation values as shown in Fig. 6.7. Trivially high correlations exist between size metrics (intensity, size, ellipse minor, ellipse major, perimeter, circularity, and ellipse aspect ratio) ranging from ±0.66 to ±0.97 correlation. We also observe the following size metrics ordered from highest to lowest correlation have a significant relationship with germination: ellipse aspect ratio (-0.86), ellipse minor (0.84), size (0.79), intensity (0.76), perimeter (0.75), and ellipse major (0.66). Through this analysis, we also see that intensity has a significant correlation with the same size metrics, ellipse minor (0.84), size (0.83), ellipse aspect ratio (-0.78), perimeter (0.77), and ellipse major (0.76).

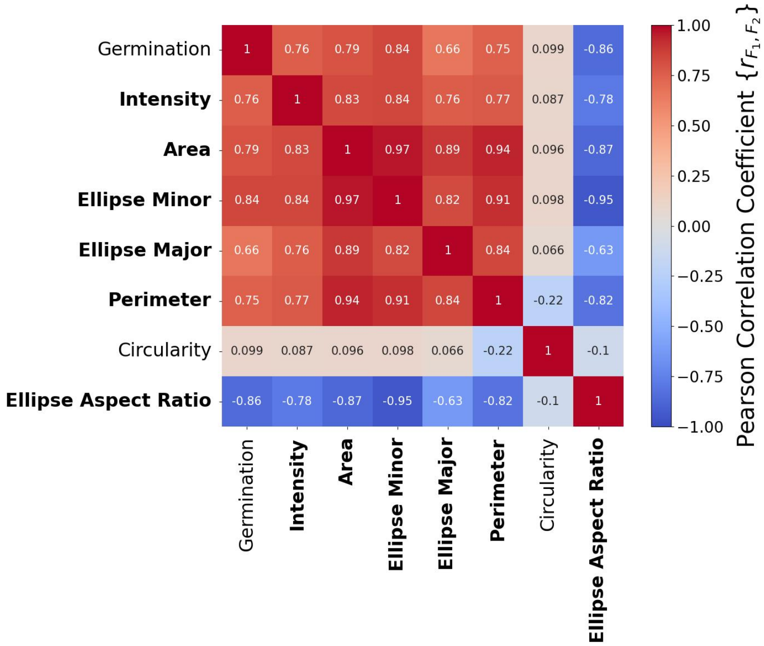


Figure 6.7: **Feature correlations between physiological features and germination**

For each combination of features, we measured the Pearson correlation coefficient to measure their relationship over time. In bold we have features that were considered highly correlated with germination.

6.3.4 LSTM can predict the germination status of the spores at single spore level

To predict germination, we consider our model with lookback N, meaning at each timestep the model considers the values of our parameters at the previous N steps, to determine whether at the next timestep a spore is germinated. To evaluate the predictive performance of the LSTM model at the spore level, we calculated the accuracy, recall, and F1 score for each spore at each timestep. The model's predictions, which were initially probabilities, were thresholded at 0.5 to obtain binary outcomes indicating germination status (1 for germinated, 0 for not germinated). We then compared these binary predictions with the actual germination statuses to determine the proportion of correct predictions, which constitutes the accuracy. The accuracy, recall, and F1 scores on the test set are 0.92, 0.91, and 0.95, respectively. These metrics provided a comprehensive evaluation of the model's performance on the test dataset.

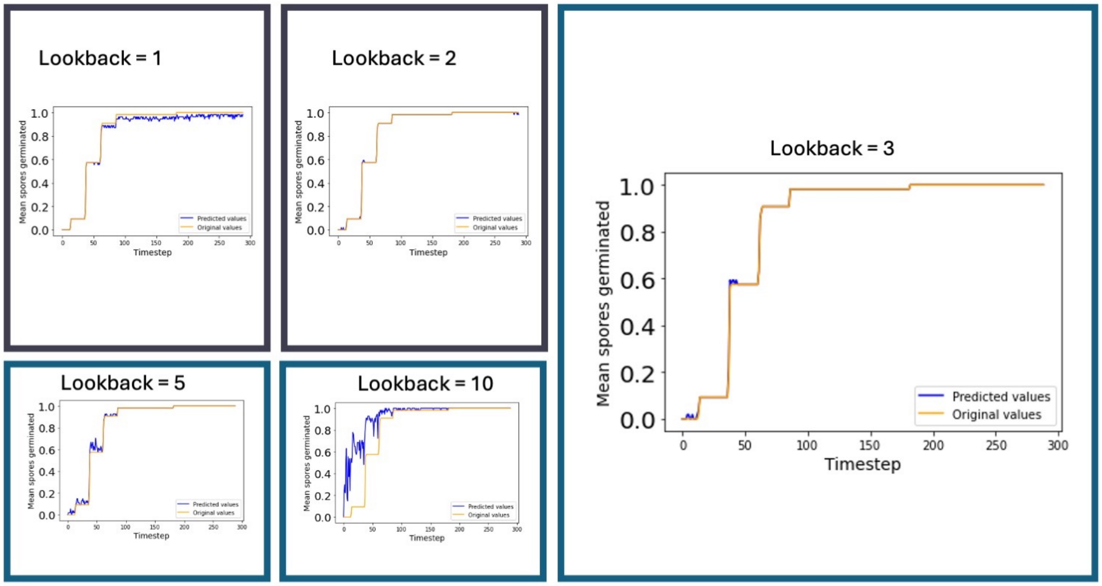


Figure 6.8: Population level predictions on spore germination

6.4 Conclusion

We extracted and analyzed spore germination times and associated physiological features. For future feature analysis, we plan to investigate varying germinant exposure intervals and concentrations, as well as explore the spatial dynamics between modified spores capable of intercellular communication. We developed an LSTM model that accurately predicts germination time using features such as intensity, area, and perimeter. Further analysis will focus on refining this model by examining highly correlated features and exploring alternative modeling approaches to understand better the biological mechanisms involved in germination.

7 | Conclusion

Our research aimed to investigate various cell types using microscopy time series images, with the primary goal of monitoring and understanding the dynamics of biological systems.

7.1 Summary

Chapter 2 explores various types of microscopy imaging and the techniques used to extract valuable information from these images. While numerous image processing, machine learning, and deep learning methods exist for analyzing microscopy images, they are often tailored to specific problems and cannot be universally applied. Our approach utilized image processing techniques to read, convert, and pre-process the images. Following this, we identified objects such as cells within the images. For time-series images, we tracked these objects over time to analyze their movement characteristics.

Chapter 3 presents a non-invasive method using image processing techniques to study cell behavior and migratory patterns, with a focus on predicting macrophage subtypes. Our aim was to extract the cells' shape and motility patterns and then map these morphological patterns to their motility behaviors. We developed a machine-learning model to classify different cell types. Unlike flow cytometry, which requires cell fixation, this technique preserves cell viability for further analysis and allows real-time observations. It is particularly useful for assessing macrophage polarization through migration and morphological changes, providing insights into cellular responses to stimuli without the need for external cues like electrical guidance or staining. This approach is valuable for high-throughput screening in biochemical research and drug discovery.

In Chapter 4, we investigated whether the (x,y) position values of cells could be used to classify different cell subtypes and created a deep-learning model for this purpose. Our focus was on macrophages, which are important for wound healing. We discovered distinct migratory patterns and morphologies among the M0, M1, and M2 subtypes. We found clear differences in migratory patterns by analyzing trajectories, perimeter, convex hull area, and pairwise distances. This method has the potential for future applications where obtaining precise cell morphology data is difficult, and it could enhance the reliability of cell identification through phase-contrast microscopy, even with lower-quality images.

In Chapter 5, we created an image analyzer within a closed-loop system to control cell migration. This analyzer measures cell speed, directionality, and the percentage of cell movement in response to different electric fields. The controller uses this data to monitor and adjust the electric current in the system, applying an electric field to macrophages in the chamber. A microscope captures images of the cells every 5 minutes, which are then processed by the image analyzer, creating a real-time closed-loop system.

In the second section of this chapter, we conducted a study to classify macrophage subtypes based on their response to electric fields. Our goal was to determine if different subtypes could be distinguished by their behavior in an electric field. We observed that the M0 and M2 subtypes moved in the opposite direction compared to the M1 subtypes. In addition to this directional difference, we examined other factors, such as trajectory patterns, to classify the macrophage subtypes. To achieve this, we developed an LSTM model that used the (x,y) trajectory values of each subtype. Our results showed that the model could accurately identify cell subtypes within the same batch of data. However, it had difficulty distinguishing between M0 and M2 cell types when applied to a different batch. By incorporating data from multiple batches of experiments, we aim to improve the model and gain a deeper understanding of macrophage characteristics.

Finally, in Chapter 6, we applied these methods to time-series images of bacterial spores to predict their germination time under interval-based germinant exposure. This demonstrates the versatility of our developed techniques, as they can be effectively used across different cell types and for both phase contrast and fluorescence microscopy images. This work highlights the generalizability of the methods while emphasizing our primary goal of controlling spore germination. To ensure effective real-time control, it is essential to leverage real-time analysis, modeling, and prediction. Significantly, this is the first time a predictive model is being introduced, representing a substantial advancement beyond previous work that was solely focused on classification.

7.2 Future Work

For future work, we propose several valuable and exciting extensions of our research, which are listed below.

- **Classification of macrophage subtypes in galvanotaxis**

  Training the deep learning model with additional data is essential for improving its accuracy. Analyzing the deep learning results suggests that morphology and trajectory patterns are interconnected. Incorporating more neural network layers and fine-tuning hyperparameters could further enhance prediction accuracy. The model's ability to classify migratory patterns would improve significantly with thousands of cell trajectories. To achieve this, we must collect extensive time-series images of cell migration and perform the necessary pre-processing.

This approach will aid in more accurately classifying cells in response to external stimuli.

- **Building a neuronal segmentation algorithm**

  We are currently developing a neuronal segmentation algorithm using U-Net, a type of convolutional neural network. This algorithm can be used to segment neurons and identify various characteristics, such as length, area, and other relevant features.

- **Extending the spore germination model results in identifying the underlying germination mechanism.**

  We are also working on developing a physics-based model to explore the dynamics of spore germination. By integrating results from our existing spore germination model, we aim to gain deeper insights into how bacterial spores make decisions to germinate. This model will help us investigate whether there are any collective decision-making processes or cell-to-cell communications influencing germination.

A | Appendix

A.1 Code and Data availability

- The complete dataset used in Chapter 3 is available at `https://datadryad.org/stash/share/vrQi68_C27sbdkP2uAXUE8kvy_BxlYkRImI5f6_9WY0` with a metadata file accompanying the dataset. The source code required to reproduce the findings of this study has been uploaded to the following github repository: https://github.com/Gomez-Lab/CellAnalysis.
- The code and the datasets to replicate the results in Chapter 4 can be found on github: https://github.com/Gomez-Lab/Macrophages_TrajectoryPatternAnalysis.

A.2 Additional figures for Chapter 3

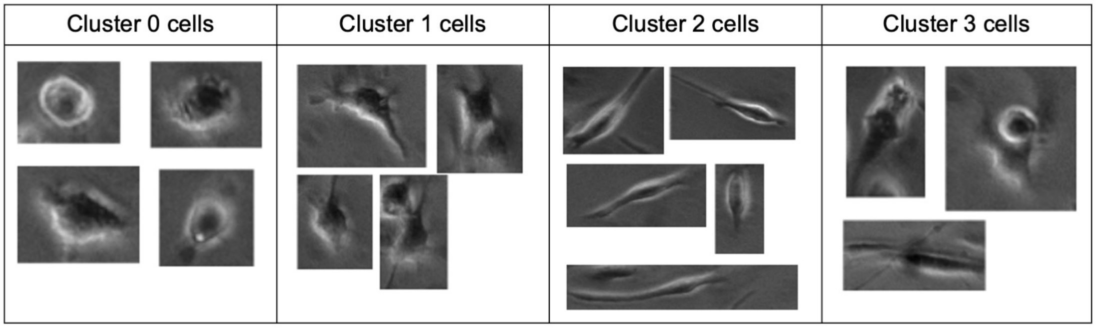


Figure A.1: **Examples of cell images from each cluster.**

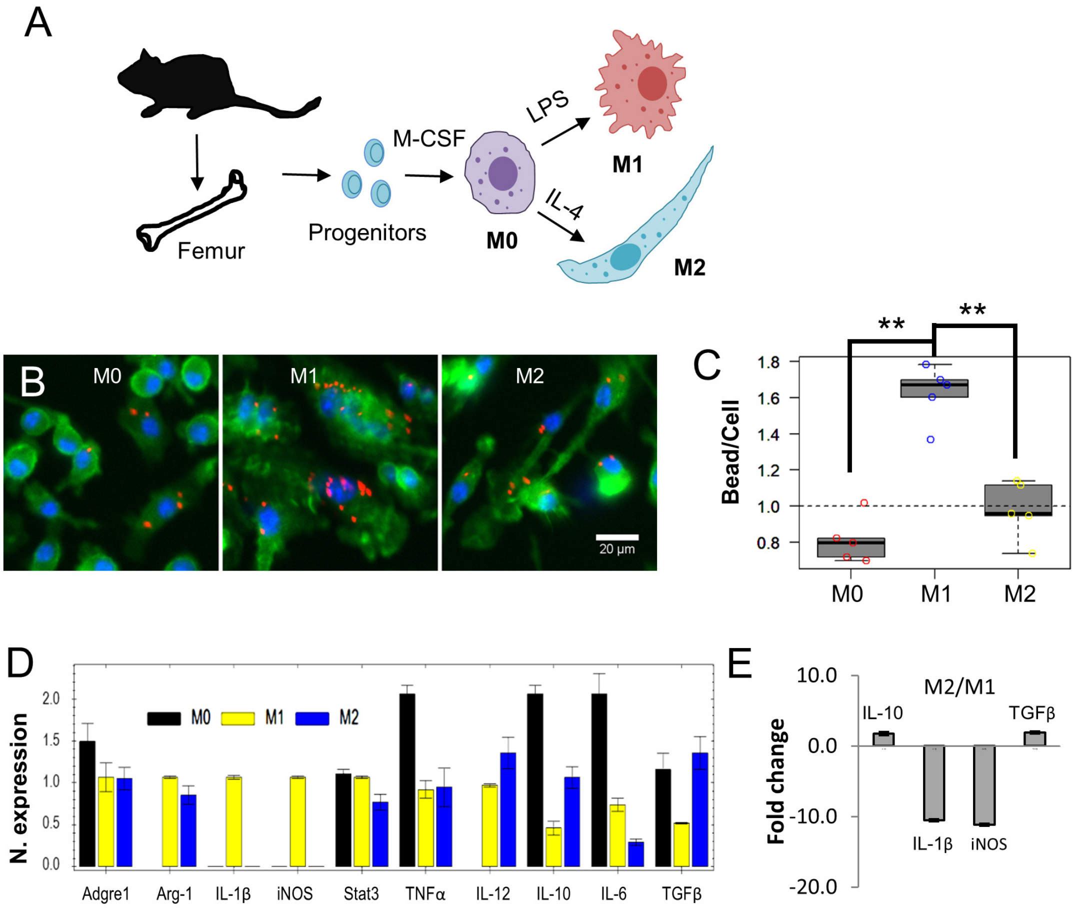


Figure A.2: **Generation and characterization of murine bone marrow derived M0, M1 and M2 macrophages.** (A) Scheme for the generation of mouse bone marrow derived macrophages (BMDMs, M0) and subsequent polarization of M1 (100 ng/ml LPS) and M2 (20 ng/ml mIL-4) macrophages. (B) Fluorescent images show M0, M1 and M2 macrophages containing latex beads (red). Actins were stained with FITC-phalloidin (green). Nuclei were counterstained with Hoechst (blue). Note the significantly increased phagocytosis capacity of M1 macrophages. Bar, 20 μm. (C) Quantification of phagocytosis. Data was calculated as bead per cell from 5 randomly chosen fields. ** $p < 0.01$ by one-way ANOVA with post-hoc Tukey HSD Test. (D) The target gene expression profiles of naïve and differentially polarized macrophages. Fold expression is calculated relative to the internal control of GAPDH mRNA expression. (E) Differential M1 (IL-1β and iNOS) or M2 marker (IL-10 and TGFβ) gene expressions induced by LPS or mIL-4, respectively. Fold change is calculated as M2/M1 ratio of mRNA expression.

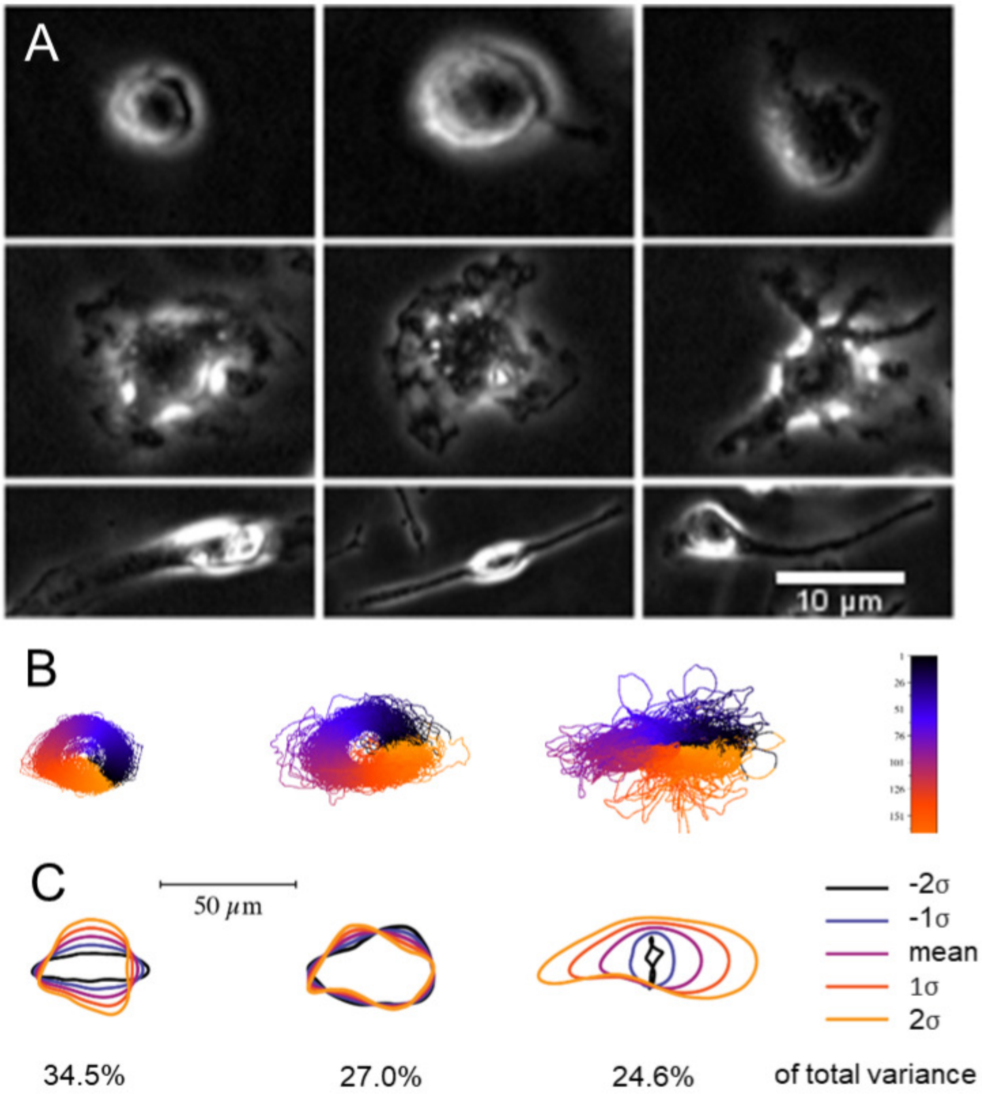


Figure A.3: **Macrophage shapes are described by three principal shape modes.** (A) Phase-contrast images of differentially polarized macrophages illustrate the phenotypic shape variation in the combined population. Bar, 10 μm. (B) The alignments of 2329 live macrophages by their outlines that are equally divided into 200 points. (C) Top three principal modes of macrophage shape variation as determined by principal components analysis. These modes—circular (mode 1), "with protrusions" shape (mode 2, one example is shown), and elongated (mode 3)—are highly reproducible; subsequent modes seem to be mixtures or noise. For each mode, the mean cell shape is shown alongside reconstructions of shapes one and two standard deviations away from the mean in each direction along the given mode. The variation accounted for by each mode is indicated. Bar, 50 μm.

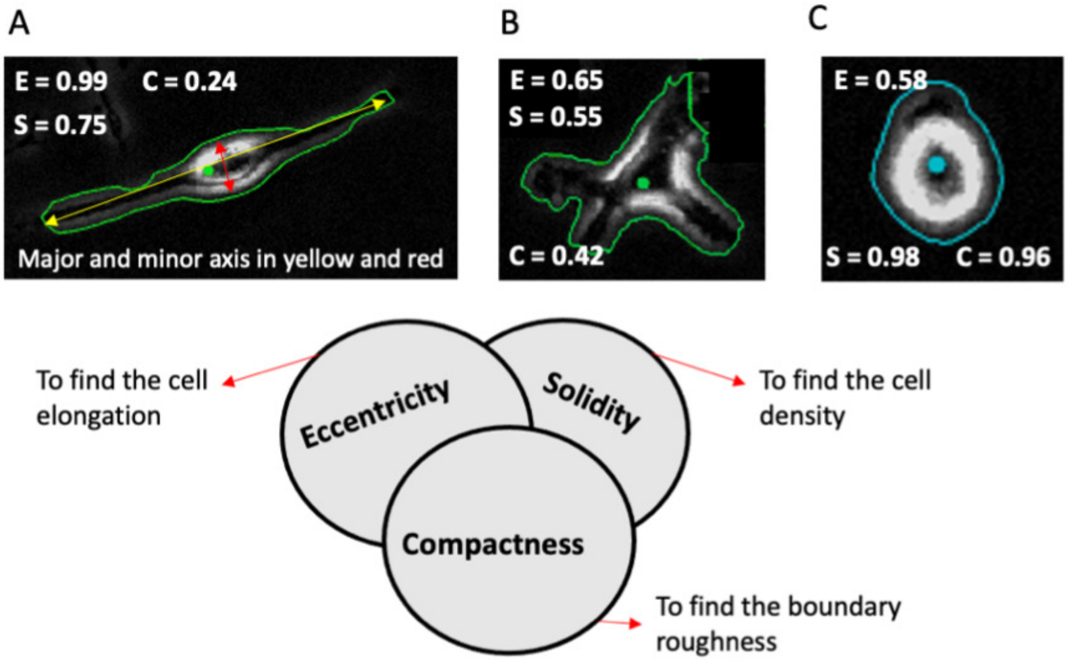


Figure A.4: **Examples of elongated/ bipolar, multipolar, and circular macrophages with their shape parameter values.** A) Elongated/Bipolar cell with compactness = 0.24, eccentricity = 0.99, solidity = 0.75 - highest eccentricity, with low compactness and higher solidity, (B) Multipolar cell with compactness = 0.42, eccentricity = 0.65, solidity = 0.55 - average eccentricity, with low compactness and solidity, (C) Circular cell with compactness = 0.96, eccentricity = 0.58, solidity = 0.98 - average eccentricity, with highest compactness and solidity.

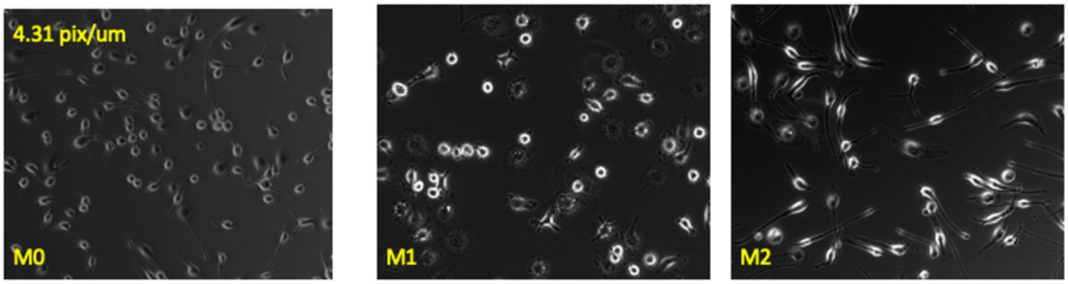


Figure A.5: **Representative M0, M1 and M2 macrophage images used in the analysis.**

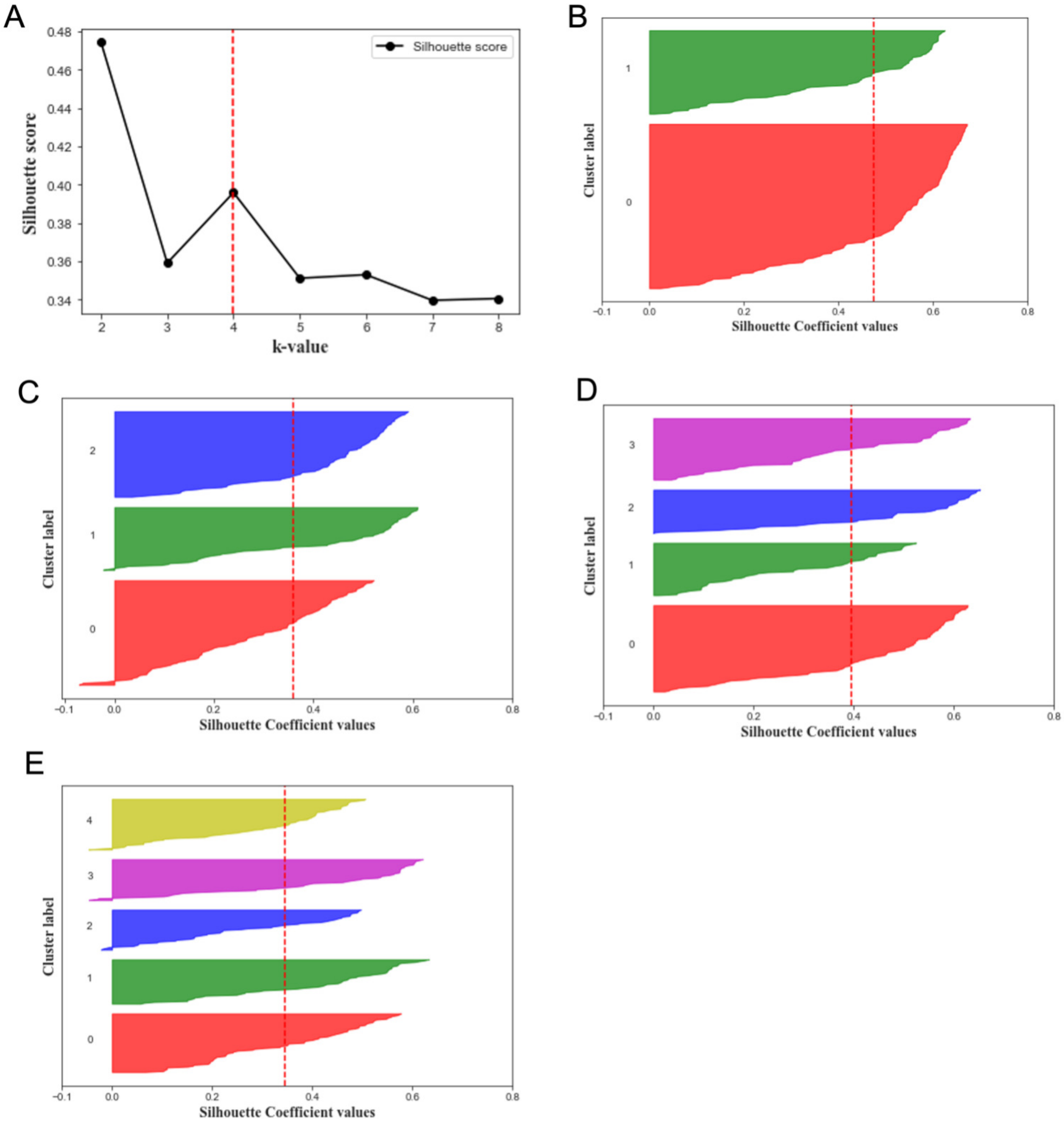


Figure A.6: **Silhouette coefficient score plot and silhouette plots to determine the optimal k-value.** A) Silhouette score plot with highest values for k = 2 and k = 4, (B) Silhouette plot for k = 2 with irregular cluster thickness showing one cluster is bigger in size than the other - suboptimal k-value, (C) Silhouette plot for k = 3 with negative values in two of the clusters indicating cells are assigned to wrong clusters and below average score- suboptimal k-value, (D) Silhouette plot for k = 4 with Silhouette score higher than k = 3 and k =5, without any negative values and somewhat uniform cluster thickness - optimal k-value, (E) Silhouette plot for k = 5 with negative values in three of the clusters indicating cells are assigned to wrong clusters and below average score - suboptimal k-value.

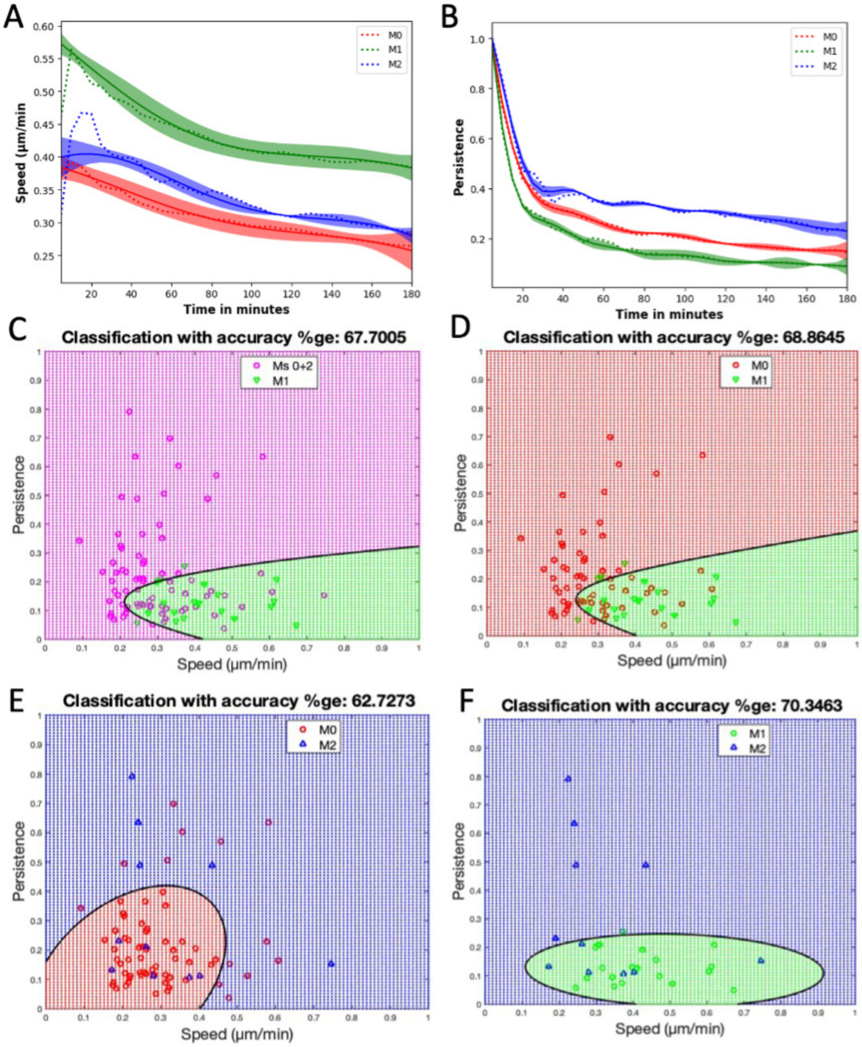


Figure A.7: **Classification and discrimination of the macrophage subtypes using motility parameters alone.** (A) Speed, (B) Persistence, where dotted lines are the mean values of the corresponding macrophage images, solid lines are the Gaussian Process Regressor model (GPR) predictions, and the color bands are their 95% confidence intervals. (C) Plot showing the quadratic classification of the cells using Speed vs Persistence between M1 image and combined (M0+M2) images with a classification accuracy of around 68%. (D) Plot showing the quadratic classification of the cells using Speed vs Persistence between M0 image and M1 image with a classification accuracy of around 69%. (E) Plot showing the quadratic classification of the cells using Speed vs Persistence between M0 image and M2 image with a classification accuracy of around 63%. (F) Plot showing quadratic the classification of the cells using Speed vs Persistence between M1 image and M2 image with a classification accuracy of around 70%. The black line in C-F is the decision boundary between the classes.

B | Additional Tables

Table B.1: Macrophage subtype classification with accuracy percentages for different analysis techniques used in the study.

Input parameters	**Method**	**M0**	**M1**	**M2**
Morphology	K-means Clustering	38%	52%	46%
Morphology	SVM	69%	71%	71%
Motility	SVM	75%	75%	41%
Morphology + Motility	SVM	81%	79%	79%

Bibliography

[1] D. W. Scott, “Sturges’ rule,” *WIREs Computational Statistics*, vol. 1, no. 3, pp. 303–306, 2009.

[2] F. Y. McWhorter, T. Wang, P. Nguyen, T. Chung, and W. F. Liu, “Modulation of macrophage phenotype by cell shape,” *Proceedings of the National Academy of Sciences*, vol. 110, no. 43, pp. 17253–17258, 2013.

[3] E. Meijering, “Cell segmentation: 50 years down the road [life sciences],” *IEEE Signal Processing Magazine*, vol. 29, no. 5, pp. 140–145, 2012.

[4] F. Zernike, “Phase contrast, a new method for the microscopic observation of transparent objects part ii,” *Physica*, vol. 9, no. 10, pp. 974–986, 1942.

[5] I. Ersoy, F. Bunyak, M. Mackey, and K. Palaniappan, “Cell segmentation using hessian-based detection and contour evolution with directional derivatives,” *Proceedings. International Conference on Image Processing*, vol. 2008, p. 1804—1807, 2008.

[6] H.-F. Tsai, J. Gajda, T. F. Sloan, A. Rares, and A. Q. Shen, “Usiigaci: Instance-aware cell tracking in stain-free phase contrast microscopy enabled by machine learning,” *SoftwareX*, vol. 9, pp. 230–237, 2019.

[7] D. A. V. V. et al., “Deep learning automates the quantitative analysis of individual cells in live-cell imaging experiments,” *PLOS Computational Biology*, vol. 12, no. 11, pp. 1–24, 2016.

[8] C. R. H. et al., “Array programming with NumPy,” *Nature*, vol. 585, pp. 357–362, Sept. 2020.

[9] G. Bradski, “The OpenCV Library,” *Dr. Dobb’s Journal of Software Tools*, 2000.

[10] S. Van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, and T. Yu, “scikit-image: image processing in python,” *PeerJ*, vol. 2, p. e453, 2014.

[11] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nature Methods*, vol. 17, pp. 261–272, 2020.

[12] D. B. Allan, T. Caswell, N. C. Keim, C. M. van der Wel, and R. W. Verweij, "soft-matter/trackpy: Trackpy v0.5.0," Apr. 2021.

[13] M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis and Machine Vision.* Springer, Boston, MA, 1993.

[14] Sternberg, "Biomedical image processing," *Computer*, vol. 16, no. 1, pp. 22–34, 1983.

[15] A. K. Jain, "Fundamentals of digital image processing," *Control of Color Imaging Systems*, 2018.

[16] S. Pizer, E. Amburn, J. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. Haar Romenij, ter, J. Zimmerman, and K. Zuiderveld, "Adaptive histogram equalization and its variations," *Computer Vision, Graphics, and Image Processing*, vol. 39, no. 3, pp. 355–368, 1987.

[17] P. S. Jean Serra, *Mathematical Morphology and Its Applications to Image Processing.* Springer Netherlands, 1994.

[18] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[19] D. Liu and J. Yu, "Otsu method and k-means," in *2009 Ninth International Conference on Hybrid Intelligent Systems*, vol. 1, pp. 344–349, 2009.

[20] G. Borgefors, "Distance transformations in digital images," *Computer Vision, Graphics, and Image Processing*, vol. 34, no. 3, pp. 344–371, 1986.

[21] N. Deo, *Graph theory with applications to engineering and computer science.* Prentice-Hall series in automatic computation, Englewood Cliffs, N.J: Prentice-Hall, 1974.

[22] P. J. Soille and M. M. Ansoult, "Automated basin delineation from digital elevation models using mathematical morphology," *Signal Processing*, vol. 20, no. 2, pp. 171–182, 1990.

[23] K. E. G. Magnusson, J. Jaldén, P. M. Gilbert, and H. M. Blau, "Global linking of cell tracks using the viterbi algorithm," *IEEE Transactions on Medical Imaging*, vol. 34, no. 4, pp. 911–929, 2015.

[24] G. Forney, "The viterbi algorithm," *Proceedings of the IEEE*, vol. 61, no. 3, pp. 268–278, 1973.

[25] M. Kesapragada, Y.-H. Sun, C. Recendez, D. Fregoso, M. Rolandi, H.-Y. Yang, R. Isseroff, M. Zhao, and M. Gomez, "A data-driven approach to establishing cell motility patterns as predictors of macrophage subtypes and their relation to cell morphology," *bioRxiv*, 2024.

[26] T. A. Wynn and K. M. Vannella, "Macrophages in tissue repair, regeneration, and fibrosis," *Immunity*, vol. 44, pp. 450–462, 2016.

[27] P. Krzyszczyk, R. Schloss, A. Palmer, and F. Berthiaume, "The role of macrophages in acute and chronic wound healing and interventions to promote pro-wound healing phenotypes," *Frontiers in Physiology*, vol. 9, p. 419, 2018.

[28] F. Martinez and S. Gordon, "The m1 and m2 paradigm of macrophage activation: time for reassessment," *F1000Prime Reports*, vol. 6, p. 13, 2014.

[29] C. Atri, F. Z. Guerfali, and D. Laouini, "Role of human macrophage polarization in inflammation during infectious diseases," *International Journal of Molecular Sciences*, vol. 19, 2018.

[30] D. M. Mosser and J. P. Edwards, "Exploring the full spectrum of macrophage activation," *Nature Reviews Immunology*, vol. 8, pp. 958–969, 2008.

[31] S. Gordon, "Alternative activation of macrophages," *Nature Reviews Immunology*, vol. 3, pp. 23–35, 2003.

[32] P. J. Murray and T. A. Wynn, "Protective and pathogenic functions of macrophage subsets," *Nature Reviews Immunology*, vol. 11, pp. 723–737, 2011.

[33] R. Zhao, H. Liang, E. Clarke, C. Jackson, and M. Xue, "Inflammation in chronic wounds," *International Journal of Molecular Sciences*, vol. 17, 2016.

[34] M. Hesketh, K. B. Sahin, Z. E. West, and R. Z. Murray, "Macrophage phenotypes regulate scar formation and chronic wound healing," *International Journal of Molecular Sciences*, vol. 18, 2017.

[35] S. Gordon and F. O. Martinez, "Alternative activation of macrophages: mechanism and functions," *Immunity*, vol. 32, pp. 593–604, 2010.

[36] A. Sindrilaru and K. Scharffetter-Kochanek, “Disclosure of the culprits: Macrophages-versatile regulators of wound healing,” *Advances in Wound Care*, vol. 2, pp. 357–368, 2013.

[37] K. M. Vannella and T. A. Wynn, “Mechanisms of organ injury and repair by macrophages,” *Annual Review of Physiology*, vol. 79, pp. 593–617, 2017.

[38] F. De Chaumont, S. Dallongeville, N. Chenouard, N. Hervé, S. Pop, T. Provoost, V. Meas-Yedid, P. Pankajakshan, T. Lecomte, Y. Le Montagner, *et al.*, “Icy: an open bioimage informatics platform for extended reproducible research,” *Nature methods*, vol. 9, no. 7, pp. 690–696, 2012.

[39] A. E. Carpenter, T. R. Jones, M. R. Lamprecht, C. Clarke, I. H. Kang, O. Friman, D. A. Guertin, J. H. Chang, R. A. Lindquist, J. Moffat, *et al.*, “Cellprofiler: image analysis software for identifying and quantifying cell phenotypes,” *Genome biology*, vol. 7, pp. 1–11, 2006.

[40] J. Chalfoun, M. Majurski, A. Dima, M. Halter, K. Bhadriraju, and M. Brady, “Lineage mapper: A versatile cell and particle tracker,” *Scientific reports*, vol. 6, no. 1, p. 36984, 2016.

[41] N. Meng, E. Y. Lam, K. K. Tsia, and H. K. H. So, “Large-scale multi-class image-based cell classification with deep learning,” *IEEE Journal of Biomedical and Health Informatics*, vol. 23, pp. 2091–2098, 2019.

[42] Y. Gu, A. C. Zhang, Y. Han, J. Li, C. Chen, and Y.-H. Lo, “Machine learning based real-time image-guided cell sorting and classification,” *Cytometry Part A*, vol. 95, no. 5, pp. 499–509, 2019.

[43] M. Shifat-E-Rabbi, X. W. Yin, C. E. Fitzgerald, and G. K. Rohde, “Cell image classification: A comparative overview,” *Cytometry Part A*, vol. 97, pp. 347–362, 2020.

[44] H. M. Rostam, P. M. Reynolds, M. R. Alexander, N. Gadegaard, and A. M. Ghaemmaghami, “Image based machine learning for identification of macrophage subsets,” *Scientific Reports*, vol. 7, p. 3521, 2017.

[45] N. Pavillon, A. J. Hobro, S. Akira, and N. I. Smith, “Noninvasive detection of macrophage activation with single-cell resolution through machine learning,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 115, pp. E2676–E2685, 2018.

[46] N. Emami, Z. Sedaei, and R. Ferdousi, “Computerized cell tracking: Current methods, tools and challenges,” *Visual Informatics*, vol. 5, pp. 1–13, 2021.

[47] X. Zhang, R. Goncalves, and D. M. Mosser, "The isolation and characterization of murine macrophages," *Current Protocols in Immunology*, vol. Chapter 14, p. Unit 14 11, 2008.

[48] W. Ying, P. S. Cheruku, F. W. Bazer, S. H. Safe, and B. Zhou, "Investigation of macrophage polarization using bone marrow derived macrophages," *Journal of Visualized Experiments*, vol. 10.3791/50323, 2013.

[49] K. E. Magnusson, J. Jalden, P. M. Gilbert, and H. M. Blau, "Global linking of cell tracks using the viterbi algorithm," *IEEE Transactions on Medical Imaging*, vol. 34, pp. 911–929, 2015.

[50] K. E. G. Magnusson, "Segmentation and tracking of cells and particles in time-lapse microscopy," *PhD dissertation, KTH Royal Institute of Technology*, 2016.

[51] J. Weischenfeldt and B. Porse, "Bone marrow-derived macrophages (bmm): Isolation and applications," *Cold Spring Harbor Protocols*, vol. 2008, p. pdb.prot5080, 2008.

[52] F. M. Marim, T. N. Silveira, D. S. Lima, and D. S. Zamboni, "A method for generation of bone marrow-derived macrophages from cryopreserved mouse bone marrow cells," *PLoS One*, vol. 5, p. e15263, 2010.

[53] M. D. Englen, Y. E. Valdez, N. M. Lehnert, and B. E. Lehnert, "Granulocyte/macrophage colony-stimulating factor is expressed and secreted in cultures of murine l929 cells," *Journal of Immunological Methods*, vol. 184, pp. 281–283, 1995.

[54] Y. Sun, H. Do, J. Gao, R. Zhao, M. Zhao, and A. Mogilner, "Keratocyte fragments and cells utilize competing pathways to move in opposite directions in an electric field," *Current Biology*, vol. 23, no. 7, pp. 569–574, 2013.

[55] F. Y. McWhorter, T. Wang, P. Nguyen, T. Chung, and W. F. Liu, "Modulation of macrophage phenotype by cell shape," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 110, pp. 17253–17258, 2013.

[56] Z. Pincus and J. A. Theriot, "Comparison of quantitative methods for cell-shape analysis," *Journal of Microscopy*, vol. 227, pp. 140–156, 2007.

[57] C. K. Williams and C. E. Rasmussen, *Gaussian processes for machine learning*, vol. 2. MIT press Cambridge, MA, 2006.

[58] C. Sammut, *Encyclopedia of Machine Learning*. Springer, 2010.

[59] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *Computational and Applied Mathematics*, vol. 20, pp. 53–65, 1987.

[60] F. Osisanwo, J. Akinsola, O. Awodele, J. Hinmikaiye, O. Olakanmi, J. Akinjobi, *et al.*, “Supervised machine learning algorithms: classification and comparison,” *International Journal of Computer Trends and Technology (IJCTT)*, vol. 48, no. 3, pp. 128–138, 2017.

[61] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, “Smote: synthetic minority over-sampling technique,” *Journal of artificial intelligence research*, vol. 16, pp. 321–357, 2002.

[62] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, *et al.*, “Scikit-learn: Machine learning in python,” *Journal of Machine Learning Re-search, 12*, 2011.

[63] C. Cortes and V. Vapnik, “Support-vector networks,” *Machine learning*, vol. 20, pp. 273–297, 1995.

[64] L. Liu, J. V. Stokes, W. Tan, and S. B. Pruett, “An optimized flow cytometry panel for classifying macrophage polarization,” *Journal of Immunological Methods*, vol. 511, p. 113378, 2022.

[65] A. Shapouri-Moghaddam, S. Mohammadian, H. Vazini, M. Taghadosi, S.-A. Esmaeili, F. Mardani, B. Seifi, A. Mohammadi, J. T. Afshari, and A. Sahebkar, “Macrophage plasticity, polarization, and function in health and disease,” *Journal of Cellular Physiology*, vol. 233, pp. 6425–6440, Sep 2018.

[66] H. M. Rostam, P. M. Reynolds, M. R. Alexander, N. Gadegaard, and A. M. Ghaemmaghami, “Image based machine learning for identification of macrophage subsets,” *Scientific Reports*, vol. 7, p. 3521, 2017.

[67] V. Ulman, M. Maška, K. E. Magnusson, O. Ronneberger, C. Haubold, N. Harder, P. Matula, P. Matula, D. Svoboda, M. Radojevic, *et al.*, “An objective comparison of cell-tracking algorithms,” *Nature methods*, vol. 14, no. 12, pp. 1141–1152, 2017.

[68] P. Matula, M. Maška, D. V. Sorokin, P. Matula, C. Ortiz-de Solórzano, and M. Kozubek, “Cell tracking accuracy measurement based on comparison of acyclic oriented graphs,” *PloS one*, vol. 10, no. 12, p. e0144959, 2015.

[69] M. Maška, V. Ulman, D. Svoboda, P. Matula, P. Matula, C. Ederra, A. Urbiola, T. España, S. Venkatesan, D. M. Balak, *et al.*, “A benchmark for comparison of cell tracking algorithms,” *Bioinformatics*, vol. 30, no. 11, pp. 1609–1617, 2014.

[70] O. Hilsenbeck, M. Schwarzfischer, S. Skylaki, B. Schauberger, P. S. Hoppe, D. Loeffler, K. D. Kokkaliaris, S. Hastreiter, E. Skylaki, A. Filipczyk, *et al.*, “Software tools for single-cell tracking and quantification of cellular and molecular properties,” *Nature biotechnology*, vol. 34, no. 7, pp. 703–706, 2016.

[71] J.-Y. Tinevez, N. Perry, J. Schindelin, G. M. Hoopes, G. D. Reynolds, E. Laplantine, S. Y. Bednarek, S. L. Shorte, and K. W. Eliceiri, “Trackmate: An open and extensible platform for single-particle tracking,” *Methods*, vol. 115, pp. 80–90, 2017.

[72] E. L. Barnhart, K. C. Lee, K. Keren, A. Mogilner, and J. A. Theriot, “An adhesion-dependent switch between mechanisms that determine motile cell shape,” *PLoS Biology*, vol. 9, p. e1001059, 2011.

[73] C. Sloas, S. Gill, and M. Klichinsky, “Engineered car-macrophages as adoptive immunotherapies for solid tumors,” *Frontiers in Immunology*, vol. 12, p. 783305, Nov 2021.

[74] M. Kesapragada, Y.-H. Sun, C. Recendez, D. Fregoso, H.-Y. Yang, E. Aslankoohi, R. Isseroff, M. Rolandi, M. Zhao, and M. Gomez, “Deep learning classification for macrophage subtypes through cell migratory pattern analysis.,” *Front Cell Dev Biol.*, 2024.

[75] C. Atri, F. Z. Guerfali, and D. Laouini, “Role of human macrophage polarization in inflammation during infectious diseases,” *International journal of molecular sciences*, vol. 19, 2018.

[76] T. Chanmee, P. Ontong, K. Konno, and N. Itano, “Tumor-associated macrophages as major players in the tumor microenvironment,” *Cancers*, vol. 6, pp. 1670–1690, 2014.

[77] S. Funes, M. Rios, J. Escobar-Vera, and A. Kalergis, “Implications of macrophage polarization in autoimmunity,” *Immunology*, vol. 154, pp. 186–195, 2018.

[78] P. Krzyszczyk, R. Schloss, A. Palmer, and F. Berthiaume, “The role of macrophages in acute and chronic wound healing and interventions to promote pro-wound healing phenotypes,” *Frontiers in physiology*, vol. 9, p. 419, 2018.

[79] T. A. Wynn and K. M. Vannella, “Macrophages in tissue repair, regeneration, and fibrosis,” *Immunity*, vol. 44, pp. 450–462, 2016.

[80] P. Murray, J. Allen, S. Biswas, E. Fisher, D. Gilroy, and S. G. et al., “Macrophage activation and polarization: nomenclature and experimental guidelines,” *Immunity*, vol. 41, pp. 14–20, 2014.

[81] F. O. Martinez and S. Gordon, “The m1 and m2 paradigm of macrophage activation: time for reassessment,” *F1000prime reports*, vol. 6, p. 13, 2014.

[82] A. Sica, P. Larghi, A. Mancino, L. Rubino, C. Porta, M. G. Totaro, M. Rimoldi, S. K. Biswas, P. Allavena, and A. Mantovani, “Macrophage polarization in tumour progression,” *Semin Cancer Biol*, vol. 18(5), pp. 349–355, 2008.

[83] D. Mosser and J. Edwards, "Exploring the full spectrum of macrophage activation," *Nature Reviews Immunology*, vol. 8, pp. 958–969, 2008.

[84] S. Chatterjee, S. Yabaji, O. Rukhlenko, B. Bhattacharya, E. Waligurski, N. Vallavoju, S. Ray, B. Kholodenko, L. Brown, A. Beeler, A. Ivanov, L. Kobzik, J. Porco, and I. Kramnik, "Channeling macrophage polarization by rocaglates increases macrophage resistance to mycobacterium tuberculosis," *iScience*, vol. 24, p. 102845, 2021.

[85] Y. Du, L. Rong, Y. Cong, L. Shen, N. Zhang, and B. Wang, "Macrophage polarization: an effective approach to targeted therapy of inflammatory bowel disease," *Expert Opin Ther Targets*, vol. 25, pp. 191–209, 2021.

[86] T. Kuntzel and D. Bagnard, "Manipulating macrophage/microglia polarization to treat glioblastoma or multiple sclerosis.," *Pharmaceutics*, vol. 14, p. 344, 2022.

[87] G. Kotwal and S. Chien, "Macrophage differentiation in normal and accelerated wound healing," *Results Probl Cell Differ*, 2017.

[88] H. Rostam, P. Reynolds, M. Alexander, N. Gadegaard, and A. Ghaemmaghami, "Image based machine learning for identification of macrophage subsets," *Scientific Reports*, 2017.

[89] N. Pavillon, A. J. Hobro, S. Akira, and N. I. Smith, "Noninvasive detection of macrophage activation with single-cell resolution through machine learning," *P Natl Acad Sci USA*, vol. 115, pp. E2676–E2685, 2018.

[90] F. Mcwhorter, T. Wang, P. Nguyen, T. Chung, and W. Liu, "Modulation of macrophage phenotype by cell shape," *Proceedings of the National Academy of Sciences of the United States of America*, 2013.

[91] M. Kesapragada, Y.-H. Sun, C. Recendez, D. Fregoso, M. Rolandi, H.-Y. Yang, R. Isseroff, M. Zhao, and M. Gomez, "A data-driven approach to establishing cell motility patterns as predictors of macrophage subtypes and their relation to cell morphology," *bioRxiv*, 2024.

[92] W. Ying, P. Cheruku, F. Bazer, S. Safe, and B. Zhou, "Investigation of macrophage polarization using bone marrow derived macrophages," *Journal of visualized experiments*, 2013.

[93] X. Zhang, R. Goncalves, and D. Mosser, "The isolation and characterization of murine macrophages," *Current protocols in immunology*, 2008.

[94] K. Magnusson, J. Jalden, P. Gilbert, and H. Blau, "Global linking of cell tracks using the viterbi algorithm," *IEEE transactions on medical imaging*, 2015.

[95] F. Chollet, "Keras: Deep learning for humans. https://github.com/fchollet/keras," 2015.

[96] V. Nair and G. Hinton, "Rectified linear units improve restricted boltzmann machines," in *In Proceedings of the 27th international conference on machine learning*, pp. 807–814, ICML-10, 2010.

[97] T. Tieleman and G. Hinton, "Rmsprop: Divide the gradient by a running average of its recent magnitude," *COURSERA: Neural Networks for Machine Learning*, pp. 26–31, 2012.

[98] M. Vicente-Manzanares and A. R. Horwitz, "Cell migration: an overview," *Cell migration*, pp. 1–24, 2011.

[99] K. Zlobina, M. Jafari, M. Rolandi, and M. Gomez, "The role of machine learning in advancing precision medicine with feedback control," *Cell Reports Physical Science*, p. 101149, 2022.

[100] J. Selberg, M. Jafari, C. Bradley, M. Gomez, and M. Rolandi, "Expanding biological control to bioelectronics with machine learning," *APL Materials*, vol. 8, no. 12, p. 120904, 2020.

[101] J. Selberg, M. Jafari, J. Mathews, M. Jia, P. Pansodtee, H. Dechiraju, C. Wu, S. Cordero, A. Flora, N. Yonas, *et al.*, "Machine learning-driven bioelectronics for closed-loop control of cells," *Advanced Intelligent Systems*, vol. 2, no. 12, p. 2000140, 2020.

[102] M. Jia, M. Jafari, P. Pansodtee, M. Teodorescu, M. Gomez, and M. Rolandi, "A multi-ion electrophoretic pump for simultaneous on-chip delivery of h+, na+, and cl-," *APL Materials*, vol. 10, no. 4, p. 041112, 2022.

[103] R. Owens, P. Kjall, A. Richter-Dahlfors, and F. Cicoira, "Organic bioelectronics-novel applications in biomedicine. preface.," *Biochimica et biophysica acta*, vol. 1830, no. 9, pp. 4283–4285, 2013.

[104] M. F. Griffin, P. E. Butler, A. M. Seifalian, and D. M. Kalaskar, "Control of stem cell fate by engineering their micro and nanoenvironment," *World journal of stem cells*, vol. 7, no. 1, p. 37, 2015.

[105] D. T. Simon, E. O. Gabrielsson, K. Tybrandt, and M. Berggren, "Organic bioelectronics: bridging the signaling gap between biology and technology," *Chemical Reviews*, vol. 116, no. 21, pp. 13009–13041, 2016.

[106] X. Strakosas, M. Seitanidou, K. Tybrandt, M. Berggren, and D. T. Simon, "An electronic proton-trapping ion pump for selective drug delivery," *Science Advances*, vol. 7, no. 5, p. eabd8738, 2021.

[107] M. Zhao, “Electrical fields in wound healing—an overriding signal that directs cell migration,” in *Seminars in cell & developmental biology*, vol. 20, pp. 674–682, Elsevier, 2009.

[108] S. Shaner, A. Savelyeva, A. Kvartuh, N. Jedrusik, L. Matter, J. Leal, and M. Asplund, “Bioelectronic microfluidic wound healing: a platform for investigating direct current stimulation of injured cell collectives,” *Lab on a Chip*, 2023.

[109] E. Shirzaei Sani, C. Xu, C. Wang, Y. Song, J. Min, J. Tu, S. A. Solomon, J. Li, J. L. Banks, D. G. Armstrong, *et al.*, “A stretchable wireless wearable bioelectronic system for multiplexed monitoring and combination treatment of infected chronic wounds,” *Science Advances*, vol. 9, no. 12, p. eadf7388, 2023.

[110] Y. Jiang, A. A. Trotsyuk, S. Niu, D. Henn, K. Chen, C.-C. Shih, M. R. Larson, A. M. Mermin-Bunnell, S. Mittal, J.-C. Lai, *et al.*, “Wireless closed-loop smart bandage for chronic wound management and accelerated tissue regeneration,” *bioRxiv*, pp. 2022–01, 2022.

[111] T. J. Zajdel, G. Shim, L. Wang, A. Rossello-Martinez, and D. J. Cohen, “Scheepdog: programming electric cues to dynamically herd large-scale cell migration,” *Cell systems*, vol. 10, no. 6, pp. 506–514, 2020.

[112] J. Wosik, W. Chen, K. Qin, R. M. Ghobrial, J. Z. Kubiak, and M. Kloc, “Magnetic field changes macrophage phenotype,” *Biophysical Journal*, vol. 114, no. 8, pp. 2001–2013, 2018.

[113] B. Song, Y. Gu, J. Pu, B. Reid, Z. Zhao, and M. Zhao, “Application of direct current electric fields to cells and tissues in vitro and modulation of wound electric field in vivo,” *Nature protocols*, vol. 2, no. 6, pp. 1479–1489, 2007.

[114] Y. Sun, G. Luxardi, G. Xu, K. Zhu, B. Reid, B. Guo, C. Lebrilla, E. Maverakis, and M. Zhao, “Surface glycans regulate salmonella infection-dependent directional switch in macrophage galvanotaxis independent of nanh,” *Infection and immunity*, vol. 90, no. 1, pp. e00516–21, 2022.

[115] R. Nuccitelli, P. Nuccitelli, C. Li, S. Narsing, D. Pariser, and K. Lui, “The electric field near human skin wounds declines with age and provides a noninvasive indicator of wound healing,” *Wound repair and regeneration : official publication of the Wound Healing Society [and] the European Tissue Repair Society*, vol. 19, p. 645–655, 2011.

[116] G. Marquez, *Machine Learning and Control Methods for Biological Systems: Towards Advancing Precision Medicine.* PhD thesis, University of California, Santa Cruz, 2023. ProQuest ID: Marquez_ucsc_0036E_12284. Merritt ID: ark:/13030/m5nx0cnn.

[117] J. Mao, L. Chen, Z. Cai, S. Qian, Z. Liu, B. Zhao, Y. Zhang, X. Sun, and W. Cui, "Advanced biomaterials for regulating polarization of macrophages in wound healing," *Advanced Functional Materials*, vol. 32, no. 12, p. 2111003, 2022.

[118] R. J. Korthuis, "Skeletal muscle circulation," in *Colloquium Series on Integrated Systems Physiology: From Molecule to Function*, vol. 3, pp. 1–144, Morgan & Claypool Life Sciences, 2011.

[119] M. Rodrigues, N. Kosaric, C. A. Bonham, and G. C. Gurtner, "Wound healing: a cellular perspective," *Physiological reviews*, vol. 99, no. 1, pp. 665–706, 2019.

[120] K. Kikuchi, L. Galera-Laporta, C. Weatherwax, J. Y. Lam, E. C. Moon, E. A. Theodorakis, J. Garcia-Ojalvo, and G. M. Süel, "Electrochemical potential enables dormant spores to integrate environmental signals," *Science*, vol. 378, no. 6615, pp. 43–49, 2022.

[121] S. Doron and S. Gorbach, "Bacterial infections: Overview," in *International Encyclopedia of Public Health* (H. K. K. Heggenhougen, ed.), pp. 273–282, Oxford: Academic Press, 2008.

[122] L. H. Duc, H. A. Hong, N. Fairweather, E. Ricca, and S. M. Cutting, "Bacterial spores as vaccine vehicles," *Infection and Immunity*, vol. 71, no. 5, pp. 2810–2818, 2003.

[123] A. Andersson, U. Rönner, and P. E. Granum, "What problems does the food industry have with the spore-forming pathogens bacillus cereus and clostridium perfringens?," *International Journal of Food Microbiology*, vol. 28, no. 2, pp. 145–155, 1995. Physiology of Food Poisoning Microorganisms, AAIR Concerted Action PL920630.

[124] K. Toyota, "Bacillus-related spore formers: Attractive agents for plant growth promotion," *Microbes and Environments*, vol. 30, no. 3, pp. 205–207, 2015.

[125] P. Chattopadhyay, G. Banerjee, and S. Mukherjee, "Recent trends of modern bacterial insecticides for pest control practice in integrated crop management system," *3 Biotech*, vol. 7, no. 1, p. 60, 2017.

[126] V. Profillidis and G. Botzoris, "Chapter 5 - statistical methods for transport demand modeling," in *Modeling of Transport Demand* (V. Profillidis and G. Botzoris, eds.), pp. 163–224, Elsevier, 2019.