Quantifying Biological Processes in Motion and Across Scales

Diego Ulisse Pizzagalli, Raffaella Fiamma Cabini, and Inge M. N. Wortel

Abstract

A fundamental characteristic of living organisms is their capacity to change over time. These changes manifest across various scales: from molecular dynamics to modifications in the shape or arrangement of cells and, ultimately, changes at the level of tissues and organs. Imaging techniques to record biomedical videos have become integral to studying these processes in motion, but are notoriously difficult to analyze. While exciting advances in computer vision technology hold the promise to solve open challenges in quantifying these videos, their application to the field of biomedical imaging is far from trivial. This book chapter summarizes the state-of-the-art in analyzing motion from

The original version of this chapter was previously published without open access. A correction to this chapter is available at https://doi.org/10.1007/978-3-031-99410-4_31

D. U. Pizzagalli (✉)
Euler Institute, USI, Lugano, Switzerland

International Center for Advanced Computing in Medicine, University of Pavia, Pavia, Italy

Theodor Kocker Institute, University of Bern, Bern, Switzerland
e-mail: pizzad@usi.ch

R. F. Cabini
Euler Institute, USI, Lugano, Switzerland

International Center for Advanced Computing in Medicine, University of Pavia, Pavia, Italy
e-mail: raffaella.fiamma.cabini@usi.ch

Inge M. N. Wortel (✉)
Data Science, Institute for Computing and Information Sciences, Radboud University, Nijmegen, Netherlands
e-mail: inge.wortel@ru.nl

I. F. Cengiz et al. (eds.), *Bioimaging Modalities in Bioengineering*,
Biomaterials, Bioengineering and Sustainability 9,
https://doi.org/10.1007/978-3-031-99410-4_29

biomedical videos, and highlights opportunities and challenges for the translation of computer vision techniques to this field.

Keywords

Movement analysis · Tracking · Motility metrics · Video analysis

1 Introduction

Historically, our understanding of the dynamics of biological processes has gone hand in hand with the development of the microscopes used to visualize them—from Antoni van Leeuwenhoek's observations of 'wee animalcules' in the seventeenth century and Metchnikoff's observations of phagocytosis and theories on cell-mediated immunity in the late 1800s (Masters 2008), to the recordings of immune cells in lymph nodes facilitated by the first two-photon microscopes in the early 2000s. Likewise, recent advances in computer vision are revolutionizing our ability to *analyze* the motion in these videos. Video analysis enables the precise quantification of dynamic processes such as cell migration and interaction, intracellular transport, tissue, and organ remodeling—but very similar techniques can be applied to analyze human behavior, movement disorders, and crowd dynamics.

This chapter outlines classical and more recent methodologies for quantifying motion across different resolutions. While the primary focus is on cells observed through time-lapse microscopy, many of the principles discussed can also be applied to other imaging systems recording motile objects at various scales—ranging from molecular diffusion and vesicle trafficking to organ-level motility and blood flow. We discuss what is needed to successfully translate advances in computer vision to the field of bioimaging.

2 Track-based Quantification of Motion

Classical approaches to quantify motion start with **tracking**, the process of reconstructing the position of an object over time in a given space (i.e., one-dimensional (1D) along one given direction, 2D movement in a plane, or 3D movement in space). They are largely applied to time-lapse microscopy to track individual cells (Fig. 1a), but can also be applied to study the motility of molecules or intracellular vesicles. This section illustrates how to obtain tracks and how to describe them using quantitative metrics.

2.1 Tracking Algorithms

The first step in obtaining tracks from time-lapse videos is to identify objects, typically via **segmentation** or centroid **detection**, followed by linking positions of each individual cell across frames (tracking, Fig. 1a). Facilitated by software tools such as the Trackmate plugin in FIJI (Tinevez et al. 2017), manual or semi-

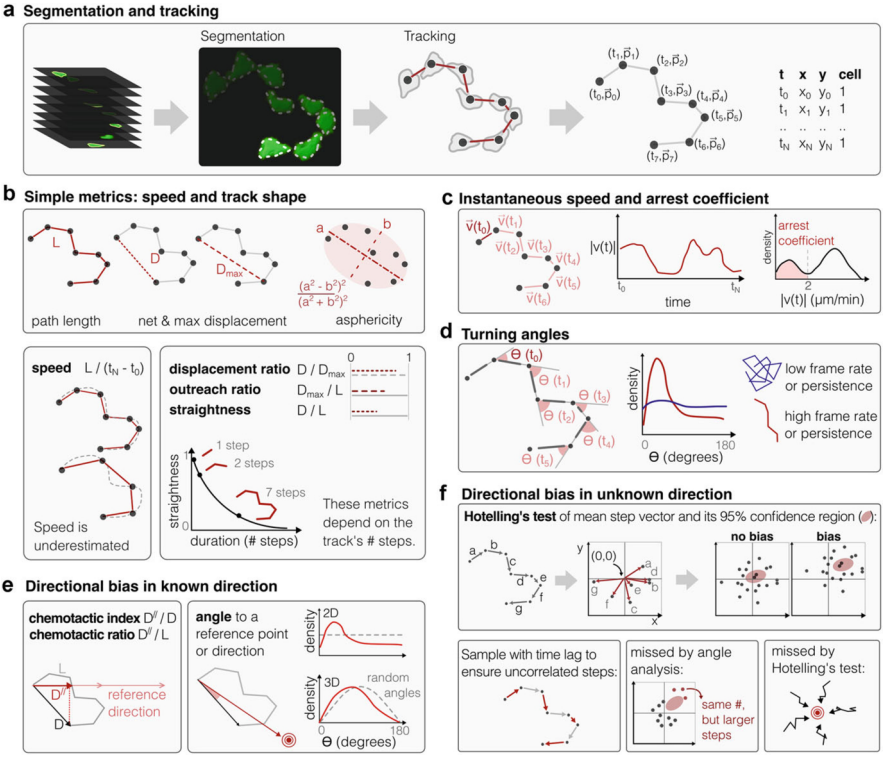

Fig. 1 Extracting track-based motility metrics. (**a**) Segmentation and tracking yields a table of cell positions over time. (**b**) Simple metrics describing tracks. (**c**) Deriving the arrest coefficient from instantaneous speeds. (**d**) Lower turning angles can indicate higher persistence or higher imaging frame rates. (**e**) Detecting bias in a known direction using metrics or angle analysis. Expected random angle distributions differ between 2D and 3D. (**f**) Hotelling's test can detect bias in an unknown direction by evaluating if the 95% confidence region of the mean step vector includes **0**. Since it assumes that steps are independent, a time lag is required for persistent cells. Compared to angle analysis, Hotelling's test is more sensitive to orthotaxis (where steps in the reference direction are faster, but not more frequent), but it can miss symmetrical biases with a mean step vector of **0**

automated tracking has long been the standard and still remains helpful in complex scenarios.

Advances in computer vision are now shifting the focus towards automated tracking methods (Hirsch et al. 2022). Despite significant algorithmic advancements, automated tracking of motile objects in biomedical imaging remains challenging due to issues like large deformations, unclear boundaries, cell interactions, and a limited field of view. Current state-of-the-art methods for automated tracking can be divided into two main approaches: **detect-and-link** and **space-time-connected**, discussed further below.

2.1.1 Detect-and-link Tracking

Detect-and-link is the most commonly used approach for cell tracking. It decomposes the problem into two steps that can be tuned independently: detecting cells in individual frames and subsequently linking them over time.

Cell Detection

Cell **detection** is the problem of finding cell centroids. This process often includes segmentation of the cell shape, which provides additional morphodynamic features but is not strictly necessary for cell tracking. In images where the (fluorescence) intensity is roughly uniform within the cell and across the field of view, a global intensity threshold can segment cells. In practice, however, the threshold often requires local adaptation, for example using background subtraction techniques. In transmitted light microscopy, higher variations of intensity are expected on the cell boundaries, and intensity thresholding may lead to hollow objects that are difficult to track. An effective technique to facilitate cell detection is **brushing**, where users annotate the foreground (cells) and background by drawing lines on the image (Berg et al. 2019; Arganda et al. 2017; Pizzagalli et al. 2022a). Points on these lines are then used to train pixel classifiers.

Once foreground pixels are detected, they are grouped into objects based on intensity (Li et al. 2010), texture features (Al et al. 2006), or size filtering using clustering, connected components, or region growing/merging methods. Challenges arise with non-globular shapes and with overlapping cells, requiring specific rules or advanced clustering algorithms (Pizzagalli et al. 2019a). When the expected size of the objects is known, methods based on intensity gradients such as the Laplacian-of-Gaussian or shape-based models such as Hough Transform, can effectively differentiate even overlapping objects.

More recently, state-of-the-art deep learning (DL) based methods for cell segmentation such as Cellpose (Stringer et al. 2021) and StarDist (Schmidt et al. 2018) have been developed to segment cells of diverse shapes and sizes, offering pre-trained and customizable models for integration into downstream tools like TrackMate. General-purpose models like YOLO have significantly improved multiple object detection and subsequent tracking in other fields, e.g., vehicles or soccer players (Markappa et al. 2024; Zuraimi et al. 2021; Krishna et al. 2021). Promising results are also expected from general-purpose models such as SAM (segment anything model), which can distinct objects within the same image. These have been recently extended to solve the Segment and Tracking Anything problem (Cheng et al. 2023; Rajič et al. 2023). For an overview of these and other cell segmentation methods, see (Schienstock et al. 2022; Moen et al. 2019).

Cell Linking

Once centroids are detected for each frame, **linking** involves finding the most likely correspondence of objects across frames. Simple heuristics, like maximal overlap or nearest neighbor, can be sufficient in simple settings, but often fail when objects are

crowded and/or displace far between frames, such that the object with the minimal displacement is no longer always the correct choice. Graph-based methods address this issue by modeling tracks as paths on a graph, where nodes represent detected objects and edges indicate possible transitions. Tracking is achieved by numerical optimization methods that minimize path-cost functions either for individual tracks or globally across all tracks (Jiang et al. 2017; Turetken et al. 2016; Magnusson et al. 2012; Chenouard et al. 2013). A key problem of detect-and-link tracking is that linking algorithms are sensitive to errors in the detection step, causing interrupted tracks or erroneous links (Berclaz et al. 2011). Graph-based approaches mitigate this issue by incorporating special nodes for object (dis)appearance. Alternatively, detection and linking can be performed simultaneously, discussed below.

2.1.2 Space-time Connected Tracking

Instead of treating detection and linking as separate problems, **space-time connected** approaches perform both simultaneously by reconstructing tracks as structures in space-time, essentially segmenting cells in both space and time at once. Contour evolution methods fit an initial contour (or 3D mesh) to object shapes by minimizing an energy functional. Usually, a data term and a regularization term are included to drive the contour towards desired properties, enabling the reconstruction of object shapes even without clear edges (Chan et al. 2001). An advantage of space-time approaches is the ability to identify special structures; for example, so-called "T-junctions" were shown to help recover the position of two interacting objects—one of the most challenging problems in tracking (Apostoloff et al. 2005, 2006). However, joint segmentation and tracking is challenging, and most methods focus on single objects or use contours from one frame as a starting points for the next. DL methods, using convolutional or recurrent neural networks, enable precise segmentation of 3D structures and show promise for reconstructing multiple space-time tracks (Milan et al. 2017).

2.2 Track-based Motility Descriptors

Once obtained, tracks can be described quantitatively using a simple yet rich set of metrics.

2.2.1 Speed and Stopping

The most commonly used descriptor of a track is its **speed**, defined as the total distance traveled divided by track duration (Fig. 1b). While seemingly straightforward, this definition will always underestimate the true speed as cells are observed at discrete intervals (Beltman et al. 2009b). This bias can be problematic when cells turn relatively far between frames (Fig. 1b); thus, reliable estimates of cell speed may require an imaging frame rate tuned to the cell's turning behavior.

Aside from average speed, "instantaneous" speeds between subsequent frames can vary by cell type and condition (Fig. 1c). For example, early in vivo microscopy studies observed immune cells in lymph nodes alternating between rapid motion (up

to 25 μm/min) and a non-motile "stop" state (Miller et al. 2002, 2003; Mempel et al. 2004; Bousso et al. 2003). Stopping is quantified using the **arrest coefficient**, the fraction of time where the speed is below some threshold (Fig. 1c; often 2 μm/min, but the ideal threshold can vary). As this fraction is computed by counting discrete steps, a more smooth, adapted version is preferable for downstream applications such as machine learning (ML) (Pizzagalli et al. 2019b).

2.2.2 Turning and Persistence Time

Another key track feature is the degree and frequency of turns in direction, which can vary substantially across cell types (Maiuri et al. 2015; Miller et al. 2003; Krummel et al. 2016) but can also reflect factors like confinement, obstacles, or cell-cell interactions (Beltman et al. 2007). A simple way to quantify turning is by analyzing **turning angles** between subsequent "steps" in the track, where lower angles indicate straighter motion (Fig. 1d). However, turning angle distributions can be difficult to interpret as they depend heavily on the imaging frame rate, limiting comparability across studies. Alternative metrics to describe turning behavior include the **displacement ratio**, **outreach ratio**, and **straightness** (Mokhtari et al. 2013) (Fig. 1b; where the latter is also referred to as "directness" Ibidi 2025, "confinement ratio" Beltman et al. 2009b; Mokhtari et al. 2013, or "directionality ratio" Gorelik et al. 2014). These metrics range between 0 and 1; however, their value depends strongly on the total duration of the track (Fig. 1b), and should be either compared between tracks of similar duration, or corrected for duration dependency (Beltman et al. 2009b; Gorelik et al. 2014). Track **asphericity** is also bounded by 0 and 1, but comes with the caveat that it is invariant to the order in which coordinates were visited (Mokhtari et al. 2013) (Fig. 1b). Another key metric is the **persistence time** representing how long motion remains relatively straight; however, this can be defined and measured in different ways as discussed in Sect. 3.1.

2.2.3 Directional Bias

Several metrics exist to detect directional bias arising when cells follow environmental structures or chemotactic gradients. The **chemotactic index** (also called "forward migration index" Ibidi 2025) is obtained by dividing the parallel component of the track's displacement ($D^{//}$) by the overall displacement (Fig. 1e); the **McCutcheon index** ("chemotactic ratio") divides $D^{//}$ by the total path length L instead (Fig. 1e) (McCutcheon et al. 1946). Another indicator of directionality is the angle of the cell's displacement vector to a reference point, direction, or plane. Without directional bias, the expected angle is 90° (although its distribution differs between 2D and 3D data; Beltman et al. 2009b, Fig. 1e). Enrichment of lower angles then indicates directed motion. This approach has proven useful to detect subtle directionality in various studies (Beltman et al. 2011; Ariotti et al. 2015).

All these approaches assume that the direction of bias is known. When it is not, directional bias can be detected using the **Rayleigh test** or **Hotelling's T^2 test**. The former is a non-parametric test against the null hypothesis that displacement directions are uniformly distributed (Ibidi 2025; Moore 1980); the latter instead takes the

average step-wise displacement across all tracks, and tests if it differs significantly from the null vector (Hotelling 1931; Oja et al. 2004; Textor et al. 2011) (Fig. 1f). A caveat of Hotelling's test is that it assumes independence of data points—an assumption that is violated when we extract subsequent steps from a track with directional persistence. This problem can be circumvented by downsampling the data, allowing sufficient spacing between steps that they are uncorrelated (Fig. 1f). It has been shown that Hotelling's test can detect directionality with similar sensitivity as angle analysis, even when the reference direction is unknown (Textor et al. 2011). Moreover, unlike angle analysis, it could detect "orthotaxis" (a mechanism where movement in the direction of the chemokine is not necessarily more frequent, but occurs at higher speed) (Fig. 1f). There are also cases where angle analysis is preferable; radial attraction towards a point may not be visible using Hotelling's (or Rayleigh's) test as displacement directions cancel out (Fig. 1f).

3 Analyzing Cell Dynamics Across Populations and Over Time

3.1 Applying Motility Metrics

3.1.1 Track-based, Tracklet-based, and Step-based Analyses

After defining a set of metrics, tracks can be described and compared between cell types or experimental conditions. These comparisons can be performed in multiple ways (Fig. 2a) (Beltman et al. 2009b; Mokhtari et al. 2013). The most straightforward approach is to compute metrics in a **track-based** fashion, with one value per cell track. However, this approach is highly misleading when tracks differ in duration. Some metrics are highly sensitive to these differences (Fig. 1b)—even if the cells do not truly differ in their underlying motion. Instead, analysis can be performed on **tracklets**, track segments of a fixed duration (also called "subtracks", Fig. 2a). This approach solves the issue of confounding by track duration. However, it discards tracks with durations below the chosen tracklet size, which can strengthen selection bias effects (see Sect. 3.3.1). A third method is to perform **step-based** analyses, computing metrics on pooled, single steps extracted from the data (Fig. 2a). This approach does not discard shorter tracks and therefore lessens selection bias, but it also has disadvantages. For example, it is sensitive to measurement noise in cell positions, less intuitive than track-based analysis, unsuitable for identifying subpopulations (Beltman et al. 2009b), and inapplicable to some metrics (e.g., "straightness" of a single step is meaningless). Statistical methods that combine strengths of both approaches have also been proposed (Letendre et al. 2015). Generally, whether an analysis should be track-, tracklet-, or step-based depends both on the research question at hand and the variation in duration among the tracks in the data.

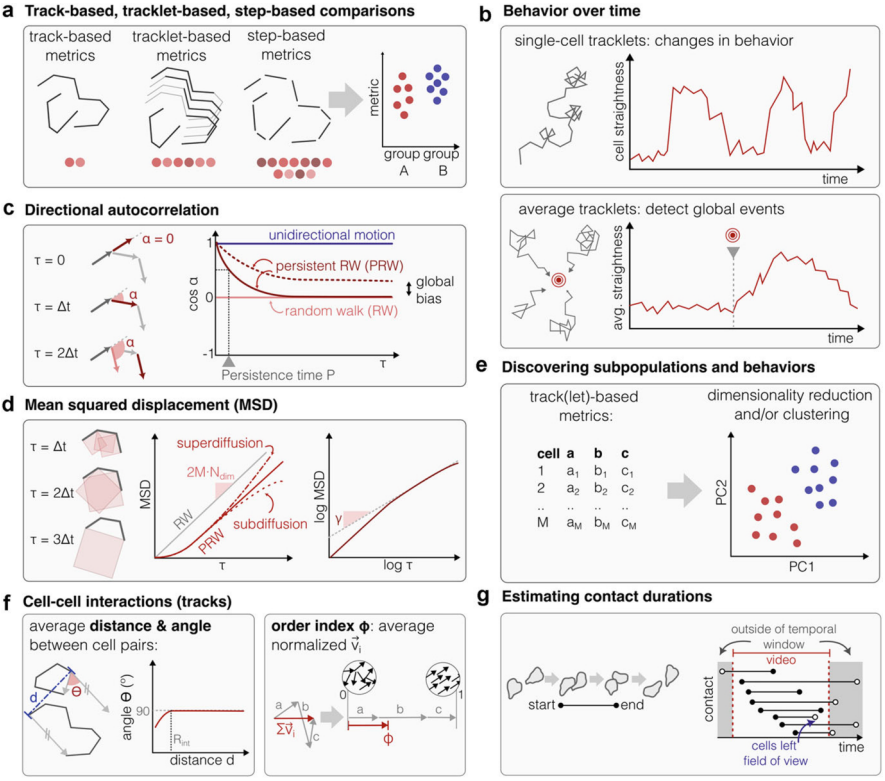

Fig. 2 Approaches to quantifying cell motility. (**a**) Metrics can be computed per track, per "tracklet" (track segment of fixed duration), or per individual "step". (**b**) Tracklet-based metrics can be analyzed over time, either for individual tracks (top, detecting changes in behavior), or for population averages (bottom, detecting global changes such as an increase in directed motion after an event of interest occurs). (**c**) The directional autocorrelation compares a cell's direction at time t to that at time $t + \tau$ to infer the persistence time. Reference curves are shown for an uncorrelated random walk (RW), persistent random walk (PRW) with and without global directionality, and unidirectional motion. (**d**) Like autocorrelation, the mean squared displacement (MSD) is computed over all possible tracklet durations τ. Expected MSD curves are shown for (P)RW motion and super-/subdiffusive motion. (**e**) Track features can be used for clustering and/or dimensionality reduction, allowing for exploration of subpopulations. (**f**) Distance-angle plots of cell pairs and the system's order index can help identify correlated, collective motion. (**g**) Physical contact can be detected from cell segmentations, but estimates for their duration are biased since we often do not observe the beginning and/or end. For unbiased estimates, contact durations should therefore be treated as time-to-event data

3.1.2 Single Cell Behavior Over Time

One advantage of step- or tracklet-based metrics is that they can be analyzed as a time series (Figs. 1c and 2b). When applied to single tracks, this approach can help identify changes in behavior over time, such as transitions between different migration modes (Crainiciuc et al. 2022). The same method can be applied to averages of tracklets from different cells in order to detect global events, for example an overall increase in directionality in response to tissue damage or sudden release of chemokines (Fig. 2b). An open question is to what extent general methodology from the field of time series can be applied to cell motility data to infer changes in behavior automatically. Simple methods such as hidden Markov models may be of use in this context (Schienstock et al. 2022), but applying them is non-trivial given the short tracks typically available in practical settings.

3.1.3 Expectations Over Time: Autocorrelation and Mean Squared Displacement

Since most metrics strongly depend on track duration (Fig. 1), several approaches involve computing these metrics on tracklets of varying duration τ in order to derive general, duration-independent motility parameters. One example is the directional **autocorrelation** function, defined as the cosine of the angle between the cell's direction at time t and $t + \tau$ (Fig. 2c). To construct the autocorrelation curve, this value is computed and averaged over all tracklets of the given duration τ. For an uncorrelated random walk (RW), the autocorrelation is zero for all $\tau > 0$, whereas for perfectly ballistic motion it always equals one.

For most cells, the autocorrelation is a function of τ, with high correlations at low τ indicating directional correlation in the short term. For such a **persistent random walk (PRW)**, the correlation decays approximately exponentially over longer time scales (Banigan et al. 2015) and becomes zero for $\tau >> P$—unless some other directional bias ensures further correlation on longer time scales. The half-life of this decay is the **persistence time** P and no longer depends on imaging frame rate or track duration. Alternative to the directional autocorrelation, the velocity autocorrelation (which includes not only the angle between vectors but also their magnitudes) is also frequently used (Banigan et al. 2015).

Similar to the autocorrelation, the **mean squared displacement (MSD)** is computed over all possible tracklet durations τ (Fig. 2d). Analytical derivations of the MSD are available for various motility models (Codling et al. 2008; Textor et al. 2013; Rosen et al. 2021). For example, the expected MSD for RW-motility is a linear function:

$$\mathrm{MSD}(\tau) = 2MN_{\mathrm{dim}}\tau \tag{1}$$

with N_{dim} the dimensionality, and M the **motility coefficient**, a parameter analogous to a diffusion coefficient. For PRW-motion, the MSD depends on both the motility

coefficient M and the persistence time P according to Fürth's equation (Fürth 1920):

$$\mathrm{MSD}(\tau) = 2MN_{\mathrm{dim}}\left[\tau - P(1 - e^{-\tau/P})\right] \quad (2)$$

which starts non-linear but becomes approximately linear for $\tau >> P$ (Fig. 2d). Fitting Fürth's equation to the empirical MSD, one can derive estimates for M and P that are mostly independent of track duration and imaging frame rate. However, for cells imaged at high frame rates, fitting a modified version of Fürth's equation may be more appropriate as cell positions (at high frame rates) are ill-defined (Thomas et al. 2020).

The linear MSD at large τ observed for (P)RW-motion is characteristic of diffusive behavior. MSD curves that grow faster or slower than linearly with τ can also be observed, indicating "superdiffusive" and "subdiffusive" behavior, respectively. These behaviors are also reflected by the slope γ of the log-log MSD plot (Fig. 2d); where $\gamma = 1$ indicates diffusive behavior, $\gamma = 2$ for ballistic motion, $1 < \gamma < 2$ is superdiffusive motion, and $\gamma < 1$ indicates subdiffusion.

While autocorrelation and MSD reduce dependency on track duration and imaging frame rate, they have some limitations. First, for larger τ, fewer tracklets are available for time averaging (Fig. 2c and d), increasing noise in both autocorrelation and MSD; exclusion of high τ values can therefore be required for robust fitting of P and M. Second, although deviations in MSD from regular (P)RW behavior can be of biological interest, they may also arise from selection bias imposed by the imaging window (see Sect. 3.3.1). Third, as fitting autocorrelation and MSD curves is less robust for individual tracks, P and M are typically derived for the entire population—potentially masking heterogeneity within the population (Mokhtari et al. 2013).

3.1.4 Exploring Subpopulations Using Clustering and Dimensionality Reduction

Finally, an interesting application of motility metrics is to discover subpopulations and behaviors in the data. After computing a set of metrics on each track (or tracklet), popular dimensionality reduction methods such as principal component analysis and UMAP (McInnes et al. 2018) can be used for data exploration, and clustering methods can help identify subpopulations (Fig. 2e). This approach can be applied both to entire tracks, revealing subpopulations of cells (McInnes et al. 2024), or to shorter tracklets, also showing cell switching between behavioral modes over time (Crainiciuc et al. 2022). However, as many metrics depend on imaging frame rate and track duration, clusters must be interpreted cautiously, especially when combining data from different experiments.

3.2 Quantifying Cell-Cell Interaction

The analyses discussed so far have mostly treated cells as independently moving objects—but a frequently asked question is to what extent cells affect the motion of their neighbors through cell-cell interaction.

3.2.1 Correlated Motion

Measurements of correlated motion, commonly studied in animal flocking and collective behavior (Vicsek et al. 2012), can also be applied to cell migration. In some settings such as wound healing and tissue development (Camley and Rappel 2017), cell-cell interaction gives rise to collective motion with cells aligning on a global scale. The resulting directional bias can then be detected as discussed in Fig. 1e and f. Another frequently used metric to detect alignment in a collective is the **order index**, or order parameter, ϕ, which ranges from 0 to 1 (Fig. 2f) (Vicsek et al. 2012) and has also been applied to cell migration (Banigan et al. 2015; Niculescu et al. 2015; Szabó et al. 2006):

$$\phi = \frac{\left|\sum_i^N \mathbf{v}_i\right|}{\sum_i^N |\mathbf{v}_i|} = \frac{1}{N v_{\text{avg}}} \left|\sum_i^N \mathbf{v}_i\right| \tag{3}$$

with N the number of tracks, tracklets, or steps and $\mathbf{v}_i$ their velocity vectors.

However, many cells do not align their directions globally, and only align with their local neighbors—at the very least, some interaction is expected for distances close to the average cell radius due to physical exclusion effects. One way to take this spatial scale into account is to consider the angles and distances between pairs of track(let)s in the data (Fig. 2f) (Beltman et al. 2009b). Interactions result in average angles below 90° within an interaction distance R_{int}. If R_{int} is substantially larger than the average cell radius, this may indicate that cells are aligning over larger distances than expected from physical exclusion alone.

3.2.2 Cell-Cell Contact Dynamics

The dynamics of cell-cell contact are key for biological processes such as immune responses (Mempel et al. 2004; Davis 2009). Unfortunately, physical contact is hard to infer directly from tracks. While cells within a minimal threshold distance (R_{min}) can be marked as "contacting", this is often inaccurate as distance distributions for contacting and non-contacting cells can overlap substantially due to variability in cell sizes and shapes (Wortel et al. 2024). When available, segmentation masks (Fig. 1a) can help detect contact more reliably. Yet even when contacts are detected with complete accuracy in single frames, estimating contact duration remains difficult. We often do not observe their beginning and/or end—for example when the recording ends or when cells leave the field of view before the contact is resolved (Fig. 2g) (Beltman et al. 2009b). This makes cell-cell contact durations a form **time-to-event data** that is both "truncated" (as we do not observe contacts at all when they fall outside the imaged time window) and "right-censored" (as we underestimate the

duration when either the beginning or end of the interaction is unobserved). These issues can cause substantial bias in the estimation of contact durations and should be corrected for; for details, see Beltman et al. (2009a).

3.3 Robustly Interpreting Track-based Analyses

Many methods exist to quantify cell motion and interactions, but they can introduce bias if not applied carefully. Here, we describe potential pitfalls in the classical track analysis pipeline, ways to mitigate them, and how simulations can facilitate robust data interpretation.

3.3.1 Artefacts from imaging, detection and tracking

Errors during image acquisition, cell detection, and tracking can introduce artifacts in downstream analysis. Common imaging errors include **tissue drift**—where movement of the tissue specimen during imaging adds a global drift velocity on top of actual cell movement, and **calibration errors** in the axial (z) dimension, causing false anisotropy. Depending on the algorithms used, common errors from cell detection and tracking include: **track gaps**—when cells are not detected in intermediate frames, resulting in tracks with either inconsistent time step (causing outliers in step-based metrics) or **track splitting** (causing an abundance of short tracks). **Track switching** occurs when cell paths cross each other, and the algorithm mistakenly switches the cell identity at the crossover point; frequent switches can bias track metrics towards preferences of the linking algorithm (e.g., a bias towards straighter tracks). Finally **border tracking** is an artefact causing artificially straight tracks at the image border. These errors have been extensively discussed elsewhere (Beltman et al. 2009b; Banigan et al. 2015), and we refer the reader to these resources for strategies to diagnose and mitigate them.

Another issue in track-based analysis is **detection noise**: as cells are deformable, their position is ill-defined (Thomas et al. 2020), and positions estimated during tracking can fluctuate even for static cells. This noise can distort motility metrics, especially for short tracklets, leading to overestimated instantaneous speeds and directional changes, an underestimated arrest coefficient, a faster decay of directional autocorrelation (Fig. 2c) and an underestimated persistence time. Importantly, as detection noise may differ between cell types with different (nuclear) shapes and sizes, this effect can create the illusion of motility differences. To diagnose this, it is useful to assess tracks of static cells, and to manually visualize detections on top of the original image. Noise can be reduced by smoothing raw tracks before track analysis using methods such as Kalman smoothing (Kalman 1960) (for an example, see Wortel et al. 2024). However, these methods also come with parameters that should be carefully tuned to avoid removing true variation.

3.3.2 Bias from Imaging Window

A key issue in time-lapse microscopy is the limited size of the imaging window, especially in the axial dimension for in vivo imaging, which introduces **selection**

bias: faster, straighter cells leave the imaging window sooner, and thus have shorter tracks. This effect has been shown to cause substantial artefacts in some analyses—especially for cell- and tracklet-based analyses where tracks below a minimum duration are removed (Beltman et al. 2009b). For example, in MSD curves it can give a false impression of subdiffusive behavior (Fig. 2d) as faster cells contribute progressively less to the average as τ increases (Beltman et al. 2009b). Similarly, directional bias may go unnoticed in MSD analysis (Textor et al. 2011). Plotting track metrics (e.g., speed) against track duration can help diagnose this bias.

3.3.3 Dealing with Short Tracks

A common problem in tracking data—especially after automated tracking—is the occurrence of short-duration tracks. Short tracks are often removed before further analysis, but this approach strongly exacerbates the effects of selection bias described above. If selection bias is suspected, step-based analyses are recommended instead (Sect. 3.1) (Beltman et al. 2009b). Short tracks can also result from undetected cells in certain frames, splitting tracks into multiple shorter segments. For analyses that require tracks of longer durations, such as MSD and directional autocorrelation, it may therefore be preferable to keep tracks intact and interpolate missing positions, to include as many cells as possible (see e.g., supplementary materials of Wortel et al. (2021) for details).

3.3.4 Video Heterogeneity

Finally, a common issue in motility analysis is the substantial variation in motility between video replicates, due to biological or technical factors like timing or temperature differences. This variation can obscure differences between populations across videos. In settings where heterogeneity is substantial, it is useful to include an internal control population that occurs in all videos and for which the motility can be assumed to be consistent (see e.g., the supplements of Wortel et al. 2024; McInnes et al. 2024).

3.3.5 Improving Interpretations Using Simulation

Cell motility analysis can reveal interesting biological mechanisms, but their sensitivity to measurement errors and biases means they should be interpreted carefully. An interesting approach to assess the robustness of track-based analysis is through **simulation**. For example, common issues such as track gaps, track splitting, track switching, or a limited imaging window can be artificially introduced in either real or simulated data to check sensitivity (see, for example, Textor et al. 2011). Similarly, detection noise can be introduced artificially to real or simulated data to evaluate its impact on analysis, helping determine how accurate the tracking needs to be. Simulations of simple models, such as random walks (Codling et al. 2008), are easy to implement, and some are available in open-source tools (e.g., MotilityLab and CelltrackR Textor et al. 2015; Wortel et al. 2021).

Simulations can also help design robust analyses to detect mechanisms of interest. For example, the stop-and-go behavior of lymphocytes in lymph nodes was originally attributed to an "internal clock" after Fourier analysis seemed to indicate

periodicity in speeds; however, later simulations with and without an internal clock (akin to a positive and negative experimental control) showed that the data were more consistent with the baseline scenario (Beltman et al. 2007).

4 Optical Flow

The quantification approaches discussed so far have all relied on object tracking. Alternatively, it is possible to bypass object tracking entirely using **Optical Flow** (OF). OF is similar to tracking, but instead of tracking object centroids, it tracks every pixel across adjacent images (Fig. 3a). It does not require object detection, making it useful when objects cannot be defined or detected, and provides additional information on morphological changes over time. OF is applied in various fields, such as studying human behavior (Perš et al. 2010), cell transportation in blood vessels (Guo 2013), and heart motion (Gröger et al. 2006), and can be combined with motion models to enhance tracking (Rodrigues et al. 2014).

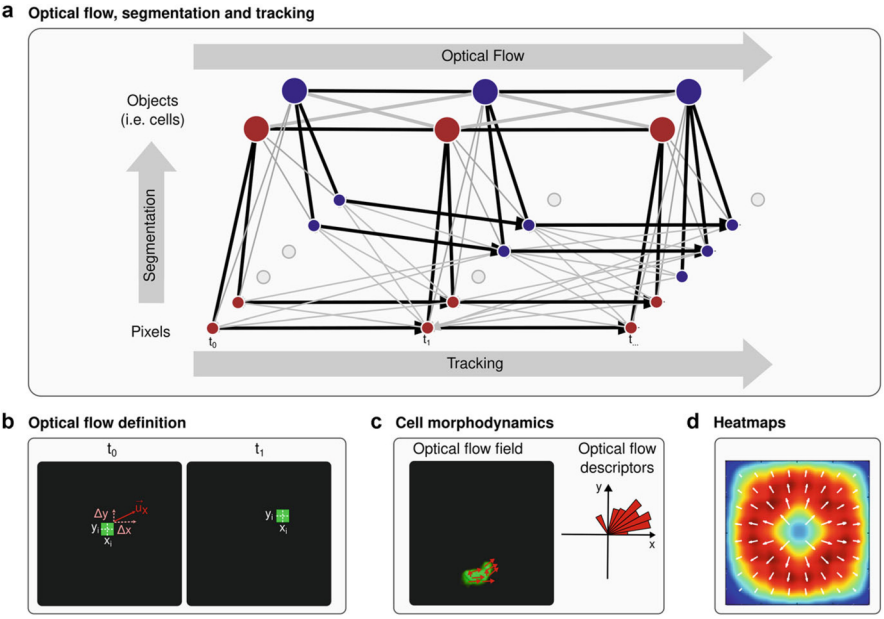

Fig. 3 Schematic overview of the OF calculation process. (**a**) OF is a pixel tracking process. (**b**) For each pixel i, the displacement vector $\mathbf{u}_x$ is calculated, mapping its position (x_i, y_i, t_0) to (x_i, y_i, t_1). (**c**) Global descriptors are derived from the OF vectors. (**d**) OF is applied to generate tissue-level motility maps

4.1 Definition and Considerations for Optical Flow in Biomedical Imaging

OF was stated in the 1940s as the apparent movement of objects due to changes in the point of view. In the 1980s, numerical methods were introduced to link pixels of an image at time t with pixels of an image at time $t + 1$ (Horn and Schunck 1981; Lucas and Kanade 1981), providing a basis for analyzing motile objects. Consider an image sequence, with space-time coordinates $\mathbf{x} = [x_1, \ldots, x_{d-1}, t]$, whose dimension d equals the number of spatial dimensions plus 1. The image series can then be represented by a function $I(\mathbf{x}) \in \mathbb{R}$ describing the brightness value at position $\mathbf{x}$. Assume further that we are imaging physical objects, whose motion between consecutive frames is described by a displacement field $U(\mathbf{x})$ mapping each point $\mathbf{x}$ at time t to a new position $\mathbf{x} + \mathbf{u}_x$ at time $t + 1$, where $\mathbf{u}_x = [\Delta x_1, \ldots, \Delta x_{d-1}, 1]$. OF is then the problem of estimating $U(\mathbf{x})$ (Fig. 3b).

In the original formulation by Horn and Schunck (Horn and Schunck 1981), the problem is solved by assuming brightness constancy—i.e., a point in the image at time t retains the same brightness value as it moves in the image at time $t + 1$:

$$I(\mathbf{x}) = I(\mathbf{x} + \mathbf{u}_x).$$

For sufficiently small displacements, the brightness constancy assumption can be approximated using a first-order Taylor expansion:

$$I(\mathbf{x} + \mathbf{u}_x) \approx I(\mathbf{x}) + \nabla I \cdot U(\mathbf{x})$$

where $\nabla I = \left[\frac{\partial I}{\partial x_1}, \ldots, \frac{\partial I}{\partial x_{d-1}}, \frac{\partial I}{\partial t}\right]^\top$ is the spatiotemporal gradient of I. Neglecting the higher-order terms of the Taylor expansion, this yields the *OF constraint*:

$$\nabla I \cdot U(\mathbf{x}) = 0. \tag{4}$$

Equation (4) provides a single constraint for each pixel, despite the $d - 1$ unknowns in $U(\mathbf{x})$. For instance, in the case of 2D images over time ($d = 3$), the constraint reduces to:

$$\frac{\partial I}{\partial x}\Delta x + \frac{\partial I}{\partial y}\Delta y + \frac{\partial I}{\partial t} \cdot 1 = 0.$$

To obtain a unique solution, additional constraints and regularization terms must be introduced. These also allow to prefer desirable properties, such as smoothness. An optimal U^* is computed by minimizing an energy functional:

$$E = E_D + \lambda E_S, \tag{5}$$

where E_D represents the data term (Eq. 4), E_S is a regularization term, and λ controls the impact of regularization.

In Horn and Schunck (1981), E_S is defined as a quadratic penalty term on the gradient of the solution, which discourages large discontinuities in the flow:

$$E(u) = \int_{\Omega} \underbrace{(I(\mathbf{x}) - I(\mathbf{x} + U(\mathbf{x})))^2}_{\text{data term}} \, d\mathbf{x} + \lambda \int_{\Omega} \underbrace{\|\nabla U(\mathbf{x})\|^2}_{\text{smoothness term}} \, d\mathbf{x}. \tag{6}$$

More recent works (Zach et al. 2007; Sánchez et al. 2013) employ different regularization techniques to allow flow discontinuities and increase robustness to noise. However, Eq. 4 is only valid within a sufficiently small local neighborhood. Errors arise in scenarios involving large displacements, texture-less regions, or displacements parallel to image edges. Various methods address these challenges. For instance, additional features to brightness (e.g., SIFT) can be used to match points (Brox and Malik 2011), whereas coarse-to-fine approaches can improve accuracy for sequences with large displacements (Chang et al. 2013). These include pyramidal image decomposition with TV-L1 regularization (Sánchez et al. 2013) and temporal superpixel clustering (Chang et al. 2013).

Computing OF with time-lapse microscopy videos presents specific challenges. First, OF operates at pixel level, but biological interpretations often require object-level abstractions. Second, the assumption of constant pixel intensity between frames may not hold for fluorescence microscopy or transmitted light imaging due to photobleaching, changes in focal planes, or repositioning of the objective in high-throughput devices. Finally, the assumption of small displacements may not hold for large time intervals between frames and rapid moving cells. Coarse-to-fine methods combined with probabilistic frameworks (Amat et al. 2013) were proposed to address these issues, enabling the quantification of large cellular movements, but required segmentation masks to exclude background regions.

More recently, DL has revolutionized the robustness of OF estimation. By leveraging large synthetic and/or real-world data sets of high quality (high frame rates, no artifacts), networks can be first trained to reproduce classical OF formulations as "ground truth", then trained to reproduce these estimates after temporal downsampling and introduction of artifacts. Among these, FlowNet (Dosovitskiy et al. 2015) used convolutional neural networks for end-to-end OF estimation, and FlowNet2 and FlowNet3 (Ilg et al. 2018a,b) enhanced accuracy through multi-stream architectures to address both small and large displacements. DeepFlow (Weinzaepfel et al. 2013) combined deep feature extraction using convolutional neural networks with traditional OF formulations, and EpicFlow (Revaud et al. 2015) incorporated edge-preserving interpolation and geodesic distance from edges to improve flow estimation with large motions and occlusions. For further details, methods, and datasets on OF, see: Shah and Xuezhi (2021).

4.2 OF Applications to Intracellular, Cellular and Tissue Dynamics

OF can quantify cellular morphodynamics at various biological scales, including organelles, cytoskeleton, and tissues, especially when tracking is not possible. For example, trafficking of intracellular organelles and actin filament remodeling are difficult to track due to the limit of optical resolution and absence of a clear centroid in these complex structures—but OF can still provide quantitative, pixel-level information on movement orientation and magnitude. These complement morphological features used in digital pathology and were shown to offer additional insights, such as changes in motor proteins in healthy or pathological conditions (Huang et al. 2017; Drechsler et al. 2020).

At the cell level, morphodynamics result from the interaction of intracellular factors (i.e., ATP availability) and extracellular factors (i.e., extracellular matrix stiffness). At this level, OF can provide insights into the dynamic processes controlling cell migration, such as polarization towards a direction, revealing the effects of therapeutic agents, or modifying extracellular factors (Boric et al. 2013). Flow vectors can also be studied inside specific cells of interest, e.g., using binary masks or segmentations (when available) or an intracellular area defined through other means. Then, selected vectors can be described using the **histogram of optical flow** (HOF, Perš et al. 2010) and/or **histogram of oriented optical flow** (HOOF, Wang and Snoussi 2013) to distinguish different cellular behaviors, as discussed in Sect. 5.3 (Fig. 3c). In addition to cellular and subcellular applications, OF techniques have been applied to analyze tissue dynamics, including embryo development (Amat et al. 2013) and immune responses in lymphoid tissues (Pizzagalli et al. 2019b). In this case, OF provides insights at larger scale, such as flow fields (Schienstock et al. 2024) for identifying collective cell behavior, or tissue velocity heatmaps to identify areas with different motility patterns (Pizzagalli et al. 2019b) (Fig. 3d).

5 Action Recognition

Action recognition (AR) is a computer vision technique for identifying where and when specific actions in video or image sequences occur. An **action** is any meaningful interaction between a subject (human, animal, or cell) and its surroundings (Herath et al. 2017). AR has been applied in video surveillance, autonomous driving, and sports analysis (Vishwakarma and Agrawal 2013; Bukht et al. 2024; Karpathy et al. 2014). It has a huge potential for analyzing cell behavior in (patho)physiological processes like immune responses, cancer metastasis, and tissue regeneration, supporting the development of diagnostics and therapies (Pizzagalli et al. 2022b; Schienstock et al. 2022; Waite et al. 2011; Hanna et al. 2015; Singer and Clark 1999).

In microscopy videos, the types of actions that can be observed and quantified vary depending on cell type and imaging technique. At the single-cell level, key actions include cell cycle events (mitosis, cell death), biological functions (autophagy, phagocytosis, endocytosis, exocytosis), and adopting specific cell migration patterns (e.g., random patrolling in confinement, directed movement towards targets, or arrested states (Auffray et al. 2007; Arasa 2021; Friedl and Weigelin 2008)). Actions at the collective level involve interactions and communication between multiple cells—either through direct physical contact (essential for processes like diapedesis or precise delivery of stimuli to target cells via synapses), or indirectly through the secretion of soluble factors such as cytokines and chemokines (regulating collective behaviors like chemotaxis and swarm formation (Celli et al. 2007; Kienle et al. 2021)). We summarize here three main categories of AR techniques: track-based, segmentation-based, and video-based AR. For details, refer (Pizzagalli et al. 2022b; Vrigkas et al. 2015; Herath et al. 2017).

5.1 Track-based Action Recognition

Track-based methods (Fig. 4a) involve three main steps:

1. **Tracking**: trajectories of cells or particles are obtained as described in Sect. 2.1.
2. **Feature Extraction**: trajectories are described using features such as motility metrics (Sect. 2.2), derived from entire trajectories or tracklets (Sect. 3.1), or other features describing single or multiple objects.
3. **Classification**: trajectories can be classified using predefined thresholds on motility metrics (for example, on track straightness to distinguish directed migration from cell contacts (Pizzagalli et al. 2022b; Pécot et al. 2018)). Actions involving more complex patterns instead can be detected via supervised learning if sufficient annotations are available (Wagner et al. 2017).

Steps 2 and 3 can be combined with DL methods that automatically learn and classify features directly from trajectory data, represented as coordinates or image-based trajectories. For example, convolutional neural networks (CNNs) have been used to classify motility patterns by training on 2D image representations of trajectories (Nicolai et al. 2024).

5.2 Segmentation-based Action Recognition

Segmentation-based methods use the surfaces of objects and their temporal dynamics to enable AR (Fig. 4b). These methods generally follow four steps:

1. **Segmentation or Detection**: objects of interest are segmented from the background in the image, either manually or automatically (section "Cell Detection"). Automatic techniques range from classical approaches like thresholding to

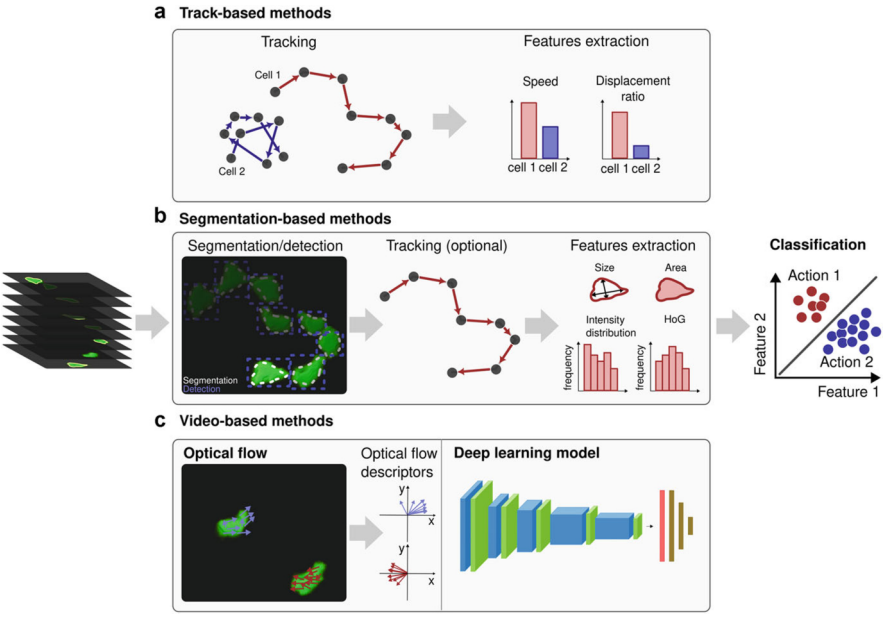

Fig. 4 Schematic representation of the three main classes of AR methods. Each method takes a microscope video as input and outputs the corresponding action class. (**a**) Track-based methods involve tracking individual cells and extracting relevant features. (**b**) Segmentation-based methods segment or detect cells, optionally track them, and extract features from the selected regions. (**c**) Video-based methods analyze the video directly, leveraging techniques such as OF (left panel) or DL approaches, exemplified here by the 3D CNN (right panel)

advanced DL methods, such as CNNs for robust object delineation (Liu et al. 2024; Lee et al. 2020; Cabini et al. 2024a, 2025a,b). In some cases, detection may involve generating bounding boxes around the objects instead of precise delineation, which is useful when exact boundaries are not needed (Pulfer et al. 2024; Anandakumaran et al. 2022; Delgado et al. 2024).

2. **(Optional) Tracking**: segmented objects are tracked across consecutive frames (see Sect. 2.1). This step is not necessary for all applications: for example, detecting contact between two cells in a microscopy video may not require tracking, but tracking is useful for identifying contact between a specific cell and its target.
3. **Feature Extraction**: once segmented (and optionally tracked), features are extracted from the objects. Single-object features describe morphological or statistical properties such as diameter, surface area, volume, intensity distribution, and OF vectors. Multiple-object features describe relationships between segmentations, such as the relative distance between surfaces or spatial arrangements, and can capture temporal changes if tracking is applied (Liu et al. 2024).
4. **Classification**: the extracted features are then classified using filters or supervised learning as outlined in Sect. 5.1.

Steps 3 and 4 can be combined using DL methods, where a single model simultaneously handles feature extraction and classification, learning relevant features directly from the data (Sillah et al. 2024).

5.3 Video-based Action Recognition

This category includes methods that process videos directly, without requiring preliminary steps such as tracking or segmentation (Fig. 4c). These methods are particularly useful in contexts where tracking or segmenting videos is impossible or prone to errors. While numerous video-based methods exist for other domains, their application for cellular AR is still limited.

5.3.1 Optical Flow Features

A major family of video-based methods relies on computing OF between consecutive frames (Fig. 4c, Sect. 4). One common OF-based approach to AR is to extract hand-crafted motion descriptors, such as HOF and HOOF (Uijlings et al. 2015) (Sect. 4), from OF vectors, which can then be used for classification using filtering-based approaches, supervised or unsupervised learning models. Another common approach is to track small image regions, called particles, over time based on OF vectors, generating trajectories (Wang et al. 2013). Features, such as chaotic invariants defined in (Wu et al. 2010; Ali and Shah 2007), are extracted from these trajectories and classified using ML models.

5.3.2 Video Classification

DL methods have become popular in AR due to their ability to automatically learn features from raw video data (Fig. 4c). Key DL architectures include **CNNs**, **multi-stream networks**, recurrent neural network (**RNN**)-based methods, and **transformer**-based methods (Herath et al. 2017; Zhu et al. 2020). CNNs use convolutional layers to capture both spatial and temporal information from raw video data, treating the input as volumetric data. These architectures typically consist of a stack of convolutional layers to learn spatiotemporal features, followed by pooling layers to reduce spatial and temporal dimensions. Convolutional layers can employ 3D convolutions, which operate simultaneously across spatial and temporal dimensions (Ji et al. 2013; Tran et al. 2014) (Fig. 4c), or a combination of 2D convolutions (for spatial dimensions) and 1D convolutions (for the temporal dimension), reducing the total number of parameters in the network (Sun et al. 2015).

Two-stream networks are a class of CNNs designed to perform object recognition and motion recognition separately. The spatial stream takes raw video frames as input to capture visual appearance information. The temporal stream takes a stack of OF images as input to capture motion information between video frames. The outputs from both networks are then fused to obtain the final prediction (Simonyan and Zisserman 2014; Tu et al. 2018).

RNN-based models exploit the capability of RNNs to handle temporal data, learning long-term relationships in the data. They consist of a 2D CNN that is simultaneously applied to each video frame separately, followed by an RNN that receives the extracted features as input and makes the final action prediction (Donahue et al. 2017; Baccouche et al. 2011). Some examples of RNNs include Long Short-Term Memory and Gated Recurrent Units.

Transformer-based methods have recently emerged as powerful tools for modeling temporal tasks, including AR. Similar to RNN-based methods, they operate on features extracted by a 2D CNN from individual video frames to produce the output prediction. However, transformers utilize a self-attention mechanism, allowing them to assess the importance of each part of the input data and capture long-range dependencies without the need for sequential computation, as in RNN-based methods (Ulhaq et al. 2022; Girdhar et al. 2019; Cabini et al. 2024b).

6 Conclusions and Future Perspectives

The development of methods for quantifying movement in biomedical videos is a current frontier of research both in computational and life sciences. In principle, the concepts and algorithms behind movement analysis methods can be applied to objects with different sizes and image modalities, quantifying movement across scales. In practice, however, many computer vision methods are currently more established for images and videos capturing natural scenes such as traffic and human behavior, and not all of them translate well to bioimaging. A more effective exchange of methods, datasets, and complementary needs between scientists in computer vision and life sciences will be crucial to overcome this gap in biomedical computer vision.

6.1 Overcoming the (Annotated) Data Gap

Modern computer vision owes its success in part to the availability of large, open, and well-curated datasets such as MNIST and ImageNet. However, a key challenge for the translation of bioimaging into the biomedical domain is that the subjects of bioimaging are substantially different from those in these classical datasets. For example, cells often lack stable features such as rigid edges or textures, ensuring that most general-purpose computer vision models trained on other types of data are not directly applicable to bioimaging applications unless trained on domain-specific data. This is especially relevant for DL methods, which typically require very high numbers of images for training. Even if models pre-trained on non-biomedical images can successfully be translated to the bioimaging domain using transfer learning, this still requires the availability of large, domain-specific and well-annotated datasets. And while these are increasingly available for medical imaging used in clinical practice (National Institute for Health Research 2025;

Stettler 2025), they are less widespread for the broad range in imaging techniques used in preclinical research.

Progress has been made in important initiatives such as the Cell Tracking Challenge (Maška et al. 2014, 2023; Ulman et al. 2017), as well as open data available in e.g., MotilityLab (Textor et al. 2015), the leukocyte tracking database (Pizzagalli et al. 2018), and Immunemap (McInnes et al. 2024), which effectively facilitate data exchange and support algorithm development. Yet annotated open datasets remain scarce given the large variety in imaging modalities, (fluorescent) labelling, and experimental settings—hampering the development of computer vision technology for more complex tasks. Especially in AR, annotations are critical for training action classifiers, and challenges arise from limited data availability and confounding motility patterns (i.e., mitosis vs. separation of two touching cells).

To overcome this issue, smart annotation strategies are likely required, e.g., by making annotations in an iterative manner: identifying cases yielding false positives / false negatives and adding them to the training dataset (Pulfer et al. 2024). A key direction of future research is how to use limited annotations more effectively, and/or how to develop active learning strategies to generate annotations where these are most needed. Likewise, an open question is how we can more effectively leverage simulations in computer vision for bioimaging, for example to generate large synthetic datasets for (pre-)training, or to generate annotated data for AR tasks where the ground truth is difficult and/or time-consuming to obtain (Brekke et al. 2019). In the past several decades, many computational models have been developed to simulate a wide range of biological dynamical processes—from cell sorting, to cell motility, to tissue development, collective migration, and many others (Graner and Glazier 1992; Beltman et al. 2007; Niculescu et al. 2015; Hirashima et al. 2017; Szabó and Merks 2013; Camley and Rappel 2017; Montagud et al. 2021). Their integration into computer vision development would be a promising direction of future research.

6.2 Overcoming Structural Barriers to Open Science Practices

A barrier to the availability of open data is the current lack of recognition and reward systems for publishing well-annotated datasets. Whereas taking and labelling pictures of cats and dogs is relatively cheap, generating specialized bioimaging data is typically (very) expensive. And even once raw data are available, providing annotations is often time-consuming as well. As long as the generation of data is not recognized as a meaningful contribution in and of itself, this provides an incentive to delay sharing of the data and/or metadata in order to capitalize on a dataset before making it available for the community to use. Solving this issue will require cultural changes as well as new perspectives on how we reward scientific contributions. Despite ongoing efforts (Devriendt et al. 2021), this area will remain in need of attention.

Aside from well-curated open datasets with rich metadata, the effective development and use of bio-image analysis methods also requires accessible (ideally open-source) analysis tools. Software should be accessible through key usability features including user interfaces, well-structured code, interoperability with different file formats and other tools, and a modular structure, and the analysis tools offered should be validated (Carpenter et al. 2021). Examples include larger platforms with community-supplied plugins such as Fiji (Schindelin et al. 2012) and Icy (Chaumont et al. 2012), but also web-based tools (Textor et al. 2015), or packages built on the existing ecosystems of popular programming languages such as R and python (Wortel et al. 2021). Despite ongoing efforts in the development of open-source software for bioimaging data, current research still often relies on custom scripts and/or proprietary software. Further supporting the transition to open-access will require the continued investment in Open Science practices such as open-access repositories, collaborative platforms, organization of scientific challenges, and release of open-source tools that can be customized or improved for different applications.

6.3 Learning Meaningful Representations of Dynamics

Finally, a promising area of ongoing research is to what extent ML can supplement hand-crafted motility metrics (Sect. 2.2) by facilitating the discovery of novel descriptors of dynamical processes, for example through self-supervised representation learning. An important caveat here is that machine-learned features will likely suffer from the same challenges and biases as hand-crafted ones (Sect. 3.3)—unless explicit care is taken to avoid issues such as confounding by track duration or bias introduced by the imaging window. A key challenge will therefore be to harness lessons learned from the application of hand-crafted features, ensuring that learned representations are meaningful.

6.4 Final Remarks

Scientific discoveries often go hand in hand with technological developments (Masters 2008). In the field of bioimaging, these developments have long been driven by the development of new microscopes and labelling techniques—but are increasingly driven by advances in the methodology used to process and analyze bioimaging data as well. Exciting developments in AI are working their way into the bioimaging domain, but some barriers remain to the optimal synergy between the fields of bioimaging and computer vision. Strengthening this synergy will be key to drive new discoveries in basic research and open up new avenues for clinical applications such as disease monitoring and precision therapeutics.

Acknowledgments DUP was supported by the Swiss National Science Foundation with grant 228512 and USI FIR grant, RFC was supported by USI FIR grant and swissuniversities with RE2VITAL CHORD B grant. IMNW was supported by AiNed Fellowship grant NGF.1607.22.020 from the Dutch Research Council (NWO).

Conflict of Interest Authors declare no conflict of interests.

References

Al-Kofahi O, Radke RJ, Goderie SK, Shen Q, Temple S, Roysam B (2006) Automated cell lineage construction: a rapid method to analyze clonal development established with murine neural progenitor cells. Cell Cycle 5(3):327–335

Ali S, Shah M (2007) A Lagrangian particle dynamics approach for crowd flow segmentation and stability analysis. In 2007 IEEE conference on computer vision and pattern recognition. IEEE, pp. 1–6

Amat F, Myers EW, Keller PJ (2013) Fast and robust optical flow for time-lapse microscopy using super-voxels. Bioinformatics 29(3):373–380

Anandakumaran PN, Ayers AG, Muranski P, Creusot RJ, Sia SK (2022) Rapid video-based deep learning of cognate versus non-cognate T cell-dendritic cell interactions. Sci Rep 12(1):559

Apostoloff N, Fitzgibbon A (2005) Learning spatiotemporal T-junctions for occlusion detection. In IEEE computer society conference on computer vision and pattern recognition, 2005 (CVPR 2005), vol. 2, IEEE, pp. 553–559

Apostoloff N, Fitzgibbon AW (2006) Automatic video segmentation using spatiotemporal T-junctions. Bmvc, pp. 1089–1098

Arasa J, Collado-Diaz V, Kritikos I, Medina-Sanchez JD, Friess MC, Sigmund EC, Schineis P, Hunter MC, Tacconi C, Paterson N et al (2021) Upregulation of VCAM-1 in lymphatic collectors supports dendritic cell entry and rapid migration to lymph nodes in inflammation. J Exp Med 218(7):e20201413

Arganda-Carreras I, Kaynig V, Rueden C, Eliceiri KW, Schindelin J, Cardona A, Sebastian Seung H (2017) Trainable weka segmentation: a machine learning tool for microscopy pixel classification. Bioinformatics 33(15):2424–2426

Ariotti S, Beltman JB, Borsje R, Hoekstra ME, Halford WP, Haanen JBAG, de Boer RJ, Schumacher TNM (2015) Subtle CXCR3-dependent chemotaxis of CTLs within infected tissue allows efficient target localization. J Immunol 195:5285–5295

Auffray C, Fogg D, Garfa M, Elain G, Join-Lambert O, Kayal S, Sarnacki S, Cumano A, Lauvau G, Geissmann F (2007) Monitoring of blood vessels and tissues by a population of monocytes with patrolling behavior. Science 317(5838):666–670

Baccouche M, Mamalet F, Wolf C, Garcia C, Baskurt A (2011) Sequential deep learning for human action recognition. In Human behavior understanding: second international workshop, HBU 2011, Amsterdam, The Netherlands, November 16, 2011. Proceedings 2. Springer, pp. 29–39

Banigan EJ, Harris TH, Christian DA, Hunter CA, Liu AJ (2015) Heterogeneous CD8+ T cell migration in the lymph node in the absence of inflammation revealed by quantitative migration analysis. PLOS Comput Biol 11:1–20

Beltman JB, Marée AF, Lynch JN, Miller MJ, de Boer RJ (2007) Lymph node topology dictates T cell migration behavior. J Exp Med 204:771–780

Beltman JB, Henrickson SE, von Andrian UH, de Boer RJ, Marée AF (2009a) Towards estimating the true duration of dendritic cell interactions with T cells. J Immunol Methods 347(1):54–69

Beltman JB, Marée AFM, de Boer RJ (2009b) Analysing immune cell migration. Nature Rev Immunol 9:789–798

Beltman JB, Allen CDC, Cyster JG, de Boer RJ (2011) B cells within germinal centers migrate preferentially from dark to light zone. Proc Natl Acad Sci 108(21):8755–8760

Berclaz J, Fleuret F, Turetken E, Fua P (2011) Multiple object tracking using k-shortest paths optimization. IEEE Trans Pattern Anal Mach Intell 33(9):1806–1819

Berg S, Kutra D, Kroeger T, Straehle CN, Kausler BX, Haubold C, Schiegg M, Ales J, Beier T, Rudy M et al (2019) Ilastik: interactive machine learning for (bio) image analysis. Nature Methods 16(12):1226–1232

Boric K, Orio P, Viéville T, Whitlock K (2013) Quantitative analysis of cell migration using optical flow. PloS One 8(7):e69574

Bousso P, Robey E (2003) Dynamics of CD8+ T cell priming by dendritic cells in intact lymph nodes. Nature Immunol 4:579–585

Brekke Å, Vatsendvik F, Lindseth F (2019) Multimodal 3D object detection from simulated pretraining. In Symposium of the Norwegian AI society. Springer, pp. 102–113

Brox T, Malik J (2011) Large displacement optical flow descriptor matching in variational motion estimation. IEEE Trans Pattern Anal Mach Intell 33(3):1–14

Bukht TFN, Rahman H, Shaheen M, Algarni A, Almujally NA, Jalal A (2024) A review of video-based human activity recognition: theory, methods and applications. Multimedia Tools Appl 1–47

Cabini RF, Tettamanti H, Zanella M (2024a) Understanding the impact of evaluation metrics in kinetic models for consensus-based segmentation. Preprint. arXiv:2412.03458

Cabini RF, Barzaghi L, Cicolari D, Arosio P, Carrazza S, Figini S, Filibian M, Gazzano A, Krause R, Mariani M et al (2024b) Fast deep learning reconstruction techniques for preclinical magnetic resonance fingerprinting. NMR Biomed 37(1):e5028

Cabini RF, Pichiecchio A, Lascialfari A, Figini S, Zanella M (2025a) A kinetic approach to consensus-based segmentation of biomedical images. Kinetic Relat Models 18(2):286–311

Cabini RF, Cozzi A, Leu S, Thelen B, Krause R, Del Grande F, Pizzagalli DU, Rizzo SMR (2025) Compositia: an open-source automated quantification tool for body composition scores from thoraco-abdominal CT scans. Eur Radiol Exp 9(1):12

Camley BA, Rappel W-J (2017) Physical models of collective cell motility: from cell to tissue. J Phys D Appl Phys 50(11):113002

Carpenter AE, Kamentsky L, Eliceiri KW (2021) A call for bioimaging software usability. Nature Methods 9(7):666–670

Celli S, Lemaître F, Bousso P (2007) Real-time manipulation of T cell-dendritic cell interactions in vivo reveals the importance of prolonged contacts for CD4+ T cell activation. Immunity 27(4):625–634

Chan TF, Vese LA (2001) Active contours without edges. IEEE Trans Image Process 10(2):266–277

Chang J, Wei D, Fisher JW (2013) A video representation using temporal superpixels. In Proceedings of the IEEE computer society conference on computer vision and pattern recognition, pp. 2051–2058

de Chaumont F, Dallongeville S, Chenouard N, Hervé N, Pop S, Provoost T, Meas-Yedid V, Pankajakshan P, Lecomte T, Le Montagner Y, Lagache T, Dufour A, Olivo-Marin J-C (2012) Icy: an open bioimage informatics platform for extended reproducible research. Nature Methods 9(7):690–696

Cheng Y, Li L, Xu Y, Li X, Yang Z, Wang W, Yang Y (2023) Segment and track anything. Preprint. arXiv:2305.06558

Chenouard N, Bloch I, Olivo-Marin J-C (2013) Multiple hypothesis tracking for cluttered biological image sequences. IEEE Trans Pattern Anal Mach Intell 35(11):2736–3750

Codling EA, Plank MJ, Benhamou S (2008) Random walk models in biology. J R Soc Interface 5(25):813–834

Crainiciuc G, Palomino-Segura M, Molina-Moreno M, Sicilia J, Aragones DG, Li JLY, Madurga R, Adrover JM, Aroca-Crevillén A, Martin-Salamanca S, del Valle AS, Castillo SD, Welch HCE, Soehnlein O, Graupera M, Sánchez-Cabo F, Zarbock A, Smithgall TE, Di Pilato M, Mempel TR, Tharaux P-L, González SF, Ayuso-Sacido A, Ng LG, Calvo GF, González-Díaz I, Díaz-de María F, Hidalgo A (2022) Behavioural immune landscapes of inflammation. Nature 601(7893):415–421

Davis DM (2009) Mechanisms and functions for the duration of intercellular contacts made by lymphocytes. Nature Rev Immunol 9:543–555 (2009)

Delgado-Rodriguez P, Sánchez RM, Rouméas-Noël E, Paris F, Munoz-Barrutia A (2024) Automatic classification of normal and abnormal cell division using deep learning. Sci Rep 14(1):14241

Devriendt T, Shabani M, Borry P (2021) Data sharing in biomedical sciences: a systematic review of incentives. Biopreservation Biobanking 19(3):219–227

Donahue J, Hendricks LA, Rohrbach M, Venugopalan S, Guadarrama S, Saenko K, Darrell T (2017) Long-term recurrent convolutional networks for visual recognition and description. IEEE Trans Pattern Anal Mach Intell 39(4):677–691

Dosovitskiy A, Fischer P, Ilg E, Hausser P, Hazirbas C, Golkov V, Van Der Smagt P, Cremers D, Brox T (2015) Flownet: learning optical flow with convolutional networks. In Proceedings of the IEEE international conference on computer vision, pp. 2758–2766

Drechsler M, Lang LF, Al-Khatib L, Dirks H, Burger M, Schönlieb C-B, Palacios IM (2020) Optical flow analysis reveals that kinesin-mediated advection impacts the orientation of microtubules in the drosophila oocyte. Molecular Biol Cell 31(12):1246–1258

Fürth R (1920) Die brownsche bewegung bei berücksichtigung einer persistenz der bewegungsrichtung. mit anwendungen auf die bewegung lebender infusorien. Z Phys 2(3):244–256

Friedl P, Weigelin B (2008) Interstitial leukocyte migration and immune function. Nature Immunol 9(9):960–969

Girdhar R, Carreira J, Doersch C, Zisserman A (2019) Video action transformer network. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 244–253

Gorelik R, Gautreau A (2014) Quantitative and unbiased analysis of directional persistence in cell migration. Nature Protocols 9:1931–1943

Gröger M, Hirzinger G (2006) Optical flow to analyse stabilised images of the beating heart. In International conference on computer vision theory and applications, vol 3, pp. 237–244

Graner FMC, Glazier JA (1992) Simulation of biological cell sorting using a two-dimensional extended potts model. Phys Rev Lett 69:2013–2016

Guo D, Van de Ven AL, Zhou X (2013) Tracking and measurement of the motion of blood cells using optical flow methods. IEEE J Biomed Health Inf 18(3):991

Hanna RN, Cekic C, Sag D, Tacke R, Thomas GD, Nowyhed H, Herrley E, Rasquinha N, McArdle S, Wu R et al (2015) Patrolling monocytes control tumor metastasis to the lung. Science 350(6263):985–990

He S, Sillah M, Cole AR, Uboveja A, Aird KM, Chen Y-C, Gong Y-N (2024) D-mains: a deep-learning model for the label-free detection of mitosis, apoptosis, interphase, necrosis, and senescence in cancer cells. Cells 13(12):1004

Herath S, Harandi M, Porikli F (2017) Going deeper into action recognition: a survey. Image Vision Comput 60:4–21

Hirashima T, Rens EG, Merks RMH (2017) Cellular potts modeling of complex multicellular behaviors in tissue morphogenesis. Develop Growth Differ 59(5):329–339

Hirsch P, Epstein L, Guignard L (2022) Chapter 20 - mathematical and bioinformatic tools for cell tracking. In: Schnoor M, Yin L-M, Sun SX (eds) Cell movement in health and disease. Academic Press, Cambridge, pp. 341–361

Horn BK, Schunck BG (1981) Determining optical flow. Artif Intell 17(1–3):185–203

Hotelling H (1931) The generalization of student's ratio. Ann Math Stat 2(3):360–378

Huang Y, Hao L, Li H, Liu Z, Wang P (2017) Quantitative analysis of intracellular motility based on optical flow model. J Healthcare Eng 2017(1):1848314

Ibidi, Ibidi application guide: Chemotaxis assays. 2019. Accessed on Jan 10, 2025. https://ibidi.com/img/cms/resources/AG/FL_AG_035_Chemotaxis_150dpi.pdf

Ilg E, Saikia T, Keuper M, Brox T, Occlusions (2018a) motion and depth boundaries with a generic network for disparity, optical flow or scene flow estimation. In Proceedings of the European conference on computer vision (ECCV), pp. 614–630

Ilg E, Çiçek Ö, Galesso S, Klein A, Makansi O, Hutter F, Brox T (2018) Uncertainty estimates and multi-hypotheses networks for optical flow. In European conference on computer vision (ECCV)

Ji S, Xu W, Yang M, Yu K (2013) 3D convolutional neural networks for human action recognition. IEEE Trans Pattern Anal Mach Intell 35(1):221–231

Jiang Z, Rozgic V, Adali S (2017) Learning spatiotemporal features for infrared action recognition with 3D convolutional neural networks. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 309–317

Kalman RE (1960) A new approach to linear filtering and prediction problems. J Basic Eng 82:35–45

Karpathy A, Toderici G, Shetty S, Leung T, Sukthankar R, Fei-Fei L (2014) Large-scale video classification with convolutional neural networks. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1725–1732

Kienle K, Glaser KM, Eickhoff S, Mihlan M, Knöpper K, Reátegui E, Epple MW, Gunzer M, Baumeister R, Tarrant TK et al (2021) Neutrophils self-limit swarming to contain bacterial growth in vivo. Science 372(6548):eabe7729

Krishna NM, Reddy RY, Reddy MSC, Madhav KP, Sudham G (2021) Object detection and tracking using YOLO. In 2021 third international conference on inventive research in computing applications (ICIRCA). IEEE, pp. 1–7

Krummel MF, Bartumeus F, Gérard A (2016) T cell migration, search strategies and mechanisms. Nature Rev Immunol 16:193–201

Lee M, Lee Y-H, Song J, Kim G, Jo Y, Min H, Kim CH, Park Y (2020) Deep-learning-based three-dimensional label-free tracking and analysis of immunological synapses of CAR-T cells. Elife 9:e49023

Letendre K, Donnadieu E, Moses ME, Cannon JL (2015) Bringing statistics up to speed with data in analysis of lymphocyte motility. PLOS ONE 10(5):1–18

Li F, Zhou X, Ma J, Wong STC (2010) Multiple nuclei tracking using integer programming for quantitative cancer cell cycle analysis. IEEE Trans Med Imaging 29(1):96–105

Liu B, Zhu Y, Yang Z, Yan HH, Leung SY, Shi J (2024) Deep learning–based 3D single-cell imaging analysis pipeline enables quantification of cell–cell interaction dynamics in the tumor microenvironment. Cancer Res 84(4):517–526

Lucas BD, Kanade T (1981) An iterative image registration technique with an application to stereo vision. In IJCAI'81: 7th international joint conference on artificial intelligence, vol. 2, pp. 674–679

Maška M, Ulman V, Svoboda D, Matula P, Matula P, Ederra C, Urbiola A, España T, Venkatesan S, Balak DM et al (2014) A benchmark for comparison of cell tracking algorithms. Bioinformatics 30(11):1609–1617

Maška M, Ulman V, Delgado-Rodriguez P, Gómez-de Mariscal E, Nečasová T, Guerrero Peña FA, Ren TI, Meyerowitz EM, Scherr T, Löffler K et al (2023) The cell tracking challenge: 10 years of objective benchmarking. Nature Methods 20(7):1010–1020

Magnusson KEG, Jaldén J (2012) A batch algorithm using iterative application of the Viterbi algorithm to track cells and construct cell lineages. In 2012 9th IEEE international symposium on biomedical imaging (ISBI). IEEE, pp. 382–385

Maiuri P, Rupprecht J-F, Wieser S, Ruprecht V, Bénichou O, Carpi N, Coppey M, De Beco S, Gov N, Heisenberg C-P, Lagev Crespo C, Lautenschlaeger F, Le Berre M, Lennon-Dumenil A-M, Raab M, Thiam H-R, Piel M, Sixt M, Voituriez R (2015) Actin flows mediate a universal coupling between cell speed and cell persistence. Cell 161(2):374–386

Markappa PSS, O'Leary C, Lynch C (2024) A review of YOLO models for soccer-based object detection. In 2024 sixth international conference on intelligent computing in data sciences (ICDS), pp. 1–7

Masters BR (2008) History of the optical microscope in cell biology and medicine. John Wiley & Sons, New York

McCutcheon M (1946) Chemotaxis in leukocytes. Physiol Rev 26(3):319–336. PMID: 20993553

McInnes L, Healy J, Melville J (2018) UMAP: Uniform manifold approximation and projection for dimension reduction. arXiv:1802.03426

Pizzagalli DU, Carrillo-Barbera P, Palladino E, Ceni K, Thelen B, Pulfer A, Moscatello E, Cabini RF, Textor J, Wortel I et al (2024) Systematic analysis of immune cell motility leveraging Immunemap, an open intravital microscopy atlas. bioRxiv

Mempel TR, Henrickson SE, von Andrian UH (2004) T-cell priming by dendritic cells in lymph nodes occurs in three distinct phases. Nature 427:154–159

Milan A, Rezatofighi SH, Dick A, Reid I, Schindler K (2017) Online multi-target tracking using recurrent neural networks. In Proceedings of the AAAI conference on Artificial Intelligence, vol. 31.

Miller MJ, Wei SH, Parker I, Cahalan MD (2002) Two-photon imaging of lymphocyte motility and antigen response in intact lymph node. Science 296(5574):1869–1873
Miller MJ, Wei SH, Cahalan MD, Parker I (2003) Autonomous T cell trafficking examined in vivo with intravital two-photon microscopy. Proc Natl Acad Sci 100(5):2604–2609
Moen E, Bannon D, Kudo T, Graf W, Covert M, Van Valen D (2019) Deep learning for cellular image analysis. Nature Methods 16(12):1233–1246
Mokhtari Z, Mech F, Zitzmann C, Hasenberg M, Gunzer M, Figge MT (2013) Automated characterization and parameter-free classification of cell tracks based on local migration behavior. PLOS ONE 8:1–20
Montagud A, de Leon MP, Valencia A (2021) Systems biology at the giga-scale: large multiscale models of complex, heterogeneous multicellular systems. Current Opinion Syst Biol 28:100385
Moore BR (1980) A modification of the Rayleigh test for vector data. Biometrika 67(1):175–180
National Institute for Health Research (2025) A collection of open source imaging data sets. Accessed on Mar 17, 2022
Nicolai M, Cabini RF, Pizzagalli DU (2024) Classification and regression of trajectories rendered as images via 2D convolutional neural networks. Preprint. arXiv:2409.18832
Niculescu I, Textor J, de Boer RJ (2015) Crawling and gliding: a computational model for shape-driven cell migration. PLOS Comput Biol 11(10):1–22
Oja H, Randles RH (2004) Multivariate nonparametric tests. Stat Sci 19(4):598–605
Pécot T, Zengzhen L, Boulanger J, Salamero J, Kervrann C (2018) A quantitative approach for analyzing the spatio-temporal distribution of 3D intracellular events in fluorescence microscopy. Elife 7:e32311
Perš J, Sulić V, Kristan M, Perše M, Polanec K, Kovačič S (2010) Histograms of optical flow for efficient representation of body motion. Pattern Recogn Lett 31(11):1369–1376
Pizzagalli DU, Farsakoglu Y, Palomino-Segura M, Palladino E, Sintes J, Marangoni F, Mempel TR, Koh WH, Murooka TT, Thelen F et al (2018) Leukocyte tracking database, a collection of immune cell tracks from intravital 2-photon microscopy videos. Sci Data 5(1):1–13
Pizzagalli DU, Gonzalez SF, Krause R (2019a) A trainable clustering algorithm based on shortest paths from density peaks. Sci Adv 5(10):eaax3770
Pizzagalli DU, Latino I, Pulfer A, Palomino-Segura M, Virgilio T, Farsakoglu Y, Krause R, Gonzalez SF (2019b) Characterization of the dynamic behavior of neutrophils following influenza vaccination. Front Immunol 10, 2621
Pizzagalli DU, Bordini J, Morone D, Pulfer A, Carrillo-Barberà P, Thelen B, Ceni K, Thelen M, Krause R, Gonzalez SF (2022a) Cancol, a computer-assisted annotation tool to facilitate colocalization and tracking of immune cells in intravital microscopy. J Immunol 208(6):1493–1499
Pizzagalli DU, Pulfer A, Thelen M, Krause R, Gonzalez SF (2022b) In vivo motility patterns displayed by immune cells under inflammatory conditions. Front Immunol 12:804159
Pulfer A, Pizzagalli DU, Gagliardi PA, Hinderling L, Lopez P, Zayats R, Carrillo-Barberà P, Antonello P, Palomino-Segura M, Grädel B, et al (2024) Transformer-based spatial–temporal detection of apoptotic cell death in live-cell imaging. Elife 12:RP90502
Rajič F, Ke L, Tai Y-W, Tang C-K, Danelljan M, Yu F (2023) Segment anything meets point tracking. Preprint. arXiv:2307.01197
Revaud J, Weinzaepfel P, Harchaoui Z, Schmid C (2015) Epicflow: edge-preserving interpolation of correspondences for optical flow. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1164–1172
Rodrigues V, Rodrigues PJ, Pereira AI, Lima R (2014) Automatic cell tracking method based on optical flow tuned by template matching. Virtual. ipb. pt
Rosen ME, Grant CP, Dallon JC (2021) Mean square displacement for a discrete centroid model of cell motion. PLOS ONE 16(12):1–19
Sánchez Pérez, J, Meinhardt-Llopis E, Facciolo G (2013) TV-L1 optical flow estimation. Image Processing On Line 3:137–150. https://doi.org/10.5201/ipol.2013.26
Schienstock D, Mueller SN (2022) Moving beyond velocity: opportunities and challenges to quantify immune cell behavior. Immunol Rev 306(1):123–136

Schienstock D, Hor JL, Devi S, Mueller SN (2024) Cecelia: a multifunctional image analysis toolbox for decoding spatial cellular interactions and behaviour. bioRxiv, pp. 2024–08

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez J-Y, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A (2012) Fiji: an open-source platform for biological-image analysis. Nature Methods 9(7):676–682

Schmidt U, Weigert M, Broaddus C, Myers G (2018) Cell detection with star-convex polygons. In Medical image computing and computer assisted intervention–MICCAI 2018: 21st international conference, Granada, Spain, September 16–20, 2018, Proceedings, Part II 11. Springer, pp. 265–273

Shah STH, Xuezhi X (2021) Traditional and modern strategies for optical flow: an investigation. SN Appl Sci 3(3):289

Simonyan K, Zisserman A (2014) Two-stream convolutional networks for action recognition in videos. Adv Neural Inf Process Syst 27

Singer AJ, Clark RA (1999) Cutaneous wound healing. New England J Med 341(10):738–746

Stettler MD (2025) Histopathology datasets for machine learning. Accessed on Mar 17, 2025. https://github.com/maduc7/Histopathology-Datasets

Stringer C, Wang T, Michaelos M, Pachitariu M (2021) Cellpose: a generalist algorithm for cellular segmentation. Nature Methods 18(1):100–106

Sun L, Jia K, Yeung D-Y, Shi BE (2015) Human action recognition using factorized spatio-temporal convolutional networks. In 2015 IEEE international conference on computer vision (ICCV), pp. 4597–4605

Szabó B, Szöllösi GJ, Gönci B, Jurányi Z, Selmeczi D, Vicsek T (2006) Phase transition in the collective migration of tissue cells: experiment and model. Phys Rev E 74:061908

Szabó A, Merks RMH (2013) Cellular Potts modeling of tumor growth, tumor invasion, and tumor evolution. Front Oncol 3

Textor J, Peixoto A, Henrickson SE, Sinn M, von Andrian UH, Westermann J (2011) Defining the quantitative limits of intravital two-photon lymphocyte tracking. Proc Natl Acad Sci 108(30):12401–12406

Textor J, Sinn M, de Boer RJ (2013) Analytical results on the Beauchemin model of lymphocyte migration. BMC Bioinf 14:S10

Textor J, Berry J, Miller MJ (2015) Motilitylab – a resource for cell migration analysis. Accessed on Jan 26, 2025

Thomas GL, Fortuna I, Perrone GC, Glazier JA, Belmonte JM, de Almeida RM (2020) Parameterizing cell movement when the instantaneous cell migration velocity is ill-defined. Phys A Stat Mech Appl 550:124493

Tinevez J-Y, Perry N, Schindelin J, Hoopes GM, Reynolds GD, Laplantine E, Bednarek SY, Shorte SL, Eliceiri KW (2017) Trackmate: an open and extensible platform for single-particle tracking. Methods 115:80–90

Tran D, Bourdev LD, Fergus R, Torresani L, Paluri M (2014) Learning spatiotemporal features with 3D convolutional networks. 2015 IEEE international conference on computer vision (ICCV), pp. 4489–4497

Tu Z, Xie W, Qin Q, Poppe R, Veltkamp RC, Li B, Yuan J (2018) Multi-stream CNN: learning representations based on human-related regions for action recognition. Pattern Recogn 79:32–43

Turetken E, Wang X, Becker CJ, Haubold C, Fua P (2016) Network flow integer programming to track elliptical cells in time-lapse sequences. IEEE Trans Med Imaging 36(4):942–951

Uijlings J, Duta IC, Sangineto E, Sebe N (2015) Video classification with densely extracted hog/hof/mbh features: an evaluation of the accuracy/computational efficiency trade-off. Int J Multimedia Inf Retrieval 4(1):33–44

Ulhaq A, Akhtar N, Pogrebna G, Mian A (2022) Vision transformers for action recognition: a survey. Preprint. arXiv:2209.05700

Ulman V, Maška M, Magnusson KE, Ronneberger O, Haubold C, Harder N, Matula P, Matula P, Svoboda D, Radojevic M et al (2017) An objective comparison of cell-tracking algorithms. Nature Methods 14(12):1141–1152
Vicsek T, Zafeiris A (2012) Collective motion. Phys Rep 517(3):71–140
Vishwakarma S, Agrawal A (2013) A survey on activity recognition and behavior understanding in video surveillance. Visual Comput 29:983–1009
Vrigkas M, Nikou C, Kakadiaris IA (2015) A review of human activity recognition methods. Front Robotics AI 2:28
Wagner T, Kroll A, Haramagatti CR, Lipinski H-G, Wiemann M (2017) Classification and segmentation of nanoparticle diffusion trajectories in cellular micro environments. PloS One 12(1):e0170165
Waite JC, Leiner I, Lauer P, Rae CS, Barbet G, Zheng H, Portnoy DA, Pamer EG, Dustin ML (2011) Dynamic imaging of the effector immune response to listeria infection in vivo. PLoS Pathogens 7(3):e1001326
Wang T, Snoussi H (2013) Histograms of optical flow orientation for abnormal events detection. In 2013 IEEE international workshop on performance evaluation of tracking and surveillance (PETS). IEEE, pp. 45–52
Wang H, Kläser A, Schmid C, Liu C-L (2013) Dense trajectories and motion boundary descriptors for action recognition. Int J Comput Vision 103:60–79
Weinzaepfel P, Revaud J, Harchaoui Z, Schmid C (2013) Deepflow: large displacement optical flow with deep matching. In Proceedings of the IEEE international conference on computer vision, pp. 1385–1392
Wortel IM, Liu AY, Dannenberg K, Berry JC, Miller MJ, Textor J (2021) CelltrackR: an R package for fast and flexible analysis of immune cell migration data. ImmunoInformatics 1–2:100003
Wortel IM, Postat J, Mihaylova M, Merino M, Bhagrath A, Harris M, Wouters L, Wiebke L, Parisi DR, Mandl JN, Textor J (2024) Cooperative motility emerges in crowds of T cells but not neutrophils. bioRxiv
Wu S, Moore BE, Shah M (2010) Chaotic invariants of Lagrangian particle trajectories for anomaly detection in crowded scenes. In 2010 IEEE computer society conference on computer vision and pattern recognition. IEEE, pp. 2054–2060
Zach C, Pock T, Bischof H (2007) A duality based approach for realtime tv-l 1 optical flow. In Pattern recognition: 29th DAGM symposium, Heidelberg, Germany, September 12–14, 2007. Proceedings 29. Springer, pp. 214–223
Zhu Y, Li X, Liu C, Zolfaghari M, Xiong Y, Wu C, Zhang Z, Tighe J, Manmatha R, Li M (2020) A comprehensive study of deep video action recognition. Preprint. arXiv:2012.06567
Zuraimi MAB, Zaman FHK (2021) Vehicle detection and tracking using YOLO and DeepSORT. In 2021 IEEE 11th IEEE symposium on computer applications & industrial electronics (ISCAIE). IEEE, pp. 23–29