Establishing cell motility patterns as predictors of macrophage subtypes and their relation to cell morphology

Manasa Kesapragada[1, *], Yao-Hui Sun[2, *], Cynthia Recendez[2], Daniel Fregoso[3], Hsin-ya Yang[3], Rivkah Isseroff[3], Min Zhao[2,3], and Marcella Gomez[1, +]

[1]Department of Applied Mathematics, University of California, Santa Cruz, Santa Cruz, CA, 95060, United States of America

[2]Department of Ophthalmology & Vision Science, School of Medicine, University of California, Davis, Sacramento, CA, 95817, United States of America

[3]Department of Dermatology, School of Medicine, UC Davis, Sacramento, CA 95816, United States of America

*These authors contributed equally to this work: Manasa Kesapragada: mkesapra@ucsc.edu, Yao-Hui Sun: yhsun@ucdavis.edu

[+]Corresponding to: Marcella Gomez: mgomez26@ucsc.edu

Key words: Macrophage, morphology, motility, clustering, machine learning

Short title: Machine learning to classify macrophages

Significance:

Previous work has shown that macrophage phenotypes can be distinguished by their morphological characteristics. We extend this work to show that distinct motility patterns are linked to macrophage morphology. Thus, motility patterns can be used to differentiate phenotypes. This can enable high-throughput classification of cell phenotypes without regard for the high-resolution images needed to back out morphological characteristics. Furthermore, combining motility-based features with morphological information improves prediction of macrophage subtypes by a machine learning based classification model.

Abstract

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, but their link to motility properties is unknown, despite M1 and M2 macrophages exhibiting distinct migratory behaviors, *in vitro*, in 3D and *in vivo*. 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 (>79%) using only phase contrast time-lapse microscopy and a small unique cell count for training (~250). Thus, the approach combining cell motility and cell morphology information can accurately assess functionally diverse macrophage phenotypes quickly and efficiently. Our approach offers a cost efficient and high through-put method for screening biochemicals targeting macrophage polarization with small datasets.

Introduction

Macrophages 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 (1, 2). Despite their native (M0) status, two major groups of macrophages that display different functional phenotypes are present in the wound bed: M1 (classically activated) and M2 (alternatively activated) macrophages. It is now understood that there in fact exists a continuum of macrophage functions (3, 4), which have been historically binned into two categories labelled M1 and M2, with primarily pro-inflammatory and anti-inflammatory functions, respectively. For simplicity in their reference, from here on out we continue to use the labels M1 and M2. 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 M1 status (5, 6). M1 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 (7). Prolonged persistence of M1 macrophages without transitioning to anti-inflammatory phenotypes, however, is rather detrimental to wound tissues, which will likely stall, or delay wound healing (8, 9). On the other hand, macrophages activated by IL-4 and IL-13 develop into so-called alternatively activated macrophages (M2) that suppress inflammatory reactions and adaptive immune responses and play reparative role during proliferation and maturation phases (10). However, like M1 macrophages, if M2 macrophages persist for too long hypertrophic scars and keloids develop due to excessive collagen production (11, 12).

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 M1/M2 macrophage phenotype markers. Traditionally, characterization of M1 and M2 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 M1and M2 macrophage phenotype *in vivo* is compounded by similar problems in *in vitro*-derived macrophages. For example, Arginase-1, which is a considered a classic M2 marker, is also upregulated in M1 macrophages (13, 14) and protein expression of Arginase-1 or CD206 is too low for reliable flow cytometry detection (4, 5, 11, 15). In conclusion, a more robust M1 vs. M2-discriminating system is critically needed to improve the detection and understanding of macrophage phenotype.

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., M1 vs. M2) 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 in M1 vs. M2 cells.

Macrophages undergo morphological changes, M0 to M1 and M0 to M2 states (2). Studies have shown that cell shape can help control macrophages — for instance, elongating them promotes behavior that enhances healing (1). 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 M1 and M2 macrophages exhibiting distinct migratory behaviors.

Machine learning methods have been successfully applied to automate the classification of different cell types from microscopy images (16, 17) 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 (18). Previous work towards classifying macrophage subtype using machine learning leverages fluorescent dyes (19, 20). Being able to identify cell types from phase contrast microscopy and a small dataset would simplify and accelerate macrophage classification. Furthermore, being able to identify cells based on their migration patterns can eliminate the need for high-quality imaging and result in methods robust to blurred images. Cell tracking can be achieved with image processing techniques in real-time (21). Motility patterns can be extracted from cell trajectories and machine learning used to classify cells based on their dynamic behavior. Finally, combining cell trajectories and morphological features can help elucidate the link between cell migratory behavior, morphology, and cellular phenotype.

Here we apply machine learning to map cell migratory patterns to macrophage subtypes and show this is possible due to the tight relationship between cell morphology and its motile behavior. We also demonstrate that both morphology and motility can be used to enhance prediction of macrophage subtypes. We begin by first validating that our derived macrophage subtypes do present distinct morphological characteristics as has been shown in previous work. We apply analysis to label-free live

cell images consisting of all macrophage subtypes. 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 migration patterns. We then apply discriminant analysis to confirm that this implies a relationship between the labeled cells (e.g., M0, M1, and M2) and motility parameters, thereby, supporting the hypothesis that we can predict macrophage subtypes from motility behavior alone. Finally, to predict macrophage subtype 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. We demonstrate predictions using motility are comparable to those using morphology but not equally good for the same subtype. This suggests that, although, they are tightly linked, they may in fact provide complimentary information. We show that a combined use of motility and morphology shows an increased prediction accuracy in identifying macrophage subtypes. Hence, by combing 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.

Results

Generation and polarization of primary murine macrophages

Bone marrow-derived macrophages (BMDMs) are primary macrophages that are differentiated from bone marrow cells (22, 23). 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 cells from freshly cut femur bones of C57/BL6 mice and grow them in the presence of murine macrophage colony-stimulating factor (M-CSF), as provided by supplementary L929-conditioned medium (24). 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 (M1) or pro-reparative (M2) macrophages by addition of bacterial lipopolysaccharide (LPS) or murine interleukin 4 (IL-4), respectively (25) (**Fig. 1*A***).

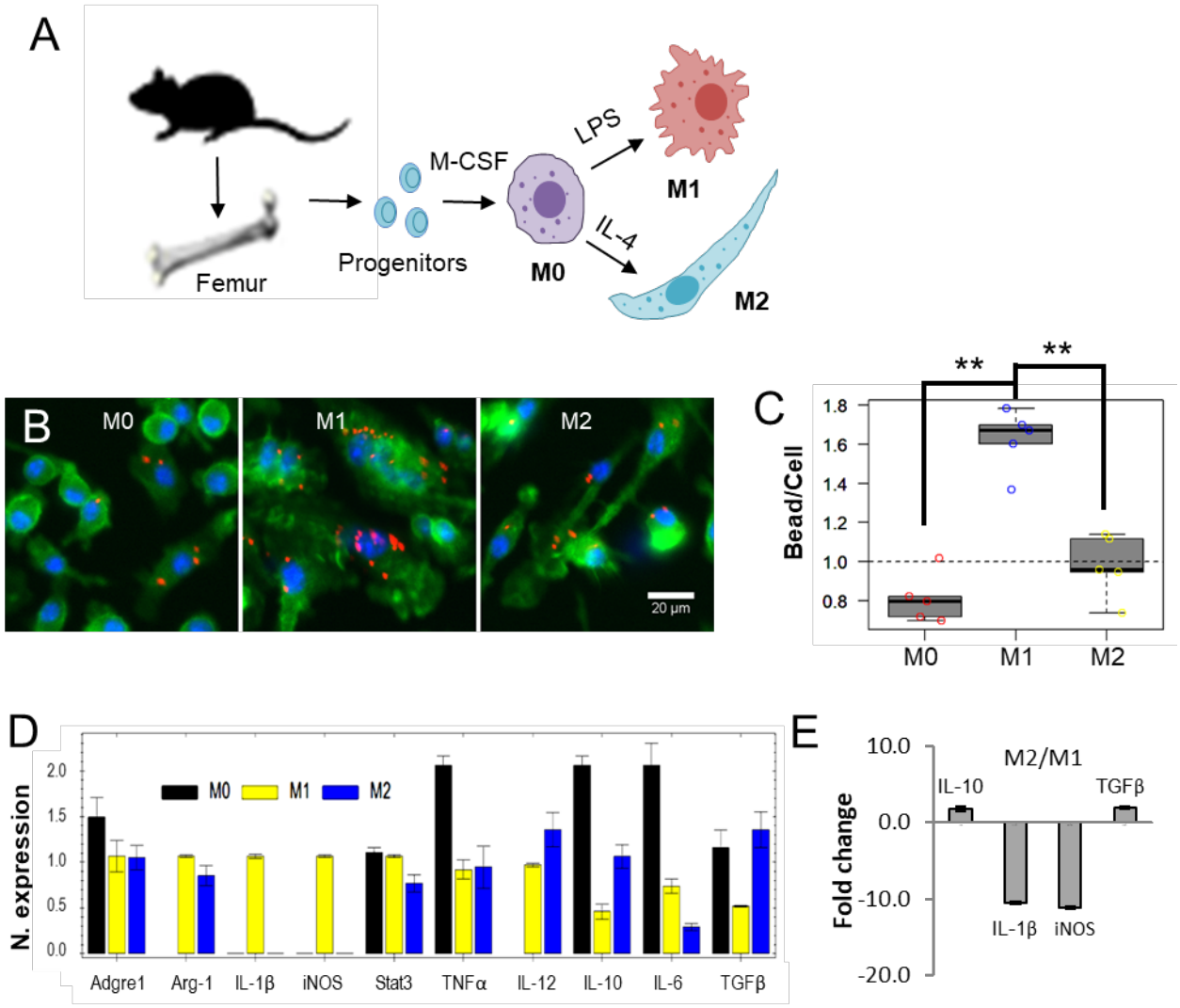


Fig. 1. Generation and characterization of murine bone marrow derived M0, M1 and M2 macrophages
(*A*) Scheme for the generation of mouse bone marrow derived macrophages (BMDMs, M0) and subsequent polarization of M1 (100 ng/ml LPS) and M2 (20 ng/ml mIL-4) macrophages. (*B*) Fluorescent images show M0, 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.

Phagocytosis is a major function of macrophages to capture and kill pathogens, as well as to remove apoptotic cells or cell debris (26). We confirm the phagocytic capacity of the naive BMDMs and LPS or IL-4 primed M1 or M2 macrophages by challenging them with fluorescently labelled latex beads (**Fig. 1*B***). While polarized M1 and M2 macrophages were both found to have enhanced phagocytic capacity, it is worth noting that the M1 macrophage that were primed by pathogenic bacterial LPS had a significantly higher number of internalized latex beads (1.624 ± 0.070, bead/cell) when compared with M0 macrophages or M2 macrophages primed by IL-4 (0.811 ± 0.057 and 0.979 ± 0.072, respectively) (**Fig. 1*C***), as has also been reported by other groups (27).

To further characterize the phenotype of polarized macrophages, we performed quantitative PCR to determine the relative mRNA expression of a panel of selected target genes and transcription factors. **Fig. 1*D*** shows typical profiles and variations of macrophage polarization-associated genes of unpolarized and differentially polarized macrophages by LPS or IL-4 for 48 h. Specifically, the transcription levels of IL-1β and inducible nitric oxide synthase (iNOS), two well-established markers of proinflammatory macrophages, are increased more than 10 folds in LPS stimulated M1 macrophages compared to that of IL-4 stimulated M2 macrophages. In contrast, the transcriptions of IL-10 and transforming growth factor beta (TGFβ), two characterized pro-healing macrophage markers, are significantly higher in M2 macrophages, compared to that of M1 macrophages (**Fig. 1*E***). Thus, we have generated and polarized primary macrophages with characterized phenotypes that are applicable and appropriate for the proposed downstream studies and analysis.

Generation of time-series data capturing spontaneous cell migratory patterns

Using time-lapse microscopy three sets of images M0, M1, and M2 are generated, named after their most representative macrophage subtypes with an interval of 5 minutes per frame for 180 minutes, for a total of 37 images in each image set (**Fig. S*1***). These are the phase contrast images without ground truths or cell labels except the image names, which are subsequently cultured. 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. To identify the existing principal shape modes, we treat all cells in each of the 3 time-lapse data sets as independent data points. With 36 time points, we have a total of 108 images with 2329 cells (non-unique). For the remainder of the analysis, we use uniquely identified cells from the time-lapse data sets. Cell segmentation and tracking details can be found in the

methods section. To uniquely identify the cells, cell tracking is followed by four cell filters in the post processing which ignore the cells partially present on the boundary of the image frames across the time-series, cells which undergo mitosis and cell death, since we want to map morphology to motility, we identify cells such that the cell area does not change more than their average cell size through the course of the experiment and we ignore the cells which are very small compared to the average group of cells. This gives us a total of 269 uniquely identified cells from the three datasets. Of the 269 cells, 83 are from M0 image, 106 are from M1 image and 80 are from M2 image. For the characterization of cell subtypes, cell morphology parameters are computed based on the first-time step and motility parameters are determined based on the time-lapse data.

Three principal shape modes characterize observed macrophage morphologies

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 (28). 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 (**Fig. 2*A***). Using Celltool (29), we extracted and aligned (30) label-free phase contrast images of a mixed population of live macrophage (n = 2329). We leveraged here all cells in each of the 3 time-lapse data sets with 36 time points each for a total of 108 images with 2329 cells. The data set contains shape information for M0 (n = 965), M1 (n = 698) and M2 (n = 666) (**Fig. 2*B***). Quantitative analysis revealed that macrophage shapes could be described with just three orthogonal modes of shape variability, which together account for 86% of the total variation in shape. In general, these modes can be characterized as either rounded 'O' shape (mode 1) or rounded with multiple random protrusions (mode 2), or an elongated 'fan' shape (mode 3) (**Fig. 2*C***). These shape modes provide a meaningful and concise quantitative description of macrophage morphology that could be defined in very few parameters, providing significant insights that may underly macrophage phenotypes and functionalities depending on their activation status.

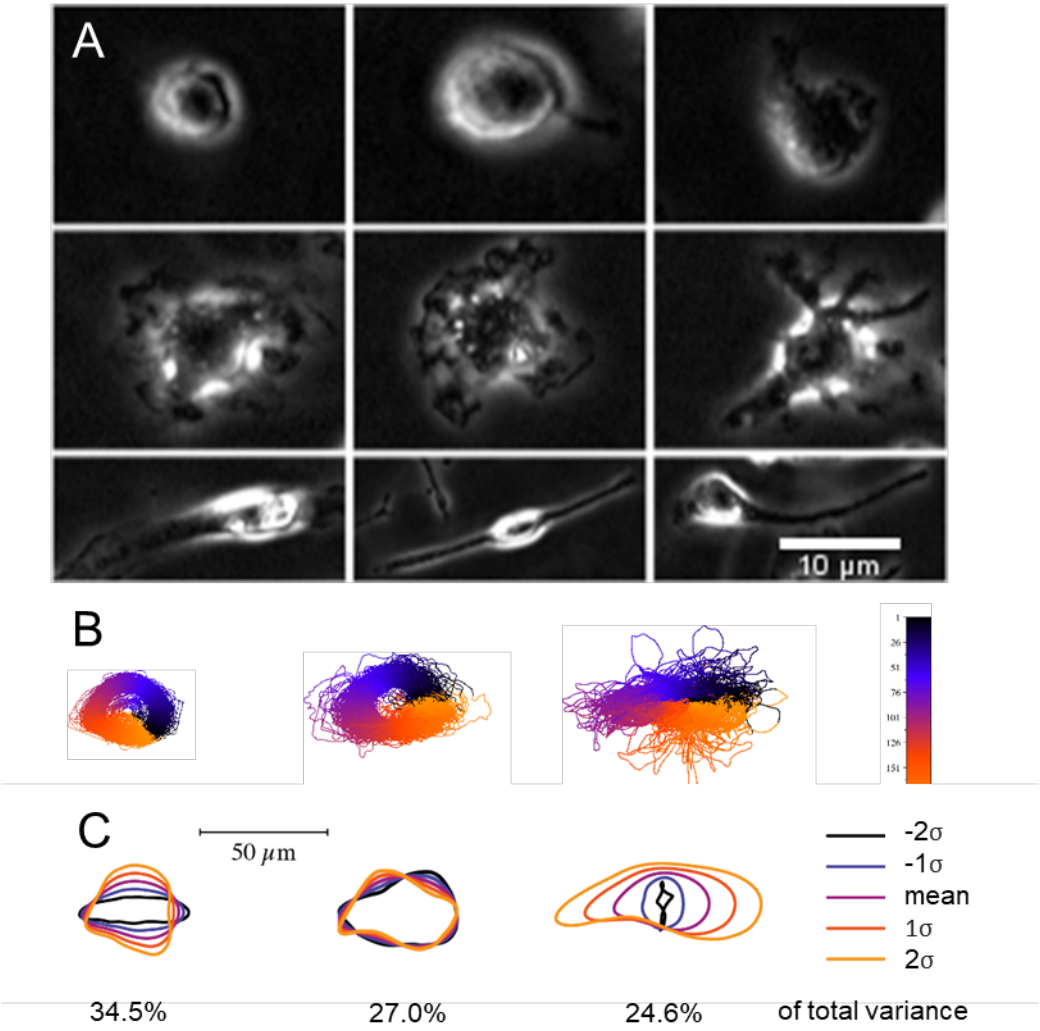
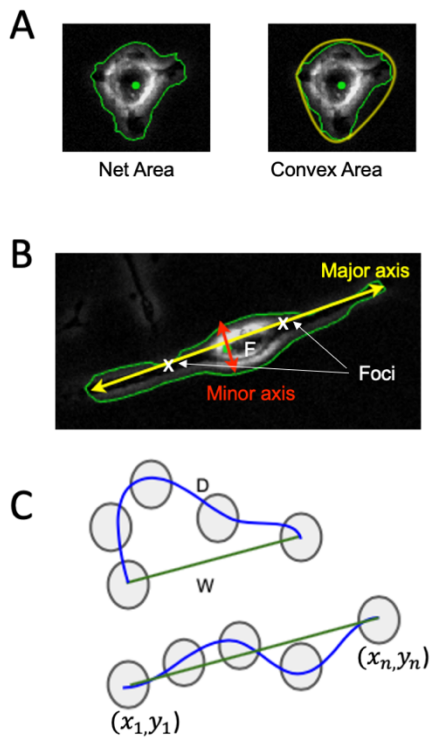


Fig. 2. Macrophage shapes are described by three principal shape modes

(*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.

Morphological characterization of macrophage phenotypes

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. Here we aim to characterize and quantify the three primary shape modes identified above through clearly defined morphological parameters. In the literature, researchers have used various metrics to describe cell shape (28, 31). 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. Different combinations and number of parameters were explored. To characterize the morphological features in defining macrophage phenotypes, we arrived at the following three parameters: compactness - to identify the complicated and irregular boundaries of the cells; eccentricity - to quantify the cell elongation; and solidity - to find the cell density, which are sufficient to describe and distinguish the three principal shape modes identified in the combined macrophage populations. These three parameters gave sufficiently independent information, whose combination distinguished the three primary modes.



D

Morphological Parameters

$$\text{eccentricity} = \frac{\text{distance between the focal points } (F)}{\text{major axis length}}$$

$$\text{solidity} = \frac{\text{area}}{\text{convex area}}$$

$$\text{compactness} = \frac{4\pi * \text{area}}{(\text{perimeter})^2}$$

Motility parameters

$$\text{speed} = \frac{\text{Distance travelled } (D)}{\text{time } (T)}$$

$$\text{persistence} = \frac{\text{Distance between first and last points } (W)}{\text{Total distance travelled } (D)}$$

$$x_{spread} = \max(x) - \min(x),\ y_{spread} = \max(y) - \min(y)$$

$$\text{trajectory spread} = \sqrt{x_{spread} * y_{spread}}$$

Fig. 3. Defining the morphological and motility parameters

Morphology is defined by metrics related to the (*A*) area and perimeter of the object and its convex hull, in additional 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.

Examples of the shape parameter values of three major cell shapes in (**Fig. S2**) with low, medium, and high parameter values for each different shape shows that the combination of these parameters is enough to identify the three primary shape modes. **Fig.3** 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 labelled as M0, M1, and M2 from their image names. We perform *k*-means clustering (32) with an optimal *k* value of 4 derived using silhouette score (33) 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 (**Fig. S3**). **Fig. 4*A*** shows the cluster assignment for the cells within the three-dimensional space of their morphological parameters' compactness, eccentricity, and solidity. The identified 269 cells from parent images M0, M1 and M2 are distributed into four clusters. For the box plots (**Fig. 4*B*-*D***), the Mann-Whitney U test (MWW) (51), a nonparametric null hypothesis test, is performed between clusters. We can observe that the p-value between any of the two clusters for compactness, eccentricity, and solidity plots is less than or equal to 1e-5, representing a statistically highly significant difference in parameter values between all the clusters. The following observations are made about each cluster where the data points are represented as swarm plots (34, 35):

Cluster 0: The eccentricity values are the lowest, with their compactness and solidity values being highest, represent that the cells in this cluster are mostly circular, have no/very less uneven boundaries, and are highly dense.

Cluster 1: The eccentricity values remain lower, with their compactness and solidity values being least, represent that the cells in this cluster are in shape closer to being circular, have the most uneven boundaries, and are not likely dense.

Cluster 2: The eccentricity values are higher, with their compactness values being above average and solidity values being higher, represent that the cells in this cluster are in a shape closer to elliptical, have fewer uneven boundaries, and are dense.

Cluster 3: The eccentricity values are higher, with their compactness values lowest and solidity values spanned from least to average, represent that the cells in this cluster are elliptical, have the most uneven boundaries, and are not likely dense enough. These cells do not coincide with any of the primary shape

modes identified. We attribute this cluster to misidentified cells in the cell segmentation process and do not use these cells going forward.

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. To check the accuracy of the macrophage classification of this clustering analysis, the parent image labels- M0/M1/M2 are used as the ground truth labels. **Fig. 4*E*** shows the distribution of the cells from the parent images to their respective clusters in percentages where in a total of 269 cells, 83 are from M0 image, 106 are from M1 image and 80 are from M2 image; the number of cells binned into clusters 0, 1, 2, and 3 are 51, 60, 91, and 67, respectively. We can observe that the highest percentage of cells from the M0 image are distributed in Cluster 0, 80% of cells from the M1 image are in Cluster 1 and nearly half of the percentage of cells from the M2 image are in Cluster 2. As we only consider the dominant representations in each image and as per our morphological observations, we conclude that the primary shape modes of M0 cells are circular in shape, M1 cells have more protrusions and M2 cells are elongated. In the following sections, we only consider clusters 0-2, which are most highly correlated to macrophage subtypes M0, M1, and M2.

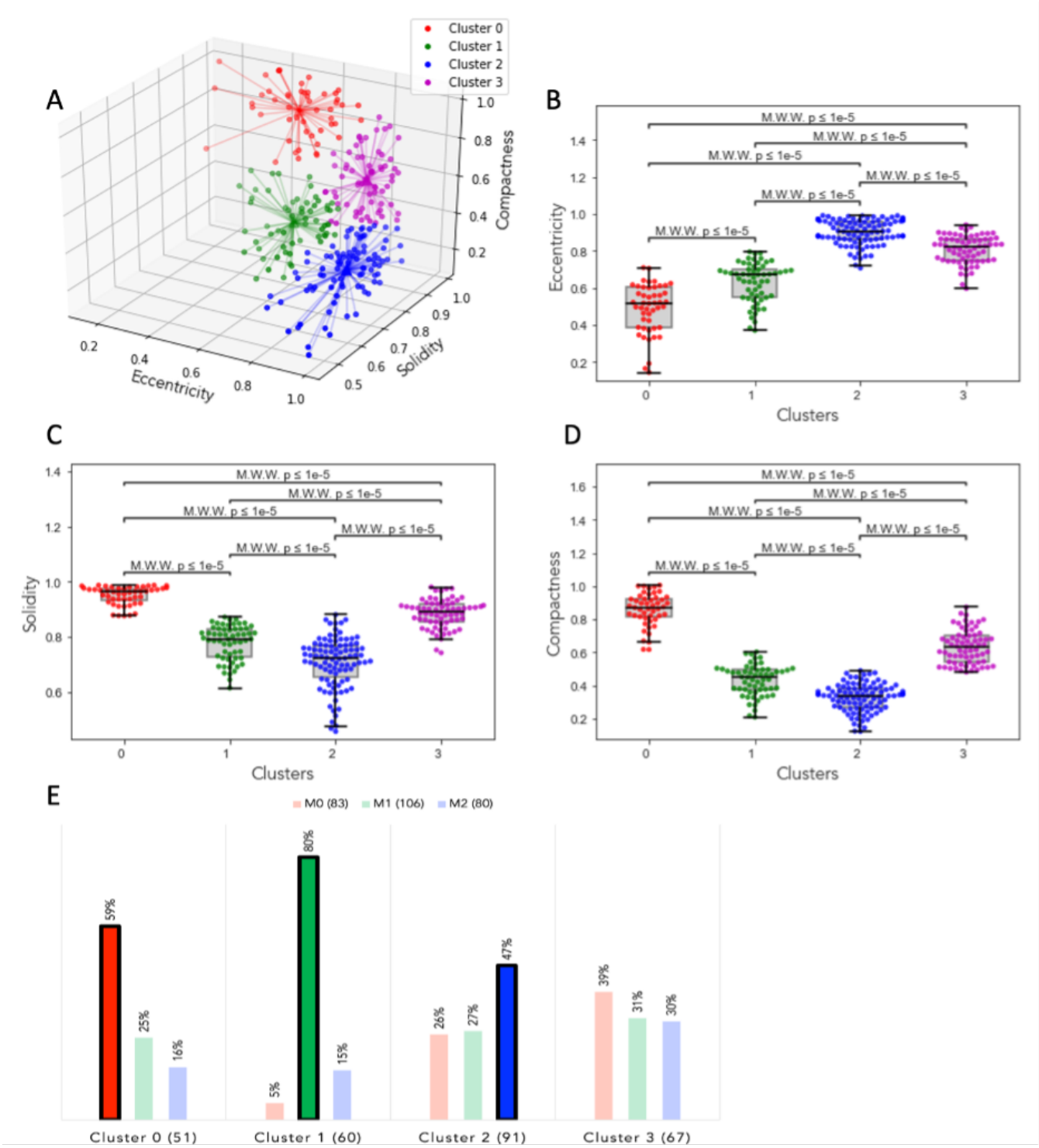


Fig. 4. Clustering macrophages based on morphological parameters

(*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*). (*E*) Percentage comparison of the cells from their clusters with respect to their parent image labels M0, M1 and M2.

Cell morphology can be predicted from its observed motility patterns

We would like to investigate the relationship between the morphological features of the cells in study and their motile behavior. 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 0-2 and quantify their corresponding motility patterns. These cells include 51 cells in cluster 0, 60 cells in cluster 1, and 91 cells in cluster 2. 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. 5*A***). We plot the dynamics of speed and persistence in the intervals between 60 minutes to 120 minutes, which best represent the steady state behavior of the cells as the cells do not show significant changes in the first 60 minutes and the last 120-180 minutes.

We use the Gaussian process regressor (GPR) model (36, 37) to plot the speed and persistence of the cells. A Gaussian process is a probability distribution over possible functions that fit a set of points, where the results for each cluster are indicated by a different color (**Fig. 5**). The solid line indicates the mean based on the GPR model, the dashed lines indicate the average value computed from the data at each time point, and the band provides a 95% confidence interval. We noticed that the speed of the cells in Cluster 1 (protruded cells) is highest, followed by the speed of the cells in Cluster 2 (elongated cells) and the lowest speed observed for the cells in Cluster 0 (circular cells) (**Fig. 5*A***). We also noticed that the persistence of the cells in Cluster 2 is highest, followed by the persistence of the cells in Cluster 0 and the lowest persistence observed for the cells in Cluster 1 (**Fig. 5*B***). 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.

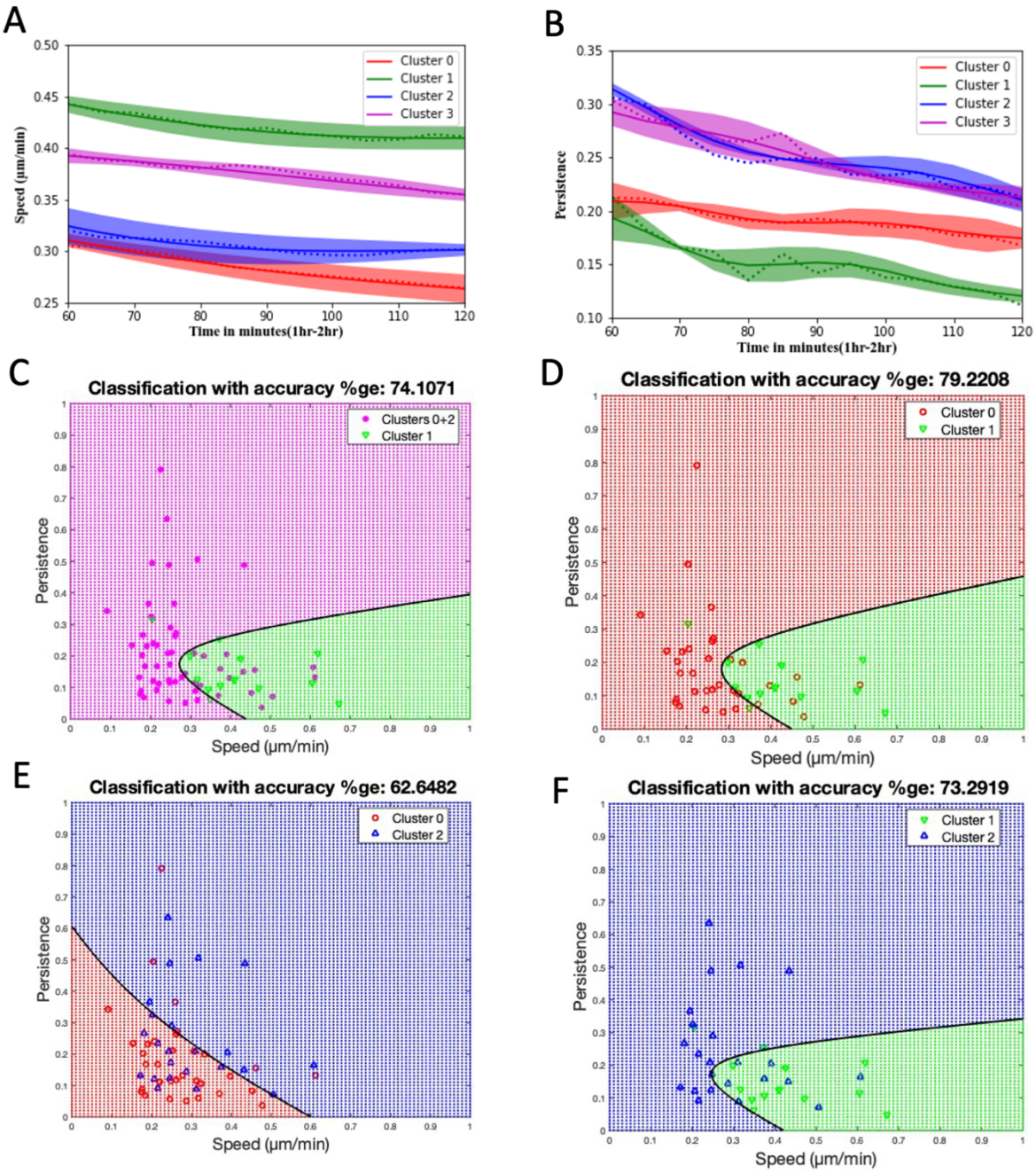


Fig. 5. Classification and discrimination of morphological groups using motility parameters

(*A*) Speed and (*B*) Persistence, where dotted lines are the mean values of the corresponding clusters, 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 Cluster 1 cells and combined clusters 0 and 2 cells with a classification accuracy of around 74%. (*D*) Plot showing the quadratic classification of the cells using Speed vs Persistence between Cluster 0 cells and Cluster 1 cells with a

classification accuracy of around 79%. (*E*) Plot showing the quadratic classification of the cells using Speed vs Persistence between Cluster 0 cells and Cluster 2 cells with a classification accuracy of around 63%. (*F*) Plot showing quadratic the classification of the cells using Speed vs Persistence between Cluster 1 cells and Cluster 2 cells with a classification accuracy of around 73%. The black line in *C-F* is the decision boundary between the classes.

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. 5**, showing the results for various pairs of classification labels). Classification error percentages confirm that the Cluster 1(protruded cells) can be distinguished combined or independently from the Cluster 0 (circular cells) and Cluster 2 (elongated cells) using the speed and persistence motility parameters (**Fig. 5*C, D, F***). However, the classification accuracy percentage between Cluster 0 (circular cells) and Cluster 2 (elongated cells) is comparatively lower around 63% (**Fig. 5*E***). The ability to predict morphological features based on cell motility imply a strong relationship between the two.

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 that in general there is a strong relationship between motile behavior and morphology. Thus, we should expect that macrophage subtypes can also be uniquely characterized by their 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 taking the supervised learning approach. Based on this motility parameter analysis given in (**Fig. S4*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 (**Fig. S4*C-F***) to quantify the motility of the cells in M0, M1 and M2 images.

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 cell trajectory spread, speed and persistence to extract motility features for all cells in M0, M1, and M2 images (**Fig. 6**). For our machine learning classifier, we consider a support vector machine (SVM) (38) model, which is suitable for small but representative datasets since the solution is unique and optimal with respect to the dataset. We have 289 cells from all the three image sets where 83 are from M0 image, 106 from M1 image and 80 from M2 image. We shuffle the data and split 80% for training and 20% for testing. The training data consists of 67 cells from M0, 83 cells from M1 and 65 cells from M2 image. We now use Synthetic Minority Oversampling Technique (SMOTE) (39) to balance the minority classes M0 and M2. The final training dataset has 83 cells each from the M0, M1 and M2 images, for a total of 249 cells. 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. 6*A*** shows the schematic of the machine learning model with the features, input labels and the expected output. To 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 (Trajectory Spread, Speed, Persistence) as inputs. The third model is trained using both morphological and motility features as inputs. The results are shown in **Fig. 6*B-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 69%, 71% and 71% respectively (**Fig. 6*B***). While the motility trained model improved the prediction accuracies for M0 (75%) and M1 (75%), the model experienced a difficulty in identifying the M2 cells (41%) (**Fig. 6*C***). Strikingly, the combined morphology and motility trained model significantly improved prediction accuracies for all the M0, M1

and M2 cells to 81%, 79% and 79% respectively (**Fig. 6*D***). 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 (**Table. S1**). To validate that the original size of the dataset, albeit small, does not lead to overfitting, we shuffle the original data (with no data balancing and augmentation) and use stratified k-fold validation with k = 2, which maintains the same M0, M1, M2 cells ratio throughout the k-folds as the ratio of the original dataset. This gives us a 50:50 split of the dataset for training and testing consisting of 41/42 M0, 60 M1 and 43 M2 cells. We can observe from (**Fig. S5**) that with just 50% training data and without any data augmentation for the minority classes, the model performance is statistically significant using both the morphology and motility parameters. We note that these results were obtained with a smaller training set overall (only 50% used for training and no data augmentation applied). This supports the claim that there exists a relationship between cell motility, cell morphology, and macrophage subtypes.

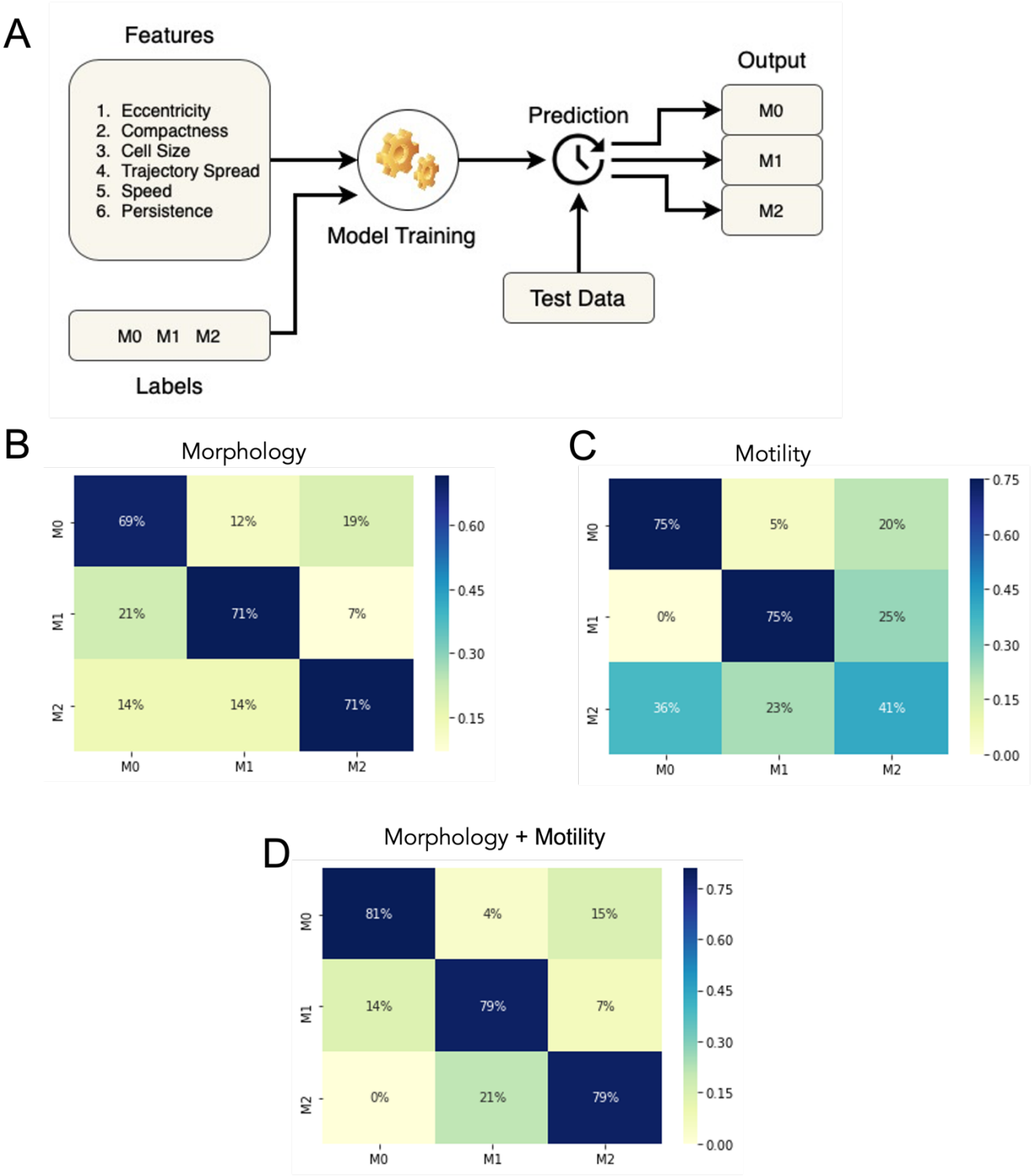


Fig. 6. Classification of macrophage subtypes using SVM on morphological and motility parameters
(*A*) Schematics of the input and output of the machine learning model. SVM classification results showing the percentage distribution of the cells in a confusion matrix with columns being the predicted labels and rows being the actual labels. (*B*) Input features with morphological parameters (Eccentricity, Compactness, Cell Size). (*C*)

Input features with motility parameters (Trajectory Spread, Speed, Persistence). (*D*) Input features with both morphological and motility parameters.

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. 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 (28, 40). 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 of 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 (41), benchmark techniques (42, 43) and software (13-15, 44, 45) have been developed and made available to the scientific community which are more ubiquitous and robust to blurry images.

Independently, motility and morphological features can provide ambiguous discriminant information depending on the macrophage subtype. For example, M2 prediction accuracies are low when using motility information (41%) but higher for M0 and M1 subtypes when compared to morphology-based prediction accuracies. Thus, combing 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.

We note that the images used for analysis required that cell densities remain low. 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. Hence, this cluster was ignored for the analysis.

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 these cells to get distributed in the above three clusters.

In summary, using image processing techniques, we can characterize cell behavior and migratory patterns that can be fed into our machine learning model to predict macrophage subtype. Analysis of these dynamics without any external cues given like electrical guidance and staining would give us insights which are critical to quantify 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.

Materials and Methods

Isolation and culture of bone marrow-derived macrophages

The mouse strains used were in a C57BL/6 background (both male and female mice were used in experiments). 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 University of California (UC), Davis. All animal experiments were performed in accordance with 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 by following standard procedures as previously described (46). Briefly, bone marrow cells were cultured in DMEM (Invitrogen) with 10% Fetal Bovine Serum (Invitrogen) and 1× Antibiotic-Antimycotic solution (Invitrogen), supplemented with 20% L-929 conditioned medium for 6 days (plus an extra feed at day 3), followed by one-day culture without the conditioned medium. Adherent macrophages were then harvested by gently scraping with a "policeman" cell scraper and used for subsequent experiments accordingly. Cell viability was determined by trypan blue staining and counting.

Activation of bone marrow-derived macrophages

BMDMs were seeded in 24-well tissue culture-treated plate (2×10^5 cells per well) and cultured in RPMI-1640 (Invitrogen) with 10% Fetal Bovine Serum (Invitrogen) and 1× Antibiotic-Antimycotic solution (Invitrogen) for overnight. Either 100 ng/ml LPS (Sigma, Cat number: L6143) or 20 ng/ml recombinant mouse IL-4 (R&D Systems, Cat number: 404-ML) were added in the culture medium for M1 or M2 activation, respectively (25). Two days after stimulation, activated M1 and M2 macrophages were used

for morphological and motility characterizations, as well as functional studies. Macrophages without stimulation were served as M0 control.

Time-lapse recording

Macrophages were seeded in 24-well tissue culture-treated plate, cultured, and activated by following procedures as described above. 48 h after stimulation, cells were monitored with a Carl Zeiss Observer Z1 inverted microscope equipped with a motorized stage and an incubation chamber (37 ºC and 5% CO2). Time-lapse contrast images were captured using MetaMorph program (Molecular Devices) with a Retiga R6 (QImaging) scientific CCD camera (6 million, 4.54 µm pixels). Typically, in each experiment, 4 fields of each condition under a long-distance 20× lens were chosen. Images were taken at 5 min intervals for up to 3 h unless stated otherwise.

Macrophage phagocytosis assay

BMDMs were seeded on 12 mm cover slips that were pre-coated with FNC Coating Mix (Athena, Cat number: 0407) for 4 hours at RT in 24-well plate (2×10^5 cells per well) and cultured in RPMI-1640 (Invitrogen) with 10% Fetal Bovine Serum (Invitrogen) and 1× Antibiotic-Antimycotic solution (Invitrogen) for overnight. Macrophages were activated by following procedures as described above. Fluorescently labeled microspheres (ThermoFisher, Cat number: F8821) were used to challenge macrophages at a multiplicity of infection of 20 (20 beads per cell). After 60 min of incubation, cells were gently washed 3× with PBS to remove excessive and un-adherent beads, and further incubated in culture media in the absence (for M0) or presence of LPS (for M1) or IL-4 (for M2) for additional 16 h. To quantify phagocytosis, macrophages were fixed with 4% paraformaldehyde and permeabilized with 0.1% Tryton-X100 in PBS. F-actin were labelled by FITC-conjugated Phalloidin (Sigma, Cat number: P5282). Nuclei were labeled by Hoechst 33342 (Sigma, Cat number: 94403). Cells were photographed using a Leica TCS SP8 confocal microscope (Leica microsystem). Images were processed using ImageJ. Quantification and comparison of fluorescent intensity were done in images taken in the same batch with the same optical setup and parameters. Beads and nuclei were identified by thresholding and counted using particle analysis function.

Shape modeling with principal component analysis

Phase contrast images of M0, M1 and M2 macrophages were converted into the binary images using custom written Matlab code (Script available upon request). Briefly, we used Matlab edge detection and a basic morphology function to outline the cells/fragments in the phase contrast image. We use the Otsu

method (47) to erase the halo artifacts. If shape was still unsatisfactory, we then used the Lasso tool in Photoshop (Adobe) to manually extract the cell shape. Contours were extracted from the binary images and plotted in Celltool (29), an open source software. For principal component analysis, polygons were extracted from a large population of cell images and mutually aligned. Principle modes of shape variation were determined by principal component analysis of the population of polygonal cell outlines and scaled in terms of the standard deviation of the population for each mode of variation, which was detailed previously (48).

Quantitative real-time PCR

Total RNAs were isolated from naïve M0, LPS-simulated M1 or IL-4-stimulated M2 macrophages using the Trizol- (Life Technologies Cat # 15596026) chloroform method, followed by RNeasy Mini Kit (Qiagen). cDNA was synthesized by QuantiTect Reverse Transcription Kit (Qiagen). qRT-PCRs were performed on the ViiA 7 Real-Time PCR System (Applied Biosystems) using PowerUp SYBR Green Master Mix (ThermoFisher) according to the manufacturer's instructions. Primer sequences are listed in **Table S2**. Relative expression values were normalized by internal housekeeping genes (GAPDHs). Fold changes were calculated as ratios of M2 *versus* M1 macrophages.

Cell segmentation and tracking

The input images are segmented and tracked 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.

Computing speed and persistence

The speed and persistence are computed at each 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:

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

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:

$$persistence = \frac{\sqrt{(x_n - x_i)^2 + (y_n - y_i)^2}}{\Sigma_{i=1}^{n-1}\sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}}$$

Acknowledgements

This research is sponsored by the DARPA Biotechnologies Office (DARPA/BTO) and was accomplished under Cooperative Agreement Number DC20AC00003. Parts of the experiments in the Zhao lab were supported by 5R21AI156409, AFOSR MURI grant FA9550-16-1-0052, and DURIP FA9550-22-1-0149.

Author contributions statement

M.K., Y.S., M.Z., and M.G. designed the research; M.K. performed image processing, parameter characterization, modelling, and machine learning. Y.S., C.R., D.F., and H.Y. conducted experiments and analyzed the data; R.R.I., M.Z., and M.G. provided resources and supervisions; M.K., Y.S., and M.G. wrote the manuscript. All authors contributed to discussion/feedback of the work and approved the manuscript.

Additional information

To include, in this order: Accession codes (where applicable); Competing interests (mandatory statement).

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. T. A. Wynn, K. M. Vannella, Macrophages in Tissue Repair, Regeneration, and Fibrosis. *Immunity* **44**, 450-462 (2016).
2. P. Krzyszczyk, R. Schloss, A. Palmer, F. Berthiaume, The Role of Macrophages in Acute and Chronic Wound Healing and Interventions to Promote Pro-wound Healing Phenotypes. *Front Physiol* **9**, 419 (2018).
3. F. O. Martinez, S. Gordon, The M1 and M2 paradigm of macrophage activation: time for reassessment. *F1000prime reports* **6**, 13 (2014).
4. C. Atri, F. Z. Guerfali, D. Laouini, Role of Human Macrophage Polarization in Inflammation during Infectious Diseases. *International journal of molecular sciences* **19** (2018).
5. D. M. Mosser, J. P. Edwards, Exploring the full spectrum of macrophage activation. *Nature reviews. Immunology* **8**, 958-969 (2008).
6. S. Gordon, Alternative activation of macrophages. *Nature reviews. Immunology* **3**, 23-35 (2003).
7. P. J. Murray, T. A. Wynn, Protective and pathogenic functions of macrophage subsets. *Nature reviews. Immunology* **11**, 723-737 (2011).
8. R. Zhao, H. Liang, E. Clarke, C. Jackson, M. Xue, Inflammation in Chronic Wounds. *International journal of molecular sciences* **17** (2016).
9. M. Hesketh, K. B. Sahin, Z. E. West, R. Z. Murray, Macrophage Phenotypes Regulate Scar Formation and Chronic Wound Healing. *International journal of molecular sciences* **18** (2017).
10. S. Gordon, F. O. Martinez, Alternative activation of macrophages: mechanism and functions. *Immunity* **32**, 593-604 (2010).
11. A. Sindrilaru, K. Scharffetter-Kochanek, Disclosure of the Culprits: Macrophages-Versatile Regulators of Wound Healing. *Advances in wound care* **2**, 357-368 (2013).
12. K. M. Vannella, T. A. Wynn, Mechanisms of Organ Injury and Repair by Macrophages. *Annu Rev Physiol* **79**, 593-617 (2017).
13. F. de Chaumont *et al.*, Icy: an open bioimage informatics platform for extended reproducible research. *Nat Methods* **9**, 690-696 (2012).
14. A. E. Carpenter *et al.*, CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol* **7**, R100 (2006).
15. J. Chalfoun *et al.*, Lineage mapper: A versatile cell and particle tracker. *Scientific reports* **6**, 36984 (2016).
16. N. Meng, E. Y. Lam, K. K. Tsia, H. K. H. So, Large-Scale Multi-Class Image-Based Cell Classification With Deep Learning. *Ieee J Biomed Health* **23**, 2091-2098 (2019).
17. Y. Gu *et al.*, Machine Learning Based Real-Time Image-Guided Cell Sorting and Classification. *Cytom Part A* **95a**, 499-509 (2019).
18. M. Shifat-E-Rabbi, X. W. Yin, C. E. Fitzgerald, G. K. Rohde, Cell Image Classification: A Comparative Overview. *Cytom Part A* **97**, 347-362 (2020).
19. H. M. Rostam, P. M. Reynolds, M. R. Alexander, N. Gadegaard, A. M. Ghaemmaghami, Image based Machine Learning for identification of macrophage subsets. *Sci Rep-Uk* **7** (2017).
20. N. Pavillon, A. J. Hobro, S. Akira, N. I. Smith, Noninvasive detection of macrophage activation with single-cell resolution through machine learning. *P Natl Acad Sci USA* **115**, E2676-E2685 (2018).
21. N. Emami, Z. Sedaei, R. Ferdousi, Computerized cell tracking: Current methods, tools and challenges. *Vis Inform* **5**, 1-13 (2021).
22. J. Weischenfeldt, B. Porse, Bone Marrow-Derived Macrophages (BMM): Isolation and Applications. *CSH Protoc* **2008**, pdb prot5080 (2008).
23. F. M. Marim, T. N. Silveira, D. S. Lima, Jr., D. S. Zamboni, A method for generation of bone marrow-derived macrophages from cryopreserved mouse bone marrow cells. *PLoS One* **5**, e15263 (2010).
24. M. D. Englen, Y. E. Valdez, N. M. Lehnert, B. E. Lehnert, Granulocyte/macrophage colony-stimulating factor is expressed and secreted in cultures of murine L929 cells. *J Immunol Methods* **184**, 281-283 (1995).
25. W. Ying, P. S. Cheruku, F. W. Bazer, S. H. Safe, B. Zhou, Investigation of macrophage polarization using bone marrow derived macrophages. *J Vis Exp* 10.3791/50323 (2013).
26. G. J. Cannon, J. A. Swanson, The macrophage capacity for phagocytosis. *J Cell Sci* **101 (Pt 4)**, 907-913 (1992).
27. R. S. Lam *et al.*, Unprimed, M1 and M2 Macrophages Differentially Interact with Porphyromonas gingivalis. *PLoS One* **11**, e0158629 (2016).

28. F. Y. McWhorter, T. Wang, P. Nguyen, T. Chung, W. F. Liu, Modulation of macrophage phenotype by cell shape. *Proc Natl Acad Sci U S A* **110**, 17253-17258 (2013).
29. Z. Pincus, J. A. Theriot, Comparison of quantitative methods for cell-shape analysis. *J Microsc* **227**, 140-156 (2007).
30. Y. Sun *et al.*, Keratocyte fragments and cells utilize competing pathways to move in opposite directions in an electric field. *Curr Biol* **23**, 569-574 (2013).
31. M. H. V. B. R. Sonka, *Image processing, analysis and machine vision* (Chapman & Hall, London, 1993).
32. C. Sammut, Encyclopedia of Machine Learning. (2010).
33. P. J. Rousseeuw, Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *CAM</cja:jid> Journal of Computational and Applied Mathematics* **20**, 53-65 (1987).
34. J. D. Hunter, Matplotlib: A 2D Graphics Environment. *Computing in science & engineering.* **9**, 90-95 (2007).
35. M. Waskom, seaborn: statistical data visualization. *JOSS Journal of Open Source Software* **6**, 3021 (2021).
36. C. E. W. C. K. I. Rasmussen, *Gaussian processes for machine learning* (MIT Press, Cambridge, Mass, 2008).
37. F. V. G. G. A. M. V. T. B. G. O. B. M. P. P. W. R. D. V. Pedregosa, Scikit-learn: Machine Learning in Python. *Journal of machine learning research : JMLR.* **12**, 2825-2830 (2012).
38. C. V. V. Cortes, Support-vector networks. *Mach Learn Machine Learning* **20**, 273-297 (1995).
39. N. V. B. K. W. H. L. O. K. W. P. Chawla, SMOTE: Synthetic Minority Over-sampling Technique. *jair Journal of Artificial Intelligence Research* **16**, 321-357 (2002).
40. H. M. Rostam, P. M. Reynolds, M. R. Alexander, N. Gadegaard, A. M. Ghaemmaghami, Image based Machine Learning for identification of macrophage subsets. *Scientific reports* **7**, 3521 (2017).
41. V. Ulman *et al.*, An objective comparison of cell-tracking algorithms. *Nat Methods* **14**, 1141-1152 (2017).
42. P. Matula *et al.*, Cell Tracking Accuracy Measurement Based on Comparison of Acyclic Oriented Graphs. *PLoS One* **10**, e0144959 (2015).
43. M. Maska *et al.*, A benchmark for comparison of cell tracking algorithms. *Bioinformatics* **30**, 1609-1617 (2014).
44. O. Hilsenbeck *et al.*, Software tools for single-cell tracking and quantification of cellular and molecular properties. *Nat Biotechnol* **34**, 703-706 (2016).
45. J. Y. Tinevez *et al.*, TrackMate: An open and extensible platform for single-particle tracking. *Methods* **115**, 80-90 (2017).
46. X. Zhang, R. Goncalves, D. M. Mosser, The isolation and characterization of murine macrophages. *Curr Protoc Immunol* **Chapter 14**, Unit 14 11 (2008).
47. N. Otsu, A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man and Cybernetics* **9**, 62-66 (1979).
48. E. L. Barnhart, K. C. Lee, K. Keren, A. Mogilner, J. A. Theriot, An adhesion-dependent switch between mechanisms that determine motile cell shape. *PLoS Biol* **9**, e1001059 (2011).
49. K. E. Magnusson, J. Jalden, P. M. Gilbert, H. M. Blau, Global linking of cell tracks using the Viterbi algorithm. *IEEE Trans Med Imaging* **34**, 911-929 (2015).
50. K. E. G. Magnusson, Segmentation and tracking of cells and particles in time-lapse microscopy. (2016).
51. H. B. Mann, D. R. Whitney "On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other," The Annals of Mathematical Statistics, Ann. Math. Statist. 18(1), 50-60, (March, 1947).

Supplementary materials:

Fig. S1. M0, M1 and M2 macrophage images used in the analysis

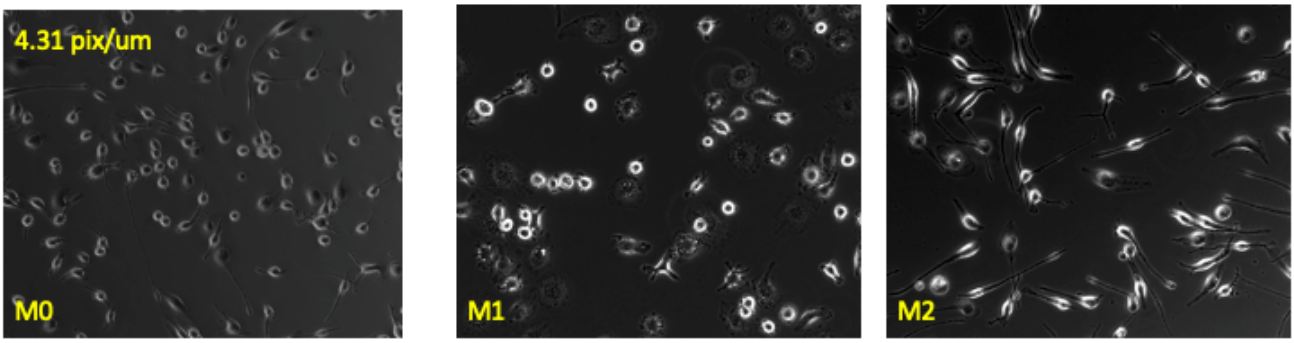
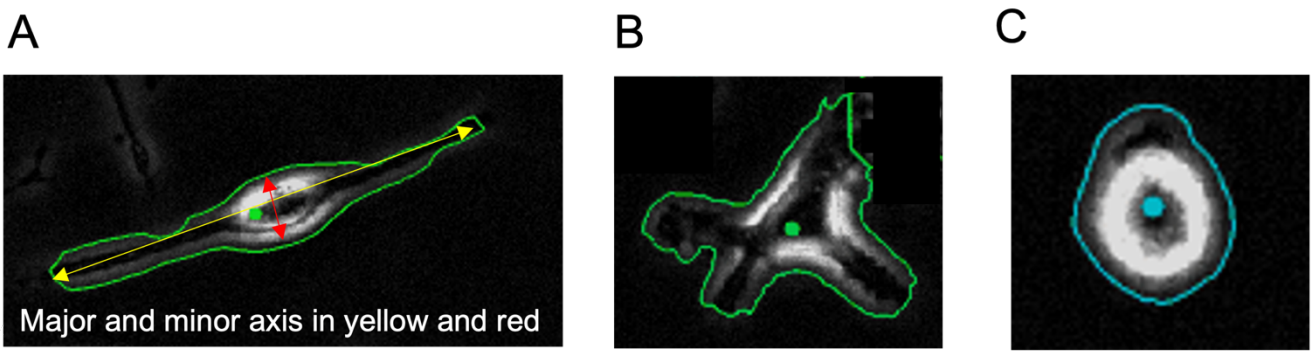


Fig. S2. Examples of elongated/ bipolar, multipolar, and circular macrophages with their shape parameter values



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

Fig. S3. 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.

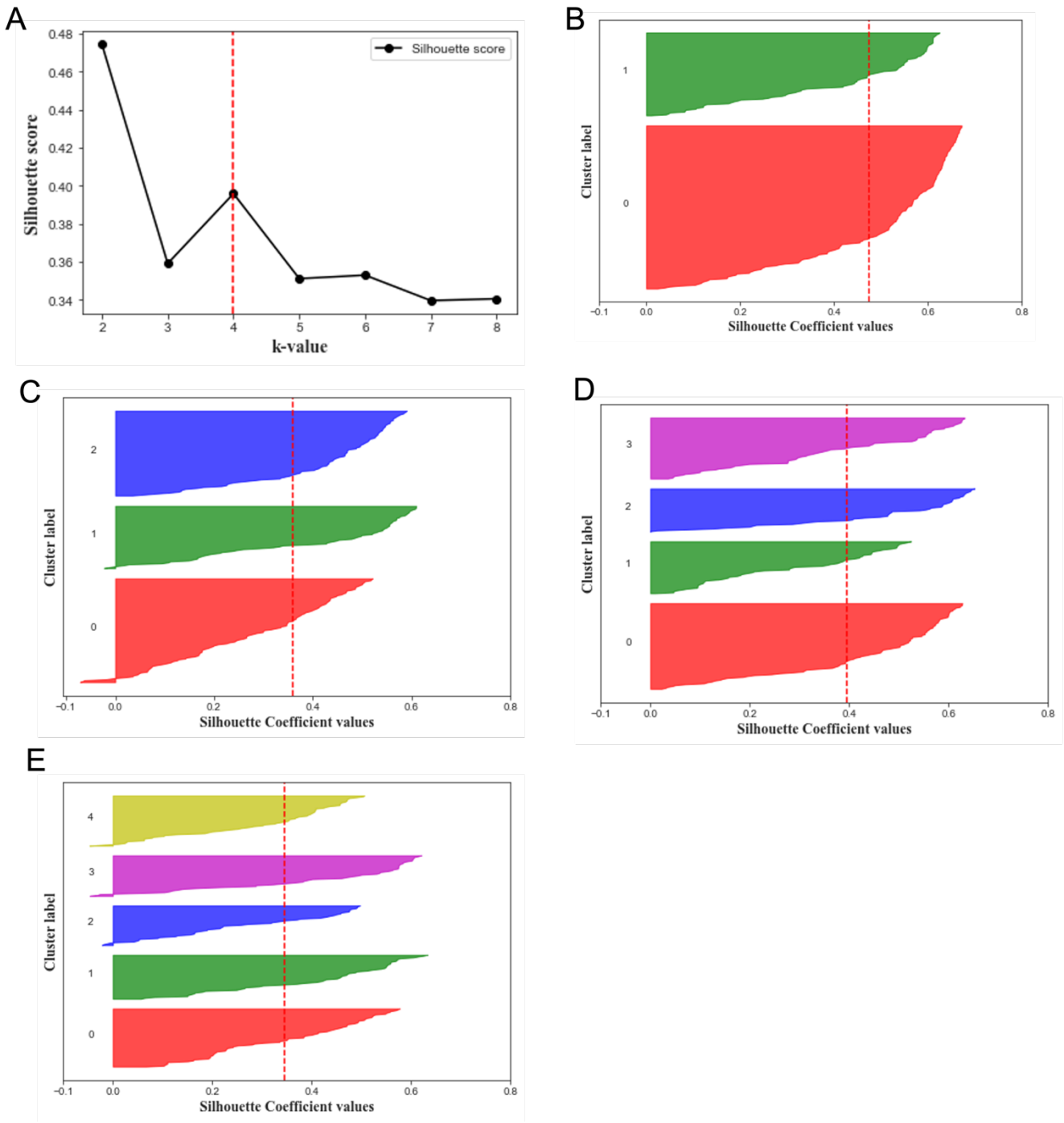
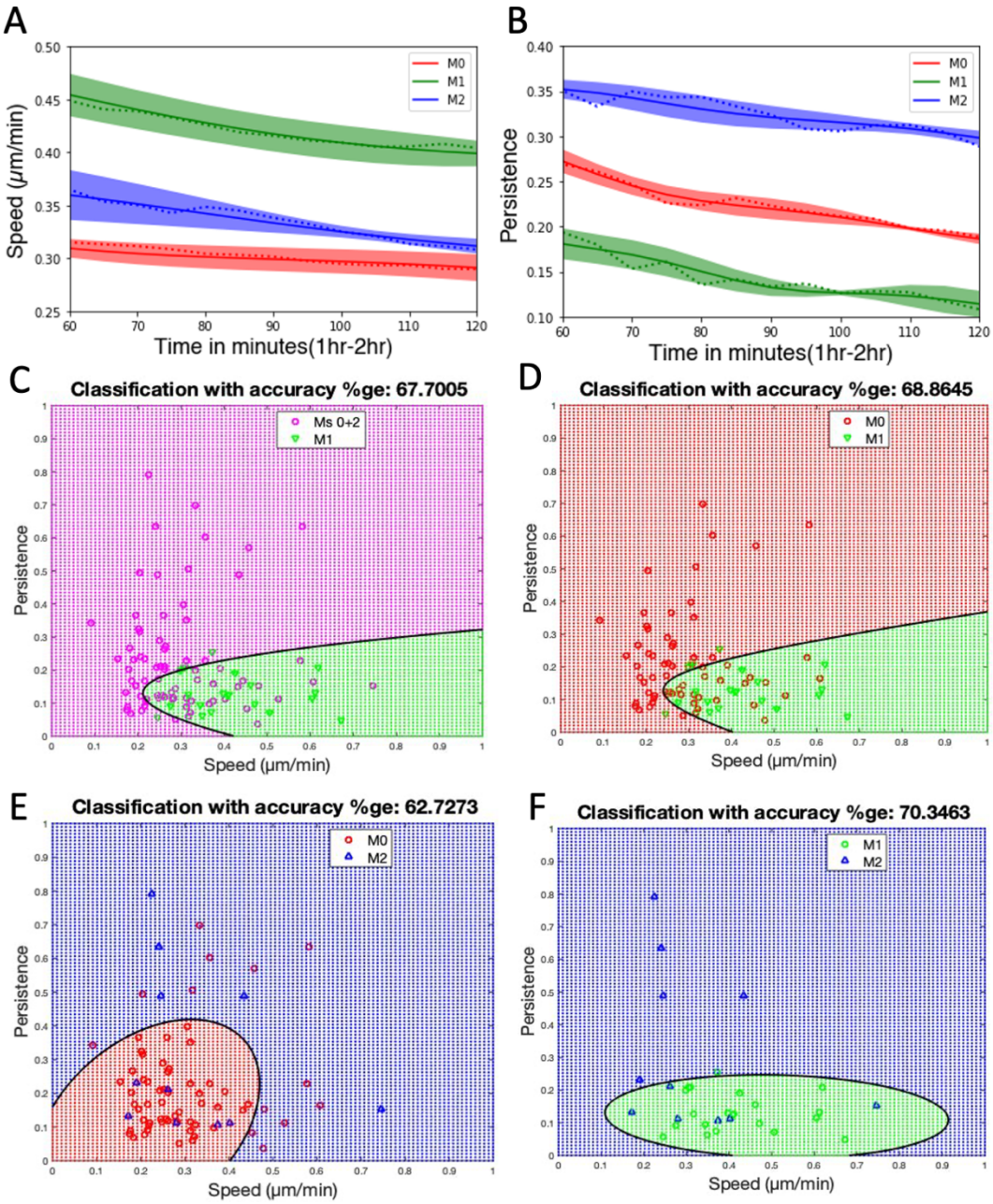


Fig. S4. Classification and discrimination of the macrophage subtypes using motility parameters alone



(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

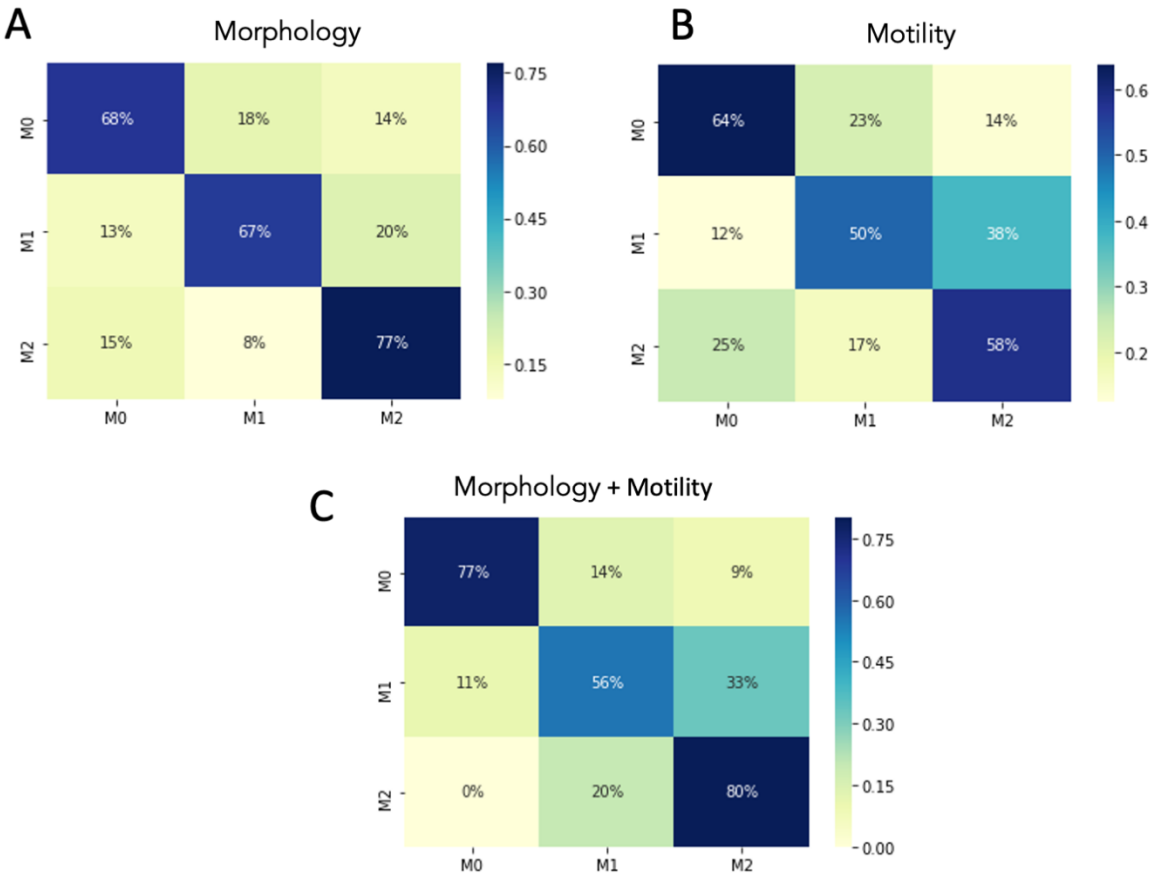
classification of the cells using Speed vs Persistence between M0 image and M2 image with a classification accuracy of around 63%. (*F*) Plot showing quadratic the classification of the cells using Speed vs Persistence between M1 image and M2 image with a classification accuracy of around 70%. The black line in *C*-*F* is the decision boundary between the classes.

We use the Gaussian process regressor model from the Scikit-learn python library (37) to plot the speed and persistence of the cells. We found that the speed of the cells in the M1 image is highest, followed by the speed of the cells in M2 and the lowest speed observed for the cells in M0 (**Fig. S4*A***). In contrast, the persistence of the cells in M2 is highest, followed by the persistence of the cells in M0 and the lowest persistence observed for the cells in M1 (**Fig. S4*B***). Based on this motility parameter analysis, we conclude that:

- M0 cells move slower and stay close to their initial time frame position.
- M1 cells move faster but still stay close to their initial time frame position.
- M2 cells move slower but reach farther away from their initial time frame position.

To better visualize and quantify the motility of the cells in M0, M1 and M2 images, we perform quadratic discriminant analysis on the median of the cells speed and persistence along its time series and plot them. This analysis provides a binary classification based on two input features (**Fig. S4**, showing the results for various pairs of classification labels). Classification accuracy percentages show that the cells in M1 image can be distinguished combined or independently from M0 and M2 cells using the speed and persistence motility parameters (**Figs. S4*C, D*, *F***). Notably, the classification accuracy percentage between M0 and M2 cells is comparatively lower around 63%, suggesting that it is comparatively difficult to classify M0 cells from M2 cells (**Fig. S4*E***).

Fig. S5. Stratified k-fold Cross Validation plots for the SVM model with 50% train and 50% test data



A) Input features with morphological parameters (Eccentricity, Compactness, Cell Size). (*C*) Input features with motility parameters (Trajectory Spread, Speed, Persistence). (*D*) Input features with both morphological and motility parameters.

Table S1. 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%**

Table S2: Target genes and primers used in q-PCR

Target gene	Forward primers (5'-3')	Reverse primers (5'-3')	Macrophage Subtypes
Arg-1	CGTTGTATGATGCACAGCCG	CCCCACCCAGTGATCTTGAC	M1
CD206	GTTCACCTGGAGTGATGGTTCTC	AGGACATGCCAGGGTCACCTTT	M2
CD31	ACAGAGCCAGCAGTATGA	AATGACAACCACCGCAAT	Angiogenesis
CXCL10	GATGGATGGACAGCAGAG	GGAAGATGGTGGTTAAGTTC	M1
Adgre1	TCTGGGGAGCTTACGATGGA	GAATCCCGCAATGATGGCAC	Macrophage
Fizz-1	CCTGCTGGGATGACTGCTACT	AGATCCACAGGCAAAGCCAC	M2
GAPDH	TCTCCTGCGACTTCAACA	TGTAGCCGTATTCATTGTCA	Housekeeping
IL-10	CAGAGCCACATGCTCCTAGA	TGTCCAGCTGGTCCTTTGTT	M2
IL-12p40	TGGTTTGCCATCGTTTTGCTG	ACAGGTGAGGTTCACTGTTTCT	M1
IL-1β	GTGCAAGTGTCTGAAGCAGC	CAAAGGTTTGGAAGCAGCCC	M1
IL-6	GGAGTCACAGAAGGAGTGGC	CGCACTAGGTTTGCCGAGTA	M2
iNOS	GTTCTCAGGCCAACAATACAAGA	GTGGACGGGTCGATGTCAC	M1
MCP-1	AGCCAACTCTCACTGAAGCC	GGACCCATTCCTTCTTGGGG	M1
MMP-9	CTGGACAGCCAGACACTAAAG	CTCGCGGCAAGTCTTCAGAG	M2 MMP
PPAR-γ	TCCTGTAAAAGCCCGGAGTAT	GCTCTGGTAGGGGCAGTGA	M2 Metabolic
Stat3	CTTGTCTACCTCTACCCCGACAT	GATCCATGTCAAACGTGAGCG	M1 TF
Stat6	TGAGGTGGGGACCAGCCGG	GTGACCAGGACACACAGCGG	M2 TF
TGF-β1	TGGAGCAACATGTGGAACTC	CAGCAGCCGGTTACCAAG	M2
TLR4	CAGAACAATAGAAGAGGAAGAC	GGCACTAACCACATAGAGAA	M1
TNF-α	GACGTGGAACTGGCAGAAGA	ACTGATGAGAGGGAGGCCAT	M1
Ym-1	CTCACTTCCACAGGAGCAGG	AGCTGCTCCATGGTCCTTC	M2
β-actin	GGCTGTATTCCCCTCCATCG	CCAGTTGGTAACAATGCCATGT	Housekeeping