

## Identification of Key Behavioral Features of Cells Based on Relational Network Analysis and Its Application in Deciphering Macrophage Calcium Signaling

Jiayang Liu

School of Biomedical Engineering, Hainan University, Sanya 572024, Hainan, China

Corresponding Email: 23210831000013@hainanu.edu.cn

**Abstract:** High-dimensional morphodynamic data acquired through intravital microscopy contain rich biological information; however, feature redundancy and noise interference substantially limit the precise dissection of cellular functional heterogeneity. This study introduces a systematic feature-engineering framework structured around “discriminative power evaluation – relational network construction – functional module screening” to automatically identify key behavioral descriptors from high-dimensional parameter sets. The approach first combines the likelihood-ratio test (LRT) and the adjusted Rand index (ARI) to quantify the discriminative efficacy of each parameter. It then integrates Pearson correlation and distance correlation to build a weighted parameter-association network that captures both linear and nonlinear dependencies. Finally, the Leiden community-detection algorithm is employed to identify co-varying functional modules, from which representative key features are selected based on dual criteria: topological centrality and discriminative power.

Applied to four independent intravital imaging datasets of hepatic parenchymal macrophages (Kupffer cells), the framework robustly identifies key morphodynamic features that are strongly associated with cellular calcium-signal intensity, while exhibiting high consistency and generalizability across experimental batches. This work not only elucidates intrinsic “morphology–function” coupling principles in macrophages, but also provides a reliable methodological tool for the unbiased quantification of immune-cell heterogeneity in complex microenvironments.

**Keywords:** Intravital microscopy; Feature engineering; Parameter-association network; Community detection; Kupffer cells; Calcium signaling

### 1. Introduction

The immune system is a highly dynamic and complex biological system composed of immune organs, immune cells, and immune molecules. It is responsible for recognizing and eliminating antigenic foreign substances, thereby maintaining internal homeostasis and physiological balance. Within this system, immune cells play a central role, and their behavior and morphology define their immunoregulatory properties, which are essential for immune responses [1-3]. Recent advances in single-cell analysis technologies, such as single-cell transcriptomics and proteomics, have revolutionized the interpretation of biological phenomena by providing spatially and temporally resolved cellular landscapes in both healthy and diseased tissues [4-6]. However, these approaches often fail to capture the dynamic scenarios in which cells continuously alter their biochemical properties and downstream “behavioral” outputs [7-9]. Growing evidence highlights the importance of behavioral morphology in macrophages [10, 11]. For example, a 2024 study in *Science Immunology* demonstrated that plasma membrane abundance and tension determine the phagocytic capacity of myeloid cells and influence their functional communication during immune responses, underscoring the critical role of biophysical properties in regulating immune-cell behavior [12, 13]. Similarly, articles published in the *Journal of Experimental Medicine* (IF=12.6, 2020) and the *International Journal of Surgery* (IF=12.5, 2023) have both indicated that macrophage morphology is associated with single-cell diversity and prognosis in colorectal liver metastasis [14, 15].

Calcium ions serve as vital intracellular second messengers and play a pivotal regulatory role in macrophage activation, polarization, and functional execution. Studies show that the spatiotemporal dynamics of calcium concentration directly affect macrophage phagocytic ability, inflammatory cytokine release, and migratory behavior. Subtle differences in calcium signal intensity can determine whether macrophages adopt a pro-inflammatory (M1) or anti-inflammatory (M2) phenotype. In the unique immune microenvironment of the liver, calcium signaling influences hepatic parenchymal macrophages (Kupffer cells) not only in mediating the recognition and clearance of pathogens and tumor cells, but also in governing immune surveillance of cancer liver metastasis and the balance between inflammatory injury and tissue repair [16-20].

Intravital microscopy enables the simultaneous acquisition of hundreds of multidimensional parameters that describe cell morphology, motility, and fluorescence characteristics [21, 22]. Nevertheless, the inherent redundancy, multicollinearity, and noise in such high-dimensional data severely limit the precise identification of essential features underlying cellular heterogeneity [23]. Traditional univariate statistical analyses often ignore complex coupling among parameters, making it difficult to uncover cooperative regulatory mechanisms behind cell behavior. To address these limitations, this chapter introduces a systematic feature-engineering framework based on “discriminative power assessment – relational network construction – functional module screening.” The framework first quantifies the discriminative efficacy of parameters through statistical testing. It then constructs a weighted parameter-association network by integrating both linear and nonlinear metrics. Finally, community-detection algorithms are employed to identify functionally co-varying modules and to select key representative parameters. The goal is to extract a compact subset of features with high biological interpretability and information content from high-dimensional heterogeneous data, which is subsequently validated using multiple datasets of hepatic parenchymal macrophages.

### 2. Materials and Methods

#### 2.1 Discriminative Power-Based Preliminary Parameter Screening Strategy

To eliminate ineffective noise from the vast number of raw parameters and identify features with significant discriminatory power, this study employs a dual-assessment strategy combining statistical hypothesis testing with cluster consistency metrics.

##### 2.1.1 Likelihood Ratio Test (LRT)

For datasets with known experimental conditions or cell type labels, a Likelihood Ratio Test (LRT) was introduced to assess the significance of distribution differences for each parameter across different groups. For the  $i$ -th parameter  $X_i$ , the null hypothesis  $H_0$  (equal means across groups) and the alternative hypothesis  $H_1$  (at least one group mean differs) were constructed. By comparing the likelihood function values of the restricted model (Null Model) and the full model (Full Model), the likelihood ratio statistic  $\Lambda_i$  was calculated:

$$\Lambda_i = -2 \ln \left( \frac{L(\hat{\theta}_0 | X_i)}{L(\hat{\theta}_1 | X_i)} \right)$$

where  $L(\cdot)$  denotes the likelihood function, and  $\hat{\theta}_1$  and  $\hat{\theta}_0$  are the maximum likelihood estimates under the alternative and null hypotheses, respectively.  $\Lambda_i$  asymptotically follows a chi-squared distribution. A smaller corresponding  $p$ -value indicates stronger discriminatory power of the parameter in distinguishing different biological states. Parameters with  $p < 0.05$  and an effect size above a predefined threshold were included in the candidate set.

##### 2.1.2 Adjusted Rand Index (ARI) Evaluation

To further verify the structure-preserving capability of parameters from an unsupervised perspective, the Adjusted Rand Index (ARI) was calculated for clustering results generated using each parameter individually. ARI measures the agreement between a clustering result based on a single parameter and the true labels (or a reference gold standard), correcting for chance:

$$ARI = \frac{RI - E[RI]}{\max(RI) - E[RI]}$$

where RI is the Rand Index and E[RI] is its expected value under random labeling. ARI ranges from [-1, 1], with values closer to 1 indicating richer class separation information contained within that parameter. Combining the p-value from LRT and the ARI score, a comprehensive discriminative power score  $S_{disc}$  was constructed. The top N high-information-content parameters were then selected based on this score for the subsequent network construction stage.

**2.2 Construction of a Multidimensionally Fused Parameter Correlation Network**

Cellular behavioral features often do not vary independently but are coupled through complex biophysical mechanisms. To comprehensively characterize these dependencies, this study moved beyond the assumption of purely linear correlation and constructed a weighted parameter correlation network that fuses linear and nonlinear metrics.

**2.2.1 Linear and Nonlinear Association Measures**

For the parameter set  $\{X_1, X_2, \dots, X_N\}$  after initial screening, pairwise association strengths were calculated as follows:

Linear Correlation: The Pearson Correlation Coefficient (PCC) was used to measure the degree of linear dependence between parameters:

$$r_{ij} = \frac{\sum_{k=1}^M (x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j)}{\sqrt{\sum_{k=1}^M (x_{ik} - \bar{x}_i)^2} \sqrt{\sum_{k=1}^M (x_{jk} - \bar{x}_j)^2}}$$

**Nonlinear Correlation:** The Distance Correlation (dCor) or a k-nearest neighbor-based Mutual Information (MI) estimator was introduced to capture potential monotonic or non-monotonic nonlinear relationships. Distance correlation has the desirable property of being zero if and only if the variables are independent, effectively identifying complex patterns that PCC may fail to detect, see Figure 1.

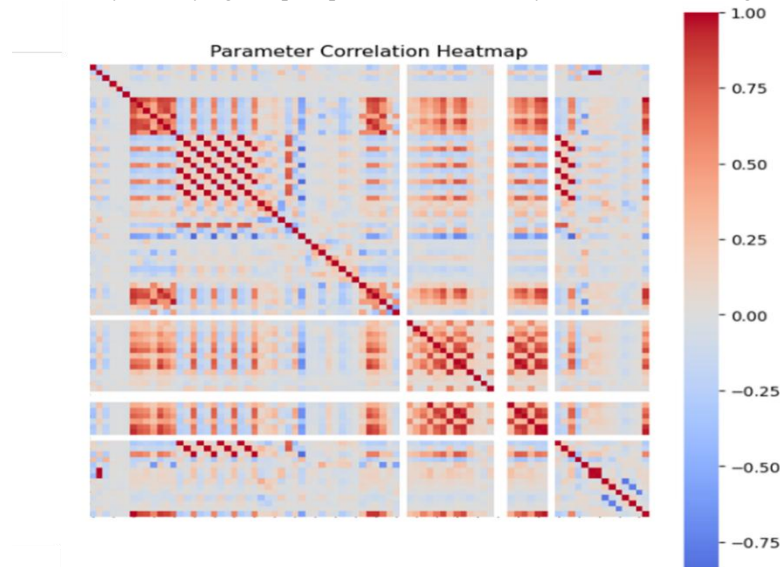


Figure 1 Distribution of Parameter Correlations.

**2.2.2 Integrated Association Strength and Network Topology Definition**

To construct a unified network model, linear and nonlinear metrics were fused into a composite association strength  $W_{ij}$ . A maximum fusion strategy (or a weighted average strategy; this study used the maximum strategy to preserve the strongest association signals) was employed:

$$W_{ij} = \max(|r_{ij}|, dCor_{ij})$$

An undirected weighted graph  $G = (V, E)$  was thus constructed, where the node set  $V$  represents the screened cellular behavioral parameters, the edge set  $E$  represents the associations between parameters, and the edge weight  $w_{ij} = W_{ij}$ . To eliminate weak noise connections, a threshold  $\tau$  was set, retaining only edges with  $w_{ij} > \tau$ , resulting in a sparsified parameter topology network. This network visually reveals the structure of coordinated regulation among cellular morphology, motility, and fluorescence parameters.

**2.3 Identification of Key Functional Modules Based on Community Detection**

Building upon the constructed parameter correlation network, this study applied community detection algorithms to identify clusters of parameters that are densely interconnected internally but sparsely connected externally—referred to as "functional modules." Each module represents a set of feature units that vary in a coordinated manner during biological processes (e.g., a "deformation-migration coupling module" or a "fluorescence-metabolism synchronization module").

**2.3.1 Module Decomposition Algorithm**

The Louvain and Leiden algorithms were compared and the Leiden algorithm was employed for community partitioning. Both algorithms optimize the objective function of modularity (Q):

$$Q = \frac{1}{2m} \sum_{i,j} \left[ w_{ij} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j)$$

where  $m$  is the total edge weight of the network,  $k_i$  is the strength of node  $i$ ,  $c_i$  is the community to which node  $i$  belongs, and  $\delta$  is the Kronecker delta function. Compared to the Louvain algorithm, the Leiden algorithm, by introducing a local refinement step, ensures community connectedness and avoids arbitrarily assigned partitions, thereby enabling more stable identification of functional submodules with clear biological significance, see Figure 2.

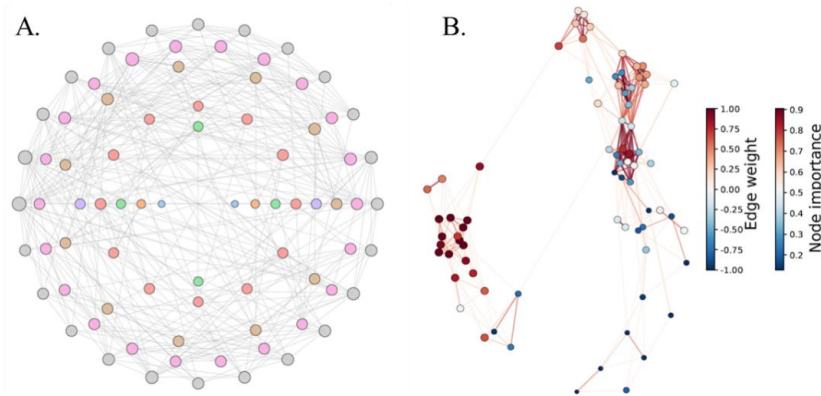


Figure 2 Network Construction, Visualization, and Key Module Identification.

### 2.3.2 Key Representative Parameter Screening

After identifying Kfunctional modules, to avoid information redundancy and reduce the dimensionality for downstream analysis, one most representative key parameter was selected from each module. The selection criteria included:

1. Topological Centrality: Calculating node degree centrality or betweenness centrality within the module, selecting the parameter with the most extensive connections and located at the core of the module.
2. Discriminative Power Priority: Incorporating the  $S_{discscore}$  calculated in Section 3.2, and prioritizing the parameter with the strongest discriminative power when topological centrality is similar.

Finally, from the Ninitially screened parameters, Kkey representative parameters ( $K \ll N$ ) were distilled to form a streamlined feature vector for downstream analyses of cellular heterogeneity, state evolution reconstruction, and tracking prediction. This method not only achieves data dimensionality reduction and denoising but also reveals the intrinsic functional synergy among cellular behavioral parameters from a network topology perspective.

### 2.4 Data Source

Following the loss of corneal and pinch reflexes, mice were placed on a temperature-controlled heating pad (RWD Life Science, China) to maintain core body temperature. Imaging was performed using an LSM 780 NLO multiphoton/confocal microscope (Zeiss, Germany) equipped with 10× (NA 0.45) or 20× (NA 1.0) water immersion objectives. The left liver lobe was exposed via a liver window for in vivoobservation. To track Kupffer cell dynamics, time-lapse imaging was initiated after modeling, acquiring z-stacks with a 10  $\mu\text{m}$  step size over a 100  $\mu\text{m}$  depth at a rate of 3 frames/minute. To simultaneously distinguish intrahepatic macrophage subsets, a chimeric mouse model was established using Cx3cr1GFP/+Lyz2RFP/+bone marrow reconstitution. In this model, GFP-labeled monocyte-derived macrophages and RFP-labeled neutrophils were recorded with a higher temporal resolution of 6 frames/minute (z-step 10  $\mu\text{m}$ ). This imaging protocol was repeated for four biologically independent experiments, yielding approximately 400 frames in total. Subsequent computational analysis identified a large number of cell masks, and after tracking and reconstruction, hundreds of single-cell migration trajectories were obtained.

### 3 Results

To extract truly informative features from the high-dimensional behavioral data of cells acquired through intravital imaging, we developed a module for automated parameter selection. The module operates in two distinct modes according to data availability: supervised (labeled) and unsupervised (unlabeled). For labeled data, the Adjusted Rand Index (ARI) and Likelihood Ratio Test (LRT) were employed to evaluate parameter importance—ARI for fully labeled data and LRT for scenarios with partially missing labels. For unlabeled data, adaptive clustering was used to compute separation scores for parameter screening. By integrating parameter-importance metrics with inter-parameter correlation scores, a weighted parameter-association network was constructed, and community-detection algorithms were subsequently applied to select representative parameters.

The framework was then partially implemented on intravital microscopy datasets of macrophages residing in the hepatic parenchyma. For the first dataset, the identified parameter set was subjected to further cluster analysis. The silhouette coefficient method determined the optimal cluster number to be 2, while the elbow method suggested an optimal range of 2–4 clusters; results for both cluster-number settings are presented in Figure 3.

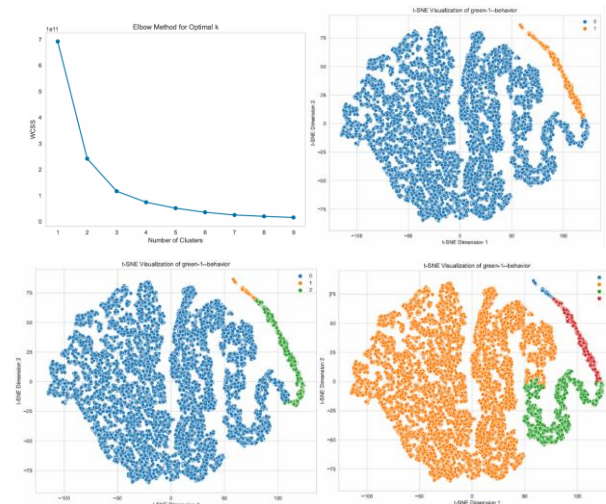


Figure 3 Dimensionality Reduction and Clustering of Macrophage Multiparameter Data.

The distribution of calcium ion concentration is illustrated in the Figure 4. Variations and a gradient relationship are discernible across different clusters, indicating a potential association between calcium concentration and cellular behavior. It is also noteworthy that, in the two-cluster setting, the two groups exhibited statistically significant differences in fluorescence intensity corresponding to calcium concentration.

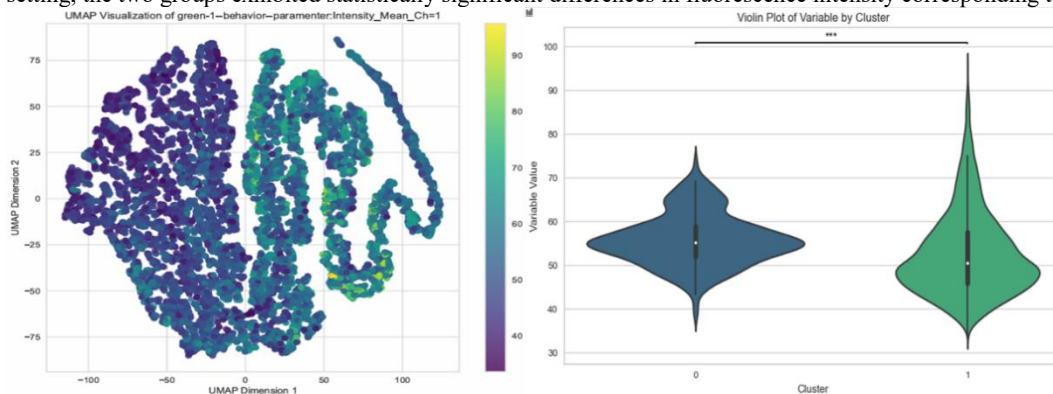


Figure 4 Calcium Concentration Differentiation Following Dimensionality Reduction and Clustering of Macrophage Multiparameter Data.

For the second dataset, the silhouette coefficient again indicated that the optimal number of clusters was 2, while the elbow method likewise suggested an ideal range of 2–4 clusters. Similarly, the calcium concentration displayed an uneven distribution across different clusters in the dimensionality-reduced visualization, see Figure 5.

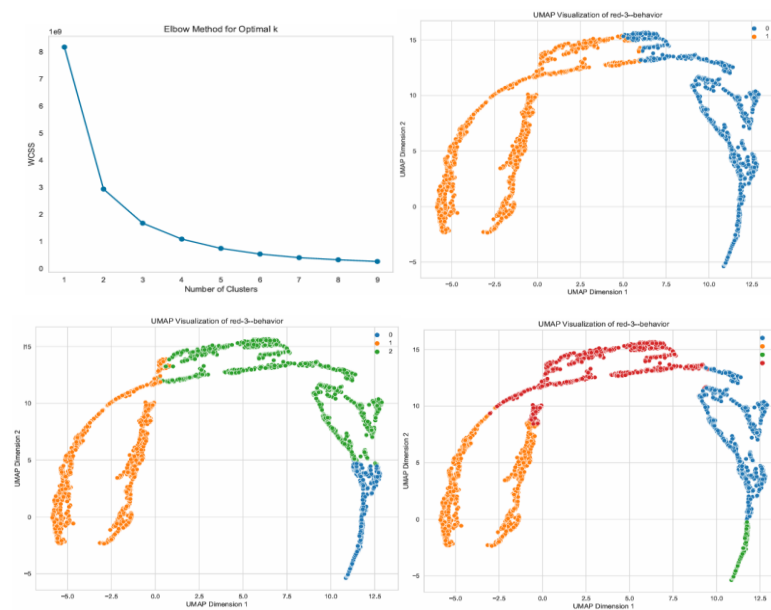


Figure 5 Dimensionality Reduction and Clustering of Macrophage Multiparameter Data in the Validation Set.

Moreover, even when grouped into four clusters, significant differences in calcium ion concentration were still observed among the clusters. The new dataset produced similar findings, reinforcing the potential relationship between calcium concentration and macrophage behavioral dynamics, see Figure 6.

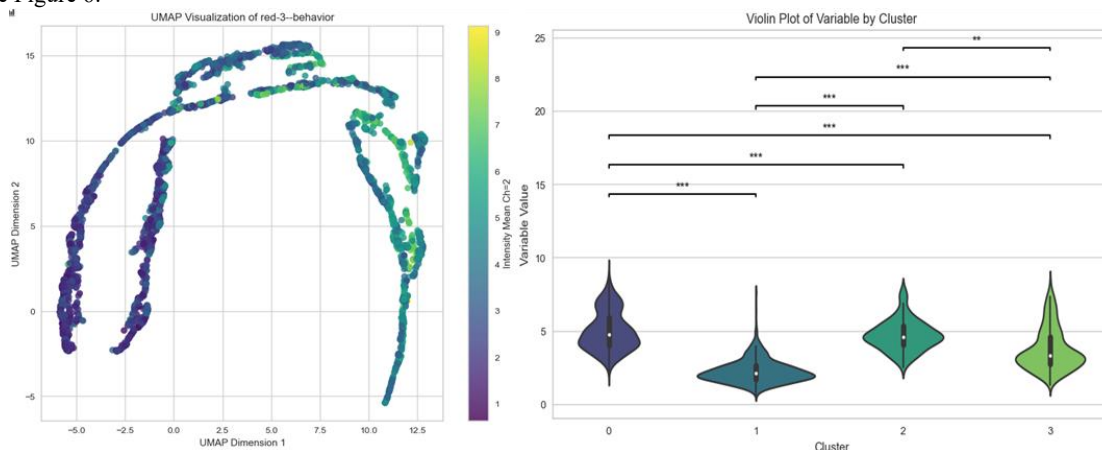


Figure 6 Calcium Concentration Differentiation Following Dimensionality Reduction and Clustering in the Validation Set.

To further elucidate the relationship between calcium concentration and macrophage behavior, a third dataset was subsequently analyzed. Cells were initially grouped according to their calcium concentration to assess potential differences in behavior and morphology among clusters. Based on the silhouette coefficient calculated after UMAP dimensionality reduction, the data could be optimally partitioned into two distinct clusters, see Figure 7.

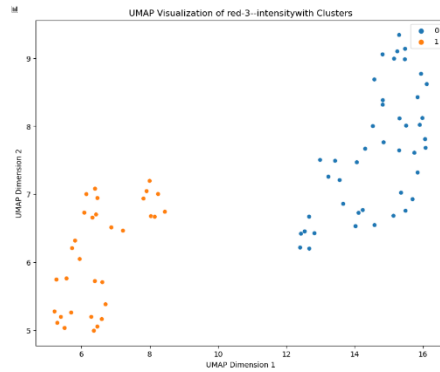


Figure 7 Dimensionality Reduction and Clustering in the Validation Set of Macrophages.

Compared to Cluster 1, cells in Cluster 0 showed a significantly higher mean calcium concentration, yet a significantly lower mean of the maximum calcium concentration per time point and a lower standard deviation in calcium concentration. This pattern suggests that hepatic parenchymal macrophages in Cluster 0 likely exist in a stable, high-calcium state, whereas those in Cluster 1 exhibit an unstable, low-calcium concentration profile, see Figure 8.



Figure 8 Violin Plots of Mean Calcium Concentration and Standard Deviation for Macrophages in the Validation Set.

Furthermore, the two clusters exhibited distinct patterns in motility, morphology, and spatial distribution. Specifically, cells in Cluster 0 displayed larger cell volume, higher ellipticity, a flatter morphology, significantly lower velocity, and closer proximity to the imaging center. These observations align with previously reported classification trends based on behavioral and morphological parameters, see Figure 9.

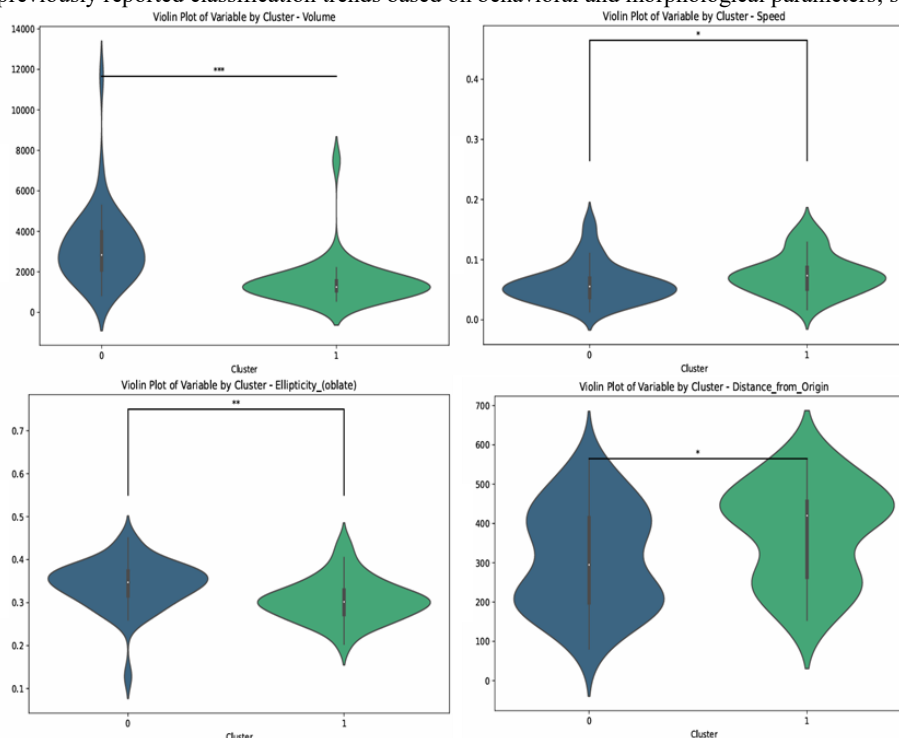


Figure 9 Violin Plots of Selected Behavioral and Morphological Parameters for Macrophages in the Validation Set.

Further analysis of the relationships between calcium concentration and four key parameters—cell volume, velocity variance, and distance from the imaging center—confirmed discernible correlative trends, see Figure 10.

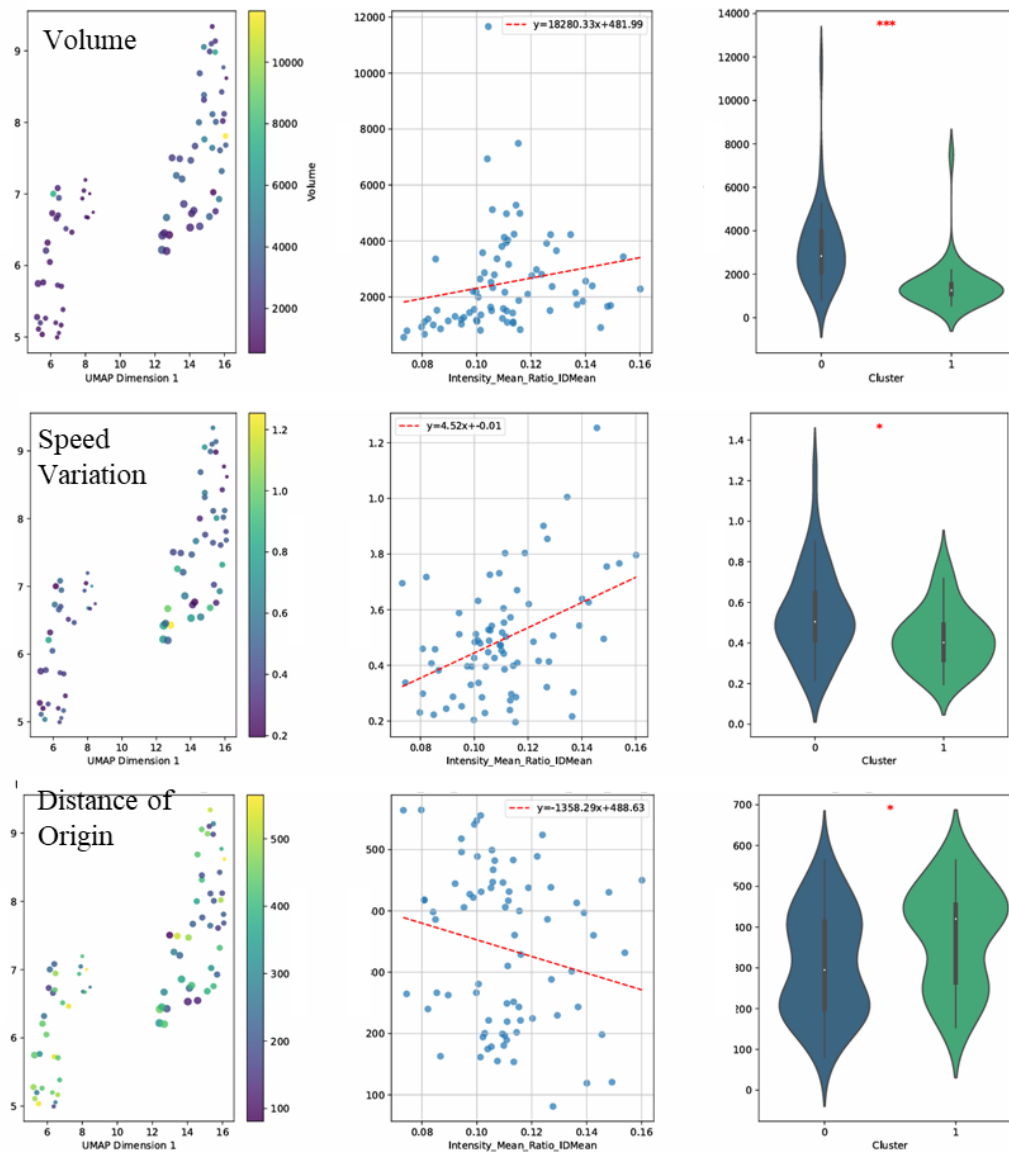


Figure 10 Relationship Between Calcium Concentration and Behavioral/Morphological Parameters for Macrophages in the Validation Set.

Analysis of the fourth dataset produced consistent classification results. Comparison between the two classifications further revealed a high degree of overlap, see Figure 11 and 12.

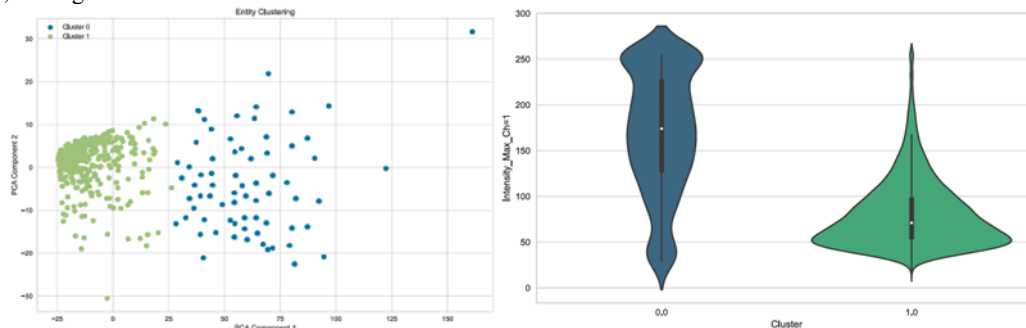


Figure 11 Classification Results and Calcium Concentration Violin Plots for the Second Validation Set of Macrophages.

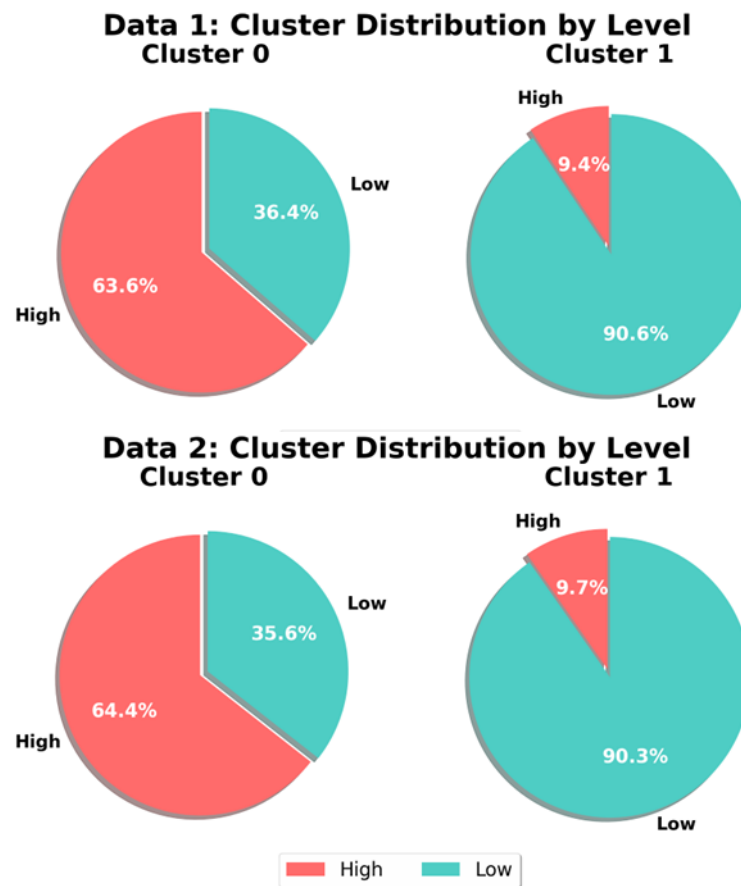


Figure 12 Pie Charts Comparing Classification Outcomes Across the Two Validation Sets.

It should be noted that in previous classifications, the group with higher mean calcium concentration (Group 0) displayed slightly lower velocity but significantly greater trajectory displacement. In the present classification, the high-calcium group (Group 0) exhibited significantly lower velocity and relatively normal trajectory displacement. While the specific numerical patterns differ, both observations support the conclusion that cells in the high-calcium concentration group (Group 0) demonstrate more directed motility. Overall, based on the three datasets analyzed, hepatic parenchymal macrophages with relatively higher calcium concentrations consistently show larger and flatter cell bodies, lower velocity, higher directional persistence in movement, and greater spatial clustering.

#### 4 Discussion

The adaptive feature-screening framework developed in this study innovatively combines supervised discriminability assessment (ARI/LRT) with unsupervised cluster-based separation analysis, and further integrates parameter-association networks with community-detection algorithms. This integrated approach effectively reduces redundancy in high-dimensional intravital imaging data and provides a novel perspective for investigating the dynamic heterogeneity of tissue-resident immune cells. Applied to four independent datasets of hepatic parenchymal macrophages, the analysis reveals a significant coupling between cellular morphodynamics and intracellular calcium signaling: the high-calcium, stable subpopulation shows larger cell volume, flatter morphology, and low-speed yet highly directional movement—characteristics indicative of an ordered “surveillance” state. In contrast, the low-calcium, fluctuating subpopulation exhibits a random exploratory pattern. Despite batch effects across datasets, the core pattern linking high calcium concentration with high directional persistence and flattened morphology remains highly consistent, strongly supporting the central role of calcium signaling in cytoskeletal remodeling and motility-strategy determination. This work not only introduces a new quantitative dimension for dissecting the functional heterogeneity of tissue-resident immune cells, but also establishes a solid methodological and theoretical foundation for deeper insights into the regulation of cellular behavior in liver immune homeostasis and inflammatory responses.

#### Data Sharing Agreement

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

#### Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

#### Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

## References

- [1] E Lelkes, MB Headley, EE Thornton, et al. The spatiotemporal cellular dynamics of lung immunity. *Trends Immunol.* 2014.35:379-386
- [2] G Crainiciuc, M Palomino-Segura, M Molina-Moreno, et al. Behavioural immune landscapes of inflammation. *Nature.* 2022.601:415-421
- [3] J Copperman, IC McLean, SM Gross, et al. Single-cell morphodynamical trajectories enable prediction of gene expression accompanying cell state change. *bioRxiv.* 2024.
- [4] I Kwok, E Becht, Y Xia, et al. Combinatorial Single-Cell Analyses of Granulocyte-Monocyte Progenitor Heterogeneity Reveals an Early Unipotent Neutrophil Progenitor. *Immunity.* 2020.53:303-318 e305
- [5] I Amit, A Regev and N Hacohen. Strategies to discover regulatory circuits of the mammalian immune system. *Nat Rev Immunol.* 2011.11:873-880
- [6] Y Zhai, L Chen and M Deng. scBOL: a universal cell type identification framework for single-cell and spatial transcriptomics data. *Brief Bioinform.* 2024.25:
- [7] T Lämmermann, PV Afonso, BR Angermann, et al. Neutrophil swarms require LTB4 and integrins at sites of cell death in vivo. *Nature.* 2013.498:371-375
- [8] V Sreeramkumar, JM Adrover, I Ballesteros, et al. Neutrophils scan for activated platelets to initiate inflammation. *Science.* 2014.346:1234-1238
- [9] A Woodfin, MB Voisin, M Beyrau, et al. The junctional adhesion molecule JAM-C regulates polarized transendothelial migration of neutrophils in vivo. *Nat Immunol.* 2011.12:761-769
- [10] M Selig, L Poehlman, NC Lang, et al. Prediction of six macrophage phenotypes and their IL-10 content based on single-cell morphology using artificial intelligence. *Front Immunol.* 2023.14:1336393
- [11] T Hourani, A Perez-Gonzalez, K Khoshmanesh, et al. Label-free macrophage phenotype classification using machine learning methods. *Sci Rep.* 2023.13:5202
- [12] Z Alraies, CA Rivera, MG Delgado, et al. Cell shape sensing licenses dendritic cells for homeostatic migration to lymph nodes. *Nat Immunol.* 2024.25:1193-1206
- [13] BY Winer, AH Settle, AM Yakimov, et al. Plasma membrane abundance dictates phagocytic capacity and functional cross-talk in myeloid cells. *Sci Immunol.* 2024.9:eadi2388
- [14] M Donadon, G Torzilli, N Cortese, et al. Macrophage morphology correlates with single-cell diversity and prognosis in colorectal liver metastasis. *J Exp Med.* 2020.217:
- [15] G Costa, C Sposito, C Soldani, et al. Macrophage morphology and distribution are strong predictors of prognosis in resected colorectal liver metastases: results from an external retrospective observational study. *Int J Surg.* 2023.109:1311-1317
- [16] Pogonyalova MY, Popov DY, Vinokurov AY. Intracellular Calcium as a Regulator of Polarization and Target Reprogramming of Macrophages. *Int J Mol Sci.* 2025;26(24):11901. Published 2025 Dec 10. doi:10.3390/ijms262411901
- [17] Desai BN, Leitinger N. Purinergic and calcium signaling in macrophage function and plasticity. *Front Immunol.* 2014 Nov 27;5:580. doi: 10.3389/fimmu.2014.00580. PMID: 25505897; PMCID: PMC4245916.
- [18] Suzuki Y, Katayama T, Fujita Y, Koide T, Sawai Y, Maeda K, Kondo R, Giles WR, Imaizumi Y, Yamamura H. Kir2.1 channels drive macrophage migration through enhancing store-operated Ca<sup>2+</sup> entry. *Am J Physiol Cell Physiol.* 2025 Aug 1;329(2):C413-C425. doi: 10.1152/ajpcell.00901.2024. Epub 2025 Jul 1. PMID: 40591253.
- [19] Schlautmann L, Burgdorf D, Ghosh S, Schieren A, Klümpen L, Stötzel I, Bremser J, Döngi M, Mass E, Stein V, Quast T, Kolanus W, Lang T, Kiermaier E, Simon MC, Burgdorf S. Glucose tasting depletes intracellular calcium stores and impairs macrophage functionality. *iScience.* 2025 Oct 14;28(11):113770. doi: 10.1016/j.isci.2025.113770. PMID: 41244562; PMCID: PMC12616019.
- [20] Sun Q, Shen M, Zhu S, Liao Y, Zhang D, Sun J, Guo Z, Wu L, Xiao L, Liu L. Targeting NAD<sup>+</sup> metabolism of hepatocellular carcinoma cells by lenvatinib promotes M2 macrophages reverse polarization, suppressing the HCC progression. *Hepatol Int.* 2023 Dec;17(6):1444-1460. doi: 10.1007/s12072-023-10544-7. Epub 2023 May 19. PMID: 37204655.
- [21] MH Qureshi, N Ozlu and H Bayraktar. Adaptive tracking algorithm for trajectory analysis of cells and layer-by-layer assessment of motility dynamics. *Comput Biol Med.* 2022.150:106193
- [22] A Arbelle, J Reyes, JY Chen, et al. A probabilistic approach to joint cell tracking and segmentation in high-throughput microscopy videos. *Med Image Anal.* 2018.47:140-152
- [23] JB Beltman, AF Marée and RJ de Boer. Analysing immune cell migration. *Nat Rev Immunol.* 2009.9:789-798