nature methods

Article

https://doi.org/10.1038/s41592-025-02665-8

Deep learning-assisted analysis of single-particle tracking for automated correlation between diffusion and function

Received: 6 December 2023

Accepted: 6 March 2025

Jacob Kæstel-Hansen [1,2,3,4], **Marilina de Sautu** [5,6], **Anand Saminathan** [7,8,9], **Gustavo Scanavachi** [7,8,9], **Ricardo F. Bango Da Cunha Correia** [7,8,9], **Annette Juma Nielsen** [1,2,3,4], **Sara Vogt Bleshøy** [1,2,3,4], **Konstantinos Tsolakidis** [1,2,3,4], **Wouter Boomsma** [10], **Tomas Kirchhausen** [3,7,8,9] & **Nikos S. Hatzakis** [1,2,3,4]

Subcellular diffusion in living systems reflects cellular processes and interactions. Recent advances in optical microscopy allow the tracking of this nanoscale diffusion of individual objects with unprecedented precision. However, the agnostic and automated extraction of functional information from the diffusion of molecules and organelles within the subcellular environment is labor intensive and poses a significant challenge. Here we introduce DeepSPT, a deep learning framework integrated in an analysis software, to interpret the diffusional two- or three-dimensional temporal behavior of objects in a rapid and efficient manner, agnostically. Demonstrating its versatility, we have applied DeepSPT to automated mapping of the early events of viral infections, identifying endosomal organelles, clathrin-coated pits and vesicles among others with F1 scores of 81%, 82% and 95%, respectively, and within seconds instead of weeks. The fact that DeepSPT effectively extracts biological information from diffusion alone illustrates that besides structure, motion encodes function at the molecular and subcellular level.

The direct observation of cellular processes is routinely achieved by fluorescence microscopy and single-particle tracking (SPT) techniques[1–6]. These techniques offer the necessary spatiotemporal resolution to localize and track diffusion of individual biomolecules—from small proteins and viruses to organelles or entire cells—in both two-dimensional (2D) and three-dimensional (3D) environments[2,3,7–9]. The observed diffusion is highly complex and exhibits considerable spatiotemporal and interparticle heterogeneity, reflecting various biological factors such as internalization stages, local environment, oligomerization states and interactions with elements such as the cytoskeleton, membranes, molecular motors, organelles and more[2,8,10]. The robust analysis of this intrinsic heterogeneity is essential for understanding the underlying biophysical processes but presents a substantial challenge and remains a major bottleneck for extracting quantitative insights from single-particle studies.

Extracting diffusional behavior from SPT experiments has relied on fitting the mean squared displacement (MSD)[11,12]. These approaches often transform entire trajectories into a single descriptor such as the diffusion coefficient, thus averaging out vital temporal information that is essential to interpret biological processes. Recent advances by

---

[1]Department of Chemistry, University of Copenhagen, Copenhagen, Denmark. [2]Center for 4D cellular dynamics, University of Copenhagen, Copenhagen, Denmark. [3]Novo Nordisk Center for Optimised Oligo Escape, University of Copenhagen, Copenhagen, Denmark. [4]Novo Nordisk foundation Center for Protein Research, University of Copenhagen, Copenhagen, Denmark. [5]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA, USA. [6]Laboratory of Molecular Medicine, Boston Children's Hospital, Boston, MA, USA. [7]Department of Cell Biology, Harvard Medical School, Cambridge, MA, USA. [8]Department of Pediatrics, Harvard Medical School, Cambridge, MA, USA. [9]Program in Cellular and Molecular Medicine Boston Children's Hospital, Boston, MA, USA. [10]Department of Computer Science, University of Copenhagen, Copenhagen, Denmark.
e-mail: kirchhau@crystal.harvard.edu; hatzakis@chem.ku.dk

us[13] and others[14] have introduced the feature extraction and the concept of diffusional fingerprinting[13] to analyze heterogeneous behavior, albeit the methods offer no temporal segmentation. Achieving temporal segmentation—a prerequisite for unlocking the rich temporal data inherent in biological processes—is a considerable challenge. For instance, manual annotation requires significant expertise and is prohibitively time consuming for large datasets, especially in 3D. Methods such as Rolling MSD[11,12,15] and divide-and-conquer[16] offer automated temporal segmentation, but they are reliant on windowing tracks, which introduces a trade-off in temporal sensitivity and accuracy and they also depend on user-defined, system-specific parameters. Hidden Markov models (HMMs)[17–19] can segment traces but only if a diffusional metric, often step length, varies significantly between states[19,20]. Over the past years, additional SPT analysis techniques have been developed such as state-transition analysis (MC-DDA, anaDDA and SMAUG)[21–23], among other analytical tools such as ExTrack[24], vbSPT[18], Momboisse et al.[25], Spot-ON[26] and TARDIS[27].

The challenge of accurate analysis of diffusion resulted in the 2021 anomalous diffusion (AnDi) challenge[28] that, in agreement with other work[20,29–31], established that current state-of-the-art in temporal segmentation is based on machine learning[13,14,20,32–34]. This aligns with a broader trend of machine learning proving exceptionally powerful for multiple tasks in biology, including protein structure prediction[35], bioimage analysis[36–39], genome engineering[40] and drug discovery[41]. The strength of machine learning in biology stems from its ability to autonomously learn meaningful feature representations tailored to optimize performance on a specific task directly from high-dimensional, noisy data. With minimal human intervention, it leverages subtle regularities, domain knowledge and nonlinear relationships often inaccessible for traditional methods. Current machine learning approaches to temporal segmentation of diffusion include a plethora of models ranging from utilizing random forests on sliding windows to end-to-end deep recurrent neural networks[20,28–31,34]. While these tools operate well for their specific sets of states with distinct, user-defined diffusional characteristics, they remain less explored for systems displaying an arbitrary number of states with broadly distributed diffusional characteristics commonly found in complex cellular environments. In addition, only a few of the current methods extend to 3D trajectories[28].

Achieving temporal segmentation of heterogeneous diffusion is crucial for overcoming the current analytical bottleneck in SPT; however, decoding correlations between heterogeneous behavior and biomolecular identity, colocalizing partners, cellular localization or the time point of a biological event may rely on more subtle feature relations. While the above toolboxes, independently of being machine learning based or not, can offer segmentation of diffusional behavior, they are not designed to correlate biological motion to biological function. Currently, identifying such biological context is a challenge in fluorescence microscopy, necessitating specialized analysis, parallelized multicolor and often super-resolution imaging[37,38,42]. Such experimental design, specialized analysis and fluorescent tagging of biomolecular entities is substantially labor and material intensive and risks impairing biological function[43]. These challenges are further compounded by the limitation of using two to three imaging channels for quantitative imaging due to spectral overlap.

Temporal analysis of diffusional behavior could overcome these challenges by acting as an orthogonal probe for extracting biological function, colocalization or identity, minimizing the need for fluorescent tagging, thereby simplifying experimental workflows. However, so far, dissecting heterogeneous diffusional behavior to inform on such biological context remains largely untapped.

Here, we introduce DeepSPT, a versatile deep learning-based toolbox designed for the rapid, accurate and automatic temporal analysis of behavior in SPT. DeepSPT facilitates extraction of biological insights from 2D or 3D traces solely based on the diffusional characteristics of the tracked objects, as demonstrated by capturing endosomal identity, colocalization partners, cellular localization of clathrin-coated pits (CCPs) and detection of the time point for viral escape to cytosol. The pretrained DeepSPT pipeline is available as open-source code on GitHub. To ensure the facile operation of DeepSPT for a broader audience, we provide a standalone executable for both Mac and Windows. The simple and intuitive graphical user interface (GUI) allows users to execute every core function of DeepSPT: segmentation, diffusional fingerprinting and training a task-specific classifier to predict biological information as well as outputting publication-quality figures.

Results

DeepSPT

DeepSPT is a deep learning framework, encompassing three sequentially connected modules: a temporal behavior segmentation module, a diffusional fingerprinting module and a task-specific downstream classifier module (Fig. 1a). As input, DeepSPT takes the output of any particle tracker: a set of x, y and z localizations over time, yielding a dataset of trajectories (Methods). Users can use any individual module or the full pipeline. The first two modules can be applied directly to any trajectory dataset characterized by x, y, (z) and time (t) coordinates across diverse biological systems. The final module capitalizes on experimental data to learn a task that is specific to the system under investigation.

The temporal behavior segmentation module consists of an ensemble of three pretrained, uncertainty calibrated U-Nets[44] adapted to accept 2D or 3D single-particle trajectories using two- or three-channel 1D convolutions (Methods). Thereby, the temporal segmentation module transforms single-particle trajectories into subsegments characterized by distinct diffusional behaviors, processing input directly from x, y, (z) and t coordinates using an ensemble of fully convolutional networks (Methods). Alongside the predicted diffusional behavior, each time point in the trajectory is assigned a probability estimate for each type of diffusion identified. This study focuses on the diffusional behaviors predominantly reported in biological system[12,18,31,45]: (1) normal diffusion, typifying unhindered random motion; (2) directed motion, as commonly exhibited by molecular motors; (3) confined motion, characterizing limited spaces with reflective boundaries, such as small membranes structures; and (4) subdiffusive motion, indicative of more restrained movement, commonly observed in densely populated cytosolic environments. We also extend DeepSPT to other diffusional behaviors reported in literature, such as the AnDi challenge[28]. The module's training utilized an extensive dataset comprising 900,000 trajectories, exhibiting broadly distributed diffusional properties, this encompasses variations spanning four orders of magnitude in diffusional coefficients, diverse trace durations, varying localization errors and trajectories displaying multiple, random length diffusional behaviors throughout their lifespan (Methods). This extensive training set expands the adaptability of DeepSPT across different biological systems and experimental conditions. It is important to note that DeepSPT can be trained to recognize other diffusional attributes and diverse motion types, or to simply predict a uniform global diffusional state in cases of homogeneous motion.

The diffusional fingerprinting module transforms each identified segment of diffusional behavior into a comprehensive set of 40 descriptive diffusional features, not just encompassing the 17 features enunciated by our work in ref. 13 but expanding the feature set to include temporal features[13] (Supplementary Table 1). Reference 13 and the diffusional fingerprinting module within DeepSPT are tools to analyze heterogeneous behavior albeit they do not offer segmentation. The diffusional fingerprinting module of DeepSPT serves a dual purpose: it facilitates quantification of individual behavior segments for user interpretation and it generates feature representations crucial for downstream classification tasks (Methods).

The task-specific downstream classification module trains and predicts directly on experimental data, which has been transformed to

https://doi.org/10.1038/s41592-025-02665-8

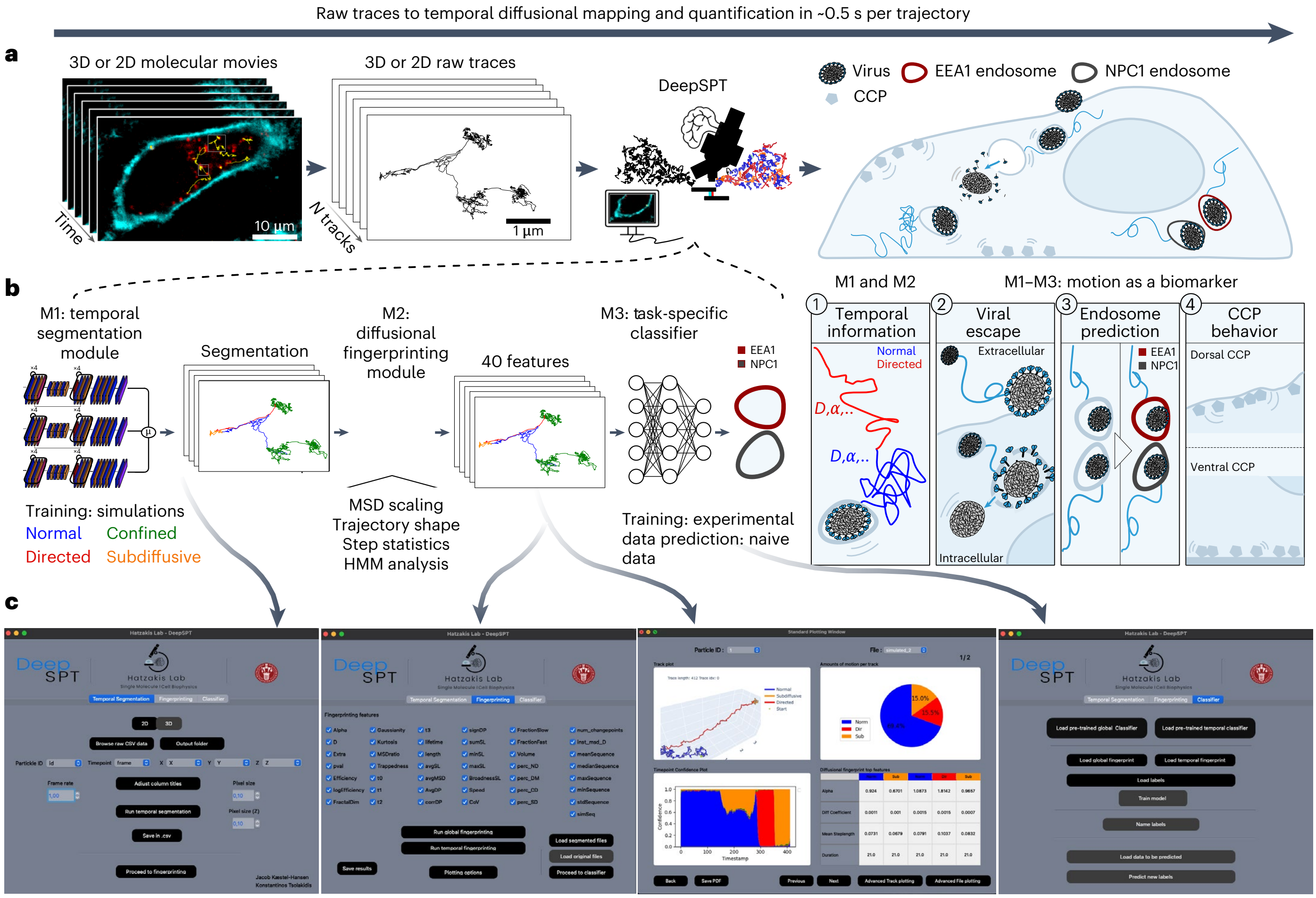


Fig. 1 | DeepSPT, an agnostic, automated approach for extraction of time-dependent behavior in dynamic systems. a, A schematic representation of the DeepSPT pipeline: 2D or 3D molecular movies from fluorescence microscopy imaging produce a set of *x*, *y*, (*z*) and *t* localizations for each particle, yielding a dataset of single-particle trajectories. These trajectories are directly fed to DeepSPT, consisting of a temporal behavior segmentation module (M1), diffusional fingerprinting module (M2) and a task-specific classifier (M3). The modules of DeepSPT appear in the zoom in. First, the temporal segmentation module classifies, per time point, the diffusional behavior adaptable to any behavior (in this case, normal, directed, confined or subdiffusive). Second, tracks segmented into diffusional behaviors are quantified by multiple diffusional descriptors by the diffusional fingerprinting module. Third, a task-specific classifier is trained utilizing the temporal information and the diffusional fingerprints for each track to learn a problem of interest, for example, identification of endosomal identity based on diffusional behavior of cargo. The entire DeepSPT pipeline has a computational time of ~500 ms per trajectory. **b**, A schematic illustration of selected biological applications enabled by the DeepSPT pipeline. (1) Temporal diffusional behavior segmentation, analysis and quantification. Applications of DeepSPT to uncover biological insights, based exclusively on diffusional behavior variation: (2) time point identification of biological events such as detection of viral escape into the cytosol; (3) prediction of endosomal identity directly using endosomal motion or solely from movement of their cargo and (4) predicting cellular localization of CCPs. **c**, Screenshots of the standalone, free-to-use, GUI that integrates the multimodular pipeline of DeepSPT. The user-friendly GUI allows users to perform all core functions of DeepSPT directly on raw traces, including segmentation, diffusional fingerprinting, live plotting and evaluation of data and training of the classifier, as well as outputting publication-quality figures.

a feature set by the temporal and diffusional fingerprinting modules. This module outputs class probability estimates solely utilizing diffusional characteristics for any domain. This is exemplified by predicting important time points during the initial phases of rotavirus infection, differentiating early and late endosomes, and by locating CCPs and vesicles to the dorsal or ventral membranes of a cell (Fig. 1b).

Rapid, automated analysis of temporal diffusional behavior

To demonstrate the effectiveness and generalizability of the temporal segmentation capabilities of DeepSPT, we employed five distinct evaluation schemes. First, we used a holdout scheme to assess performance on trajectories withheld during training (Fig. 2a–c). Second, we tested the model's generalizability using simulated trajectories with a wider distribution in the values of diffusion parameters than those used during training. Third, we compared the performance of DeepSPT against existing state-of-the-art temporal segmentation algorithms. Fourth, we benchmarked DeepSPT to two deep learning models on the five different diffusion behaviors from the AnDi challenge[28]. Last, we investigated the ability of DeepSPT to classify classes of behavior beyond diffusional behaviors.

In the holdout validation, we assessed the temporal segmentation on a test set comprising 20,000 simulated trajectories, of which 80% exhibited heterogeneous motion and 20% showed homogeneous motion (Methods). The trajectories spanned a broad range of diffusional parameters and four motion types (Fig. 2a, Methods and Supplementary Figs. 1 and 2). DeepSPT not only accurately identified temporal change points (Fig. 2a), but also yielded time-resolved probability estimates that may serve as an adjustable postprocessing parameter (DeepSPT output versus time; Fig. 2a). We calibrated these probability estimates using temperature scaling[46] (Supplementary Fig. 3) to enhance reliability and mitigate overconfidence. The fact that short segments of trajectories in Fig. 2a of Brownian motion may

https://doi.org/10.1038/s41592-025-02665-8

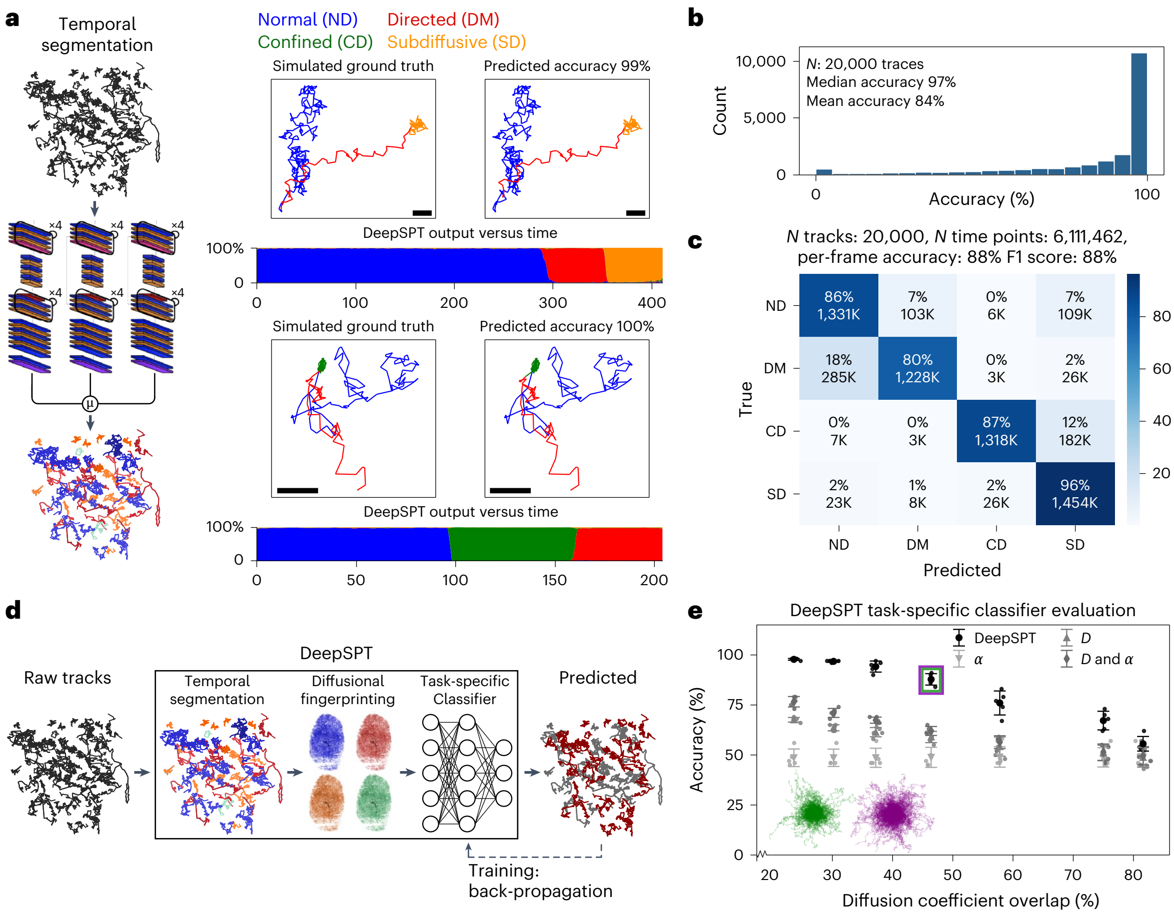


Fig. 2 | Evaluation of DeepSPT's temporal behavior segmentation. **a**, An illustration of the temporal segmentation module of DeepSPT showing two examples of DeepSPT prediction on simulated 3D trajectories with heterogeneous diffusion. Left: a 2D projection of the simulated ground truth, color coded to underlying diffusional behavior. Right: the trajectory color coded to DeepSPT's predictions. Scale bar 500 nm. Bottom: uncertainty calibrated probability estimates (DeepSPT output versus time) for each modeled diffusion type per time point, providing transparency into model certainty (see Supplementary Figs. 2 and 3 for additional examples and uncertainty calibration). **b**, A histogram of the accuracies associated with each individual 3D trajectory in the test set (*N* traces, 20,000; *N* time points, 6,111,462; Methods). **c**, A confusion matrix based on all predictions (*N* time points, 6,111,462) within the 20,000 test set trajectories in **b** totaling >6M individual time point predictions. The diagonal entries are correct predictions and off-diagonal indicates confused classes. Each entry reports the absolute number of predictions ($K = 1{,}000$) and normalization to the number of labels in the given class. **d**, An illustration of the DeepSPT classification pipeline. Each track is temporally segmented to each of the four diffusional behaviors by the segmentation module, transformed into descriptive features by the diffusional fingerprinting module, which combines to a unique feature set of temporal and diffusional features and which is subsequently is fed to a task-specific downstream classifier. **e**, Benchmarking of the DeepSPT classification pipeline against a classifier using MSD features: diffusion coefficient (D), the anomalous diffusion exponent term (α) or both (D and α) on simulated data of two classes of trajectories with diffusional properties with overlapping distributions (Methods). Classification accuracy is evaluated at incrementing degrees of overlap in the instantaneous diffusion coefficients. Purple and green trajectories depict trajectories at ~45% overlap in diffusion coefficients indicated by the purple and green boxes. Error bars depict s.d. DeepSPT significantly outperforms all three MSD feature-based approaches up to 75% overlap in diffusion coefficient (all P values <0.001 using a two-sided Welch's t-test, $N = 5$ per condition; Supplementary Table 2) and at 82% overlap DeepSPT significantly outperforms D and α (all P values <0.05; Supplementary Table 2).

resemble directed motion highlights the challenges associated with precise segmentation (Supplementary Fig. 4).

Quantification of DeepSPT's classification performance revealed a median accuracy of 96% per trace and 84% mean accuracy per frame for all four motion types (Fig. 2b,c). The model achieved 91% mean accuracy for three motion types—normal, directed and confined/subdiffusive—and 97% for two motion types normal/directed versus confined/subdiffusive (Extended Data Fig. 1), and 91% for homogeneous motion (Supplementary Fig. 5). DeepSPT achieved an F1 value of 88% for both 3D and 2D datasets (Fig. 2c and Extended Data Fig. 1). In all cases, it has an inference time of less than 40 ms per trajectory. Subdiffusive motion was classified with 96% accuracy, directed motion with 80%, normal and confined motion with 86% and 87%, respectively (Fig. 2c). Minimal confusion existed between dissimilar motion types, highlighting the capability of DeepSPT to differentiate between restricted and free motion types (Fig. 2c and Extended Data Fig. 1). Increasing the size of confinement increases model confusion and reduces accuracy, but DeepSPT retains above 91% median accuracy (Supplementary Fig. 6). However, the strength in differentiating dissimilar behavior types become more apparent when the model was tasked with identifying fewer motion categories (Extended Data Fig. 1). The robust performance of DeepSPT was further confirmed across a variety of diffusional properties, state transitions rates, track durations, tracking errors and localization errors, even for parameter ranges not included in the training set. DeepSPT excelled for traces longer than 20 frames, localization errors equal or smaller than the actual diffusional step lengths and on trajectories with an abundance of tracking errors (Extended Data Figs. 2–4 and Supplementary Fig. 7), demonstrating its adaptability and robustness to various experimental setups and that optimized imaging is essential for accurate segmentation and precise outputting of diffusional metrics.

We benchmarked the ability of DeepSPT to segment heterogeneous diffusion with four diffusional behaviors against a high-performing long short-term memory (LSTM)-based method[29] and the widely used rolling MSD approach[11,12,15,47] (Extended Data Fig. 1). HMM[17–19,48] focused on methods that do not require large variations in step lengths to detect changes in diffusional behavior (Supplementary Figs. 8 and 9). Each method was tested on 2D trajectories, given the LSTM-based technique's 2D limitation. The LSTM-based method achieved classification accuracies of 44%, 58% and 72%, outperforming the rolling MSD's 34%, 51% and 65% for four, three and two diffusional behaviors, respectively. DeepSPT, in contrast, attained accuracies of 88%, 91% and 97% for the same categories, outperforming current state-of-the-art and highlighting its improved competence in agnostic segmentation and classification of heterogeneous diffusion.

In addition, we benchmark DeepSPT against the two best-performing models featured in the 2021 AnDi challenge[28]. These were the deep learning methods E and J. Method E uses a recurrent neural network combined with fully connected layers networks while method J is based on convolutional neural networks. The AnDi challenge investigates five behaviors: annealed transient time motion, continuous-time random walk, fractional Brownian motion, Lévy walk and scaled Brownian motion. DeepSPT was retrained on these diffusion behaviors (Methods) and tested on two sets of data. Benchmarking on data with constant length and a single change point (AnDi task 3 (ref. 28); Methods) for 2D traces resulted in median accuracies of 79% and 80% for methods E and J, while DeepSPT obtained 94%. For 3D trajectories, method J is not applicable, method E obtains a median accuracy of 68%, while DeepSPT achieves 98% (Supplementary Fig. 10). Benchmarking on heterogeneous trajectories with constant track duration but multiple change points between the five AnDi diffusion behaviors (Methods) for 2D traces resulted in 64% and 59% median accuracy for methods E and J, respectively, with DeepSPT achieving 80%. For 3D trajectories, method E obtains 66% median accuracy as compared with 82% for DeepSPT (Supplementary Fig. 11). DeepSPT's improved segmentation accuracy stems from its design and training to segment an arbitrary number of change points for random trajectory lengths, while the AnDi challenge focused on modeling a single change point for fixed-length tracks. DeepSPT's ease of adaptability and improved accuracy for both AnDi task 3 and the heterogeneous diffusion highlights its strength in agnostic segmentation of heterogeneous diffusion and extendibility to multiple systems (additional metrics in Extended Data Figs. 5 and 6).

We then qualitatively evaluated DeepSPT on 2D experimental datasets (Supplementary Figs. 12 and 13). Specifically, for human insulin (HI), we labeled it with Atto-655 and recorded its spatiotemporal localization in HeLa cells using 2D live-cell spinning disk confocal fluorescence microscopy (Methods). Using DeepSPT, we report insulin intracellular transport mainly exhibited subdiffusive behavior but included segments of directed motion. The directed motion aligns with motor-protein diffusion patterns indicative of active cellular trafficking, establishing DeepSPT as a potential tool for studying transport mechanisms across diverse experimental contexts.

The classification of biomolecular identity requires an addition to temporal segmentation. We combine all modules of DeepSPT to demonstrate its capabilities to leverage subtle diffusional variations to classify heterogeneous behavior. The integration of DeepSPT's segmentation and fingerprinting modules allows for the transformation of any trajectory into a feature representation containing both temporal and diffusional features, which can then be fed to a downstream classifier (Fig. 2d). To demonstrate the descriptive power of DeepSPT's integrated approach, we assessed the classification performance of DeepSPT on two classes of simulated trajectories (1,000 tracks) with diffusional properties with overlapping distributions, that is, similarly distributed parameters underlie the diffusion of each population (Methods). Keeping all diffusional features except the diffusion coefficient constant, we evaluated the classification accuracy by stratified fivefold cross-validation for varying degrees of overlap in diffusion coefficients between the classes (Methods). DeepSPT achieved up to 98% accuracy and maintained 76% accuracy even when the overlap in diffusion coefficients was around 57%. This performance significantly (by Welch's *t*-test; Fig. 2e) outperformed that of basic MSD features, which attained accuracies ranging from approximately 49% to 76% (Fig. 2e). Comparison of the change point prediction accuracies to five benchmark approaches on trajectories switching once between simulated trajectories with approximately 75% diffusional overlap (Fig. 2e), revealed that DeepSPT significantly outperforms the benchmarks (Extended Data Fig. 7), highlighting the ability of DeepSPT to discern and exploit subtle differences in diffusional properties.

Accelerated detection of viral escape using motion

We validated the operational efficacy of DeepSPT to extract information from 3D live-cell SPT data of rotavirus (Fig. 3a and Methods). The entry process of rotavirus into cells involves glycolipid-mediated membrane association of the virus, vesicular engulfment and internalization, virus-initiated membrane permeabilization, calcium-dependent uncoating of outer proteins, membrane disruption and cytosolic delivery of the viral genome for subsequent RNA production[49] (Fig. 3a). Previous SPT in BSC-1 cells using confocal imaging indicated that the uncoating step correlates with a change in diffusional behavior, hinting at motion as a potential marker of biological behavior[49,50].

To test the capacity of DeepSPT to detect the uncoating and cytosolic delivery using only motion, we used 3D live-cell lattice light-sheet microscopy[3] (LLSM) to image the cell entry of a reconstituted rotavirus[49,51] labeled with either Atto565 on VP7, an outer shell protein, and Atto647 on the double-layered particle (DLP), or with Atto565 on the entire virus including both VP7 and DLP (Fig. 3a,b). Trajectories captured via LLSM (Fig. 3b) and output using the widely used particle tracker u-track[9] underwent temporal segmentation and diffusional fingerprinting using DeepSPT's modules in rolling windows for sequential representation (Methods). Optimized imaging conditions allowed localization error to be well below the step length (Supplementary Fig. 14). These processed trajectories were then classified as either 'before uncoat' or 'after uncoat' through a sequence-to-sequence based model, transforming coordinates in time into time-resolved predictions, serving as the task-specific classifier (Methods and Supplementary Videos 1 and 2). The ground truth for uncoating events for dual-labeled rotavirus was established by determining the extent of colocalization of differentially labeled DLP and VP7 (Methods). A representative example of rotavirus uncoating alongside DeepSPT's consistent prediction and snapshots of the raw data are shown (Fig. 3b, zoom and c).

DeepSPT correctly identified 89% of 'pre-uncoating' and 75% of 'post-uncoating' time points, yielding a mean accuracy of 85% and a median accuracy of 88% across 100 dual-labeled rotavirus trajectories. This high level of accuracy translated to a median error of just six frames in determining the uncoating time point (Fig. 3d,e). Benchmarking of DeepSPT with HMM, rolling MSD, ref. 13 and the best-performing method of the AnDi challenge showed DeepSPT obtained more accurate identification of rotavirus escape (Extended Data Fig. 8 and Supplementary Figs. 15 and 16). In addition, the characteristics of rotavirus behavior before and after uncoating can be investigated using the heuristics in DeepSPT (Supplementary Fig. 15). Unlike traditional methods, which often require manual analysis taking several minutes to hours per trajectory, DeepSPT automated the identification process, reducing the time to milliseconds per viral trajectory, offering massive acceleration and minimizing any human bias.

Importantly, DeepSPT outputs these predictions, in 500 ms per trajectory, based solely on the motion captured in the DLP trajectories, rendering the secondary VP7 channel redundant (Supplementary Videos 1 and 2). When tested on rotavirus labeled with the same

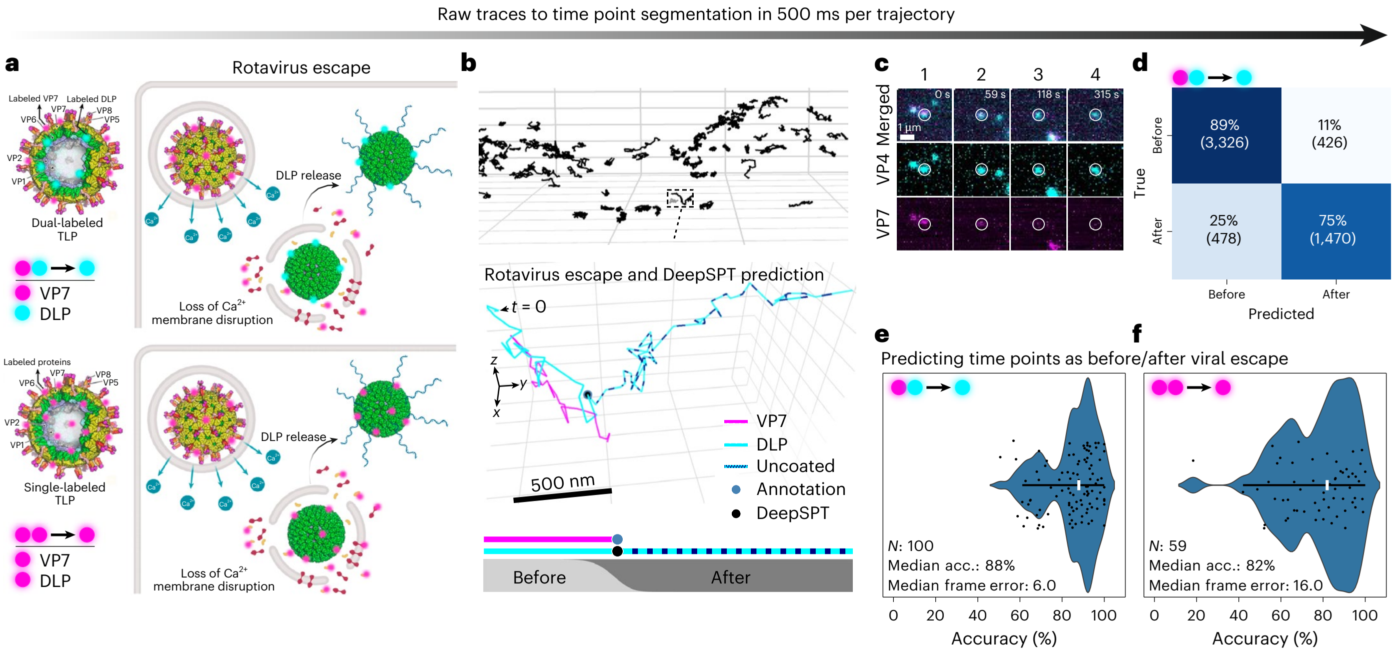


Fig. 3 | Rapid and precise classification of rotavirus uncoating by DeepSPT based exclusively on diffusional behavior. a, A schematic illustration of the typical stages of rotavirus cell-entry pathway from interactions at the plasma membrane, membrane engulfment, membrane permeabilization, calcium-dependent uncoating and escape to the cytosol where RNA production can begin. The experimental approaches for single rotavirus tracking are shown. Top: dual-labeled by recombinant construction of the rotavirus with fluorescently tagged DLP and VP7. Bottom: monochromatic, chemical labeling of free lysines by Atto565. **b**, The 3D tracks of individual rotavirus particles acquired by live-cell LLSM. The zoomed-in image shows an example of rotavirus by parallel multicolor imaging (Methods). The time point for loss of VP7 signal, indicative of uncoating and viral escape (blue dot), is correctly identified by DeepSPT (black dot). The bottom insets show a 1D representation of DLP and VP7 signal with annotations for loss of VP7 and DeepSPT prediction. SoftMax output provides time-resolved probability estimates of 'before uncoating' and 'after uncoating'. **c**, Sum intensity projections of the 3D live-cell LLSM data from a region surrounding the track in the zoom in. The insets contain parallel imaging of DLP (cyan) and VP7 (magenta). The numbered columns show different observed stages of the virus's lifetime for DLP and VP7 from colocalization to uncoating. **d**, A confusion matrix displaying DeepSPT classification performance of predicting time points as 'before uncoating' or 'after uncoating' as compared with ground truth colocalization analysis, entries are normalized to true labels for dual-labeled rotavirus (Methods) (top left, true before; top right, false after; bottom left, false before; bottom right, true after). **e**,**f**, Histograms of DeepSPT classification accuracies as the percentage of time points correctly predicted 'before uncoating' or 'after uncoating' in individual tracks: dual-labeled rotavirus showing median accuracy (acc.) of 88% (100 tracks, *N* = 1 coverslip experiments, 4 movies) (**e**) and monochromatically labeled rotavirus showing median accuracy of 82% (59 tracks, *N* = 5 coverslip experiments, 13 movies) (**f**). DeepSPT requires 500 ms processing time per trajectory to transit from trajectories to feature representations to classification of time points (classification ~1 ms), thus accelerating the analysis by a minimum of four orders of magnitude as compared with manual annotations. Panel **a** created with BioRender.com.

fluorophore on both DLP and VP7 (Methods), DeepSPT exhibited similar performance, achieving a median accuracy of 82% and a mean accuracy of 78% (Methods and Fig. 3f). Notably, DeepSPT needed less than a minute compared with ~8 working days for manual annotation to acquire the ground truth annotations based on intensity loss of the 560 nm channel (Supplementary Fig. 17). By using motion as a marker for viral uncoating, DeepSPT simplifies the experimental design and preparation, avoiding the need for constructing dual-labeled viruses. Thus, DeepSPT frees up one of the two to three available imaging channels, thereby increasing the information content in fluorescence microscopy experiments. To the best of our knowledge, these results constitute the first instance of detecting viral escape into the cytosol based solely on motion and without the need for multicolor labeling.

Colocalization and cellular localization from motion

The capacity to identify biomolecular identity, colocalization partners or to infer subcellular localization based solely on diffusional properties could minimize the need for multicolor imaging and the labor-intensive efforts associated with the creation of cell lines expressing the relevant fluorescent cellular markers. Early and late endosome differentiation, for example, requires multicolor labeling as they might appear to exhibit similar dimensions, are distributed with similar spatial density and display nearly identical diffusion coefficients[52,53]. Traditionally, their identification requires labeling of each endosomal type through antibodies specific for endogenous protein markers enriched in a given type of endosome or by ectopic expression of these markers. Based on multicolor labeling, efforts have been made using analysis of internalized cargo distribution and compartment morphology in fixed samples to deduce general principles of the endocytic machinery[54].

Here, we assess whether DeepSPT can determine endosomal identity based solely on diffusional characteristics, reducing the need for multicolor labeling (Fig. 4a). We used two-color live-cell LLSM to track early endosomes endogenously tagged by gene editing with EEA1-mScarlett and late endosomes tagged with NPC1-Halo-JFX646 (Supplementary Video 3). Their trajectories display indistinguishable diffusion coefficients and alpha values (Fig. 4b and Supplementary Figs. 18 and 19), challenging endosomal identity prediction[37]. In a tenfold stratified cross-validation scheme with varying decision confidence thresholds, DeepSPT achieved accuracies ranging from 70 ± 1.3% to 82 ± 1.8% in classifying EEA1-positive from NPC1-positive compartments. Increasing the confidence threshold improved accuracy but reduced the number of accepted tracks (Fig. 4c). At a 60% confidence threshold, DeepSPT identified endosomal types with an accuracy of 72 ± 1.4% (Fig. 4d) and a recall of 72 ± 3% for EEA1-positive compartments and 72 ± 1.4% for NPC1-positive compartments (Fig. 4d). DeepSPT significantly outperformed the commonly used MSD analysis that reached accuracies of 48 ± 4%, 55 ± 1.6% and 60 ± 1.4% (Supplementary Figs. 20 and 21) in endosomal classification by using the variation

https://doi.org/10.1038/s41592-025-02665-8

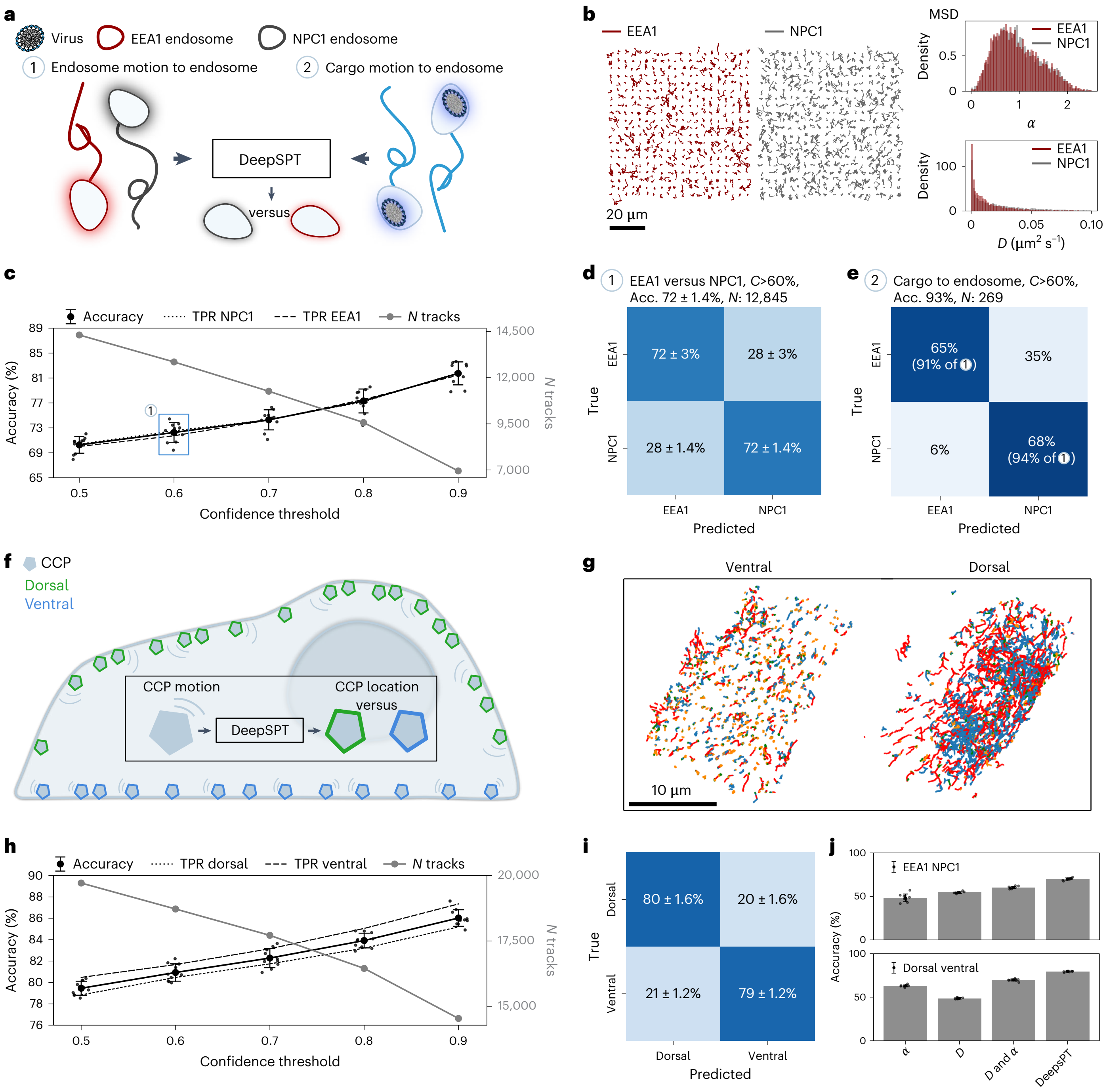


Fig. 4 | Prediction of endosomal identity and AP2 cellular localization based exclusively on temporal diffusional behavior. a, An illustration of endosomal identity prediction solely using diffusion of endosomes or their cargo. **b**, Left: a 2D projection of EEA1-mScarlett- (red) and NPC1-Halo-JFX646-positive (gray) trajectories acquired by LLSM revealing visually similar trajectories. Right: the distribution of the anomalous diffusion exponent (top) and the diffusion coefficient (bottom) for EEA1-positive (red) and NPC1-positive (gray) compartments displaying practically indiscernible distributions ($N = 4{,}770$ EEA1 positive and $N = 9{,}534$ NPC1 positive). **c**, A twin axes plot of accuracy and true positive rate (TPR) (left) and number of tracks (right) versus the confidence threshold (Methods) of the DeepSPT classification (Fig. 2d). Increasing the threshold enhances accuracy but reduces accepted traces. The error bars depict s.d. **d**, A confusion matrix of DeepSPT classification of EEA1- versus NPC1-positive compartments at 60% threshold showing accuracy of 72 ± 1.4% ($N = 4$ experiments, 35 movies). The error bars depict s.d. **e**, A confusion matrix of DeepSPT classification of EEA1- versus NPC1-positive compartments using the rotavirus cargo trajectories with confidence threshold (C) at 60% threshold. Accuracies in parentheses are normalized to the results of **d** for comparison (269 tracks from $N = 12$ experiments, 44 movies). **f**, An illustration of the DeepSPT prediction of AP2 complexes' cellular localization exclusively using diffusional behavior. **g**, The 2D projections of AP2 trajectories below and above 500 nm from the coverslip. Trajectories spending >20% of their lifetime above the 500 nm are considered dorsal and the remaining are considered ventral. Trajectories are color coded by DeepSPT segmentation. **h**, A twin axes plot of accuracy (left) TPR and number of tracks (right) versus confidence threshold (Methods) for the DeepSPT classification of AP2. The error bars depict s.d. **i**, A confusion matrix of the DeepSPT prediction of AP2 showing 79.5 ± 0.6% accuracy at 50% confidence threshold (19,712 tracks; 12,213 dorsal and 7,499 ventral from $N = 5$ experiments, 13 movies). The error bars depict s.d. **j**, Benchmark of DeepSPT versus conventional features based on MSD: diffusion coefficient (D) and anomalous diffusion exponent (α) for EEA1- versus NPC1-positive compartments and dorsal versus ventral AP2. The error bars depict s.d. DeepSPT significantly outperforms MSD analysis (P values <0.00001 using two-sided Welch's t-test, $N = 10$ per condition; Supplementary Table 3).

in α values, variation in the diffusion coefficient (D) or combining both α and D, respectively. DeepSPT, solely using the diffusion traits of endosomal cargo, achieved 91% and 94% of the recall observed in direct prediction of EEA1-positive and NPC1-positive compartments, respectively (Fig. 4e). Benchmarking DeepSPT against ref. 13, method E of the AnDi challenge, the classification using the variation in α values, in the diffusion coefficient (D) or combining both α and D on identifying endosomal identity solely from diffusion showcased that DeepSPT clearly outperformed all the methods (Extended Data Fig. 9). The ability of DeepSPT to differentiate early from late endosomes based solely on their motion, or that of their cargo, accelerates data acquisition and analysis while minimizing potential perturbations and/or the need for multicolor tagging.

To assess whether DeepSPT can infer identity for additional systems, we first applied it to a new dataset of single-particle trajectories of the assembly of CCPs and coated vesicles forming at the cell surface. The dynamic assembly and intracellular location of these structures was obtained by tracking the clathrin AP2 adapter complex, gene edited at its sigma subunit with eGFP, using 3D live LLSM. A 2D projection of the acquired AP2 trajectories qualitatively indicated that the diffusional properties of AP2 were correlated with cellular location[55]—dorsal versus ventral cell surface (Fig. 4f,g), which was quantitatively confirmed by DeepSPT's temporal segmentation of the 3D traces (Supplementary Fig. 19). DeepSPT accurately predicted AP2 cellular location in a tenfold cross-validation scheme, yielding accuracies from 79.5 ± 0.6% to 86.0 ± 0.8% at different confidence thresholds (Fig. 4h). Without applying any confidence filter (that is, a 50% threshold), DeepSPT classified the cellular location of AP2 with recalls of approximately 80% for both classes (Fig. 4i). In contrast, pinpointing the cellular location of AP2 using MSD features reached accuracies of 62.5 ± 1.8%, 48.7 ± 0.8% and 70 ± 1.3%, with a maximum recall for dorsal tracks of 60 ± 2% (Fig. 4g and Supplementary Figs. 20 and 21). While ref. 13 and method E of the AnDi challenge classified AP2 better than the MSD-based features, DeepSPT clearly outperforms all benchmark methods (Extended Data Fig. 10).

DeepSPT also displays higher or similar accuracy than the benchmark approaches in identifying biological information in four experimental datasets of tracking individual enzymes (Supplementary Figs. 22 and 23), transcription factors (Supplementary Fig. 24) and drug nanocarriers (Supplementary Fig. 25), highlighting the applicability of DeepSPT to diverse systems. Subtle diffusional variations across systems, while missed by common tools, are utilized by DeepSPT to precisely output biological information in complex systems, while also allowing investigation into the importance of each feature, thus providing mechanistic insights (Supplementary Fig. 26). DeepSPT achieves this across various biological contexts, imaging protocols and experimental conditions.

Discussion

The diffusion of biomolecules within cells exhibits both spatial and temporal heterogeneity and varies across biological systems and functionalities. Extracting quantitative temporal information from live-cell imaging is currently an analytical bottleneck and often relies on system-specific analysis or even manual annotation. DeepSPT overcomes this bottleneck by providing a framework to transition from raw trajectories to quantitative temporal information rapidly, precisely and with minimal human intervention both for 2D and 3D diffusion. The DeepSPT's multimodular architecture capitalizes on—and extends beyond—our previous work on diffusional fingerprinting[13] or others on deep learning, by combining a segmentation and a fingerprinting module as well as a downstream classifier module. The segmentation module, trained on trajectories with broadly distributed diffusional properties, consistently outperformed existing state-of-the-art toolboxes in segmenting diverse heterogeneous diffusional behaviors in simulated and experimental data. The implementation of uncertainty-calibrated probability estimates enhances the transparency of DeepSPT's output, enabling users with limited a priori knowledge of the biological system to gauge model certainty. The diffusional fingerprinting module expands from 17 to 40 features, ensuring a more holistic quantification of diffusional metrics and temporal information from which users can easily employ feature selection. The implementation of the classification module offers, to the best of our knowledge, the first output of biological information based on diffusional properties alone.

DeepSPT currently considers as input the output of a particle tracker such as u-track[9] or Trackpy[56], which already account for photoblinking, and DeepSPT is not designed to solve blinking or loss of tracking errors. However, DeepSPT demonstrates robust and accurate predictions even for imaging settings producing localization errors at the size of the average step length, multiple linking errors and trajectories with durations around just 20 frames or less (Extended Data Figs. 2–4 and Supplementary Figs. 7 and 14). Being trained on diffusion coefficients spanning four orders of magnitude and behavior segments of highly varying duration, DeepSPT is designed to capture a large number of transition rates but is not trained to capture state transitions faster than the imaging frame rate. The multimodality of DeepSPT can provide valuable insights for unsupervised data exploration by supplying pretrained deep learning-based diffusional segmentation and diffusional heuristics for any 2D or 3D trajectories. At the same time, using the extensive feature set, DeepSPT enables accurate and interpretable downstream supervised learning with potential for feature selection and feature importance evaluation for determining key mechanistic insights. End-to-end models could potentially obtain similar or improved predictions, given sufficient training data, albeit at the cost of potentially longer training times and the loss of the interpretable features[13]. Altering the DeepSPT architecture or/and the network's structure by, for example, utilizing localization shape, motion blur, localization intensity, signal-to-background ratio or axis-specific uncertainty, might improve performance; the accessible source code of DeepSPT is instrumental for that. The minimal requirement for human intervention highlights the potential of DeepSPT to enhance both the reproducibility and robustness of conclusions across different laboratories. Being open source and freely available to the public allows future users to perform customized analyses according to individual research needs.

The precise temporal segmentation combined with the comprehensive quantification of diffusional properties of DeepSPT, coupled with its trained downstream classifier, facilitate rapid prediction of viral uncoating events—achieving results in seconds as opposed to weeks as required for manual annotation. These four orders of magnitude acceleration, not only marks deep learning-assisted identification of viral uncoating, but also shifts the bottleneck in single-particle discoveries from data analysis to data acquisition. It even introduces the potential for virtually real-time analysis of early stages of viral infection.

Subtle diffusional variations in 2D or 3D, while missed by common tools, are utilized by DeepSPT to precisely output biological information in complex systems across various biological contexts, imaging protocols and experimental conditions. For example, DeepSPT discerned EEA1-positive from NPC1-positive compartments solely based on their respective 3D diffusional characteristics, or that of their cargo, with accuracies of 72% significantly outperforming the commonly used MSD analysis that reached accuracies of 50–60%. These findings prompt further mechanistic studies to explore whether divergent diffusional behaviors stem from distinct external interaction partners, inherent physical differences between endosomal compartments or other variables. DeepSPT similarly pinpointed the cellular location of AP2 on 3D data with an accuracy of 80%, significantly outperforming common analysis reaching ~50–70% accuracy. The distinct diffusional behaviors of AP2 highlighted the importance of careful selection in imaging setups. Applied on 2D data of insulin internalization, DeepSPT found insulin mainly exhibits subdiffusive behavior but included segments of directed motion indicative of active transport (Supplementary Figs. 12 and 13). The ability of DeepSPT to accurately quantify heterogeneous behaviors in both 2D and 3D, across diverse biological systems and under varying experimental and imaging conditions, attests to its utility as a platform for characterizing heterogeneous diffusion across systems[57].

The capacity of DeepSPT for predicting viral uncoating events, identifying endosomal types and discerning colocalization partners and cellular localization solely based on diffusion extends the traditional structure-to-function paradigm in proteins to a novel motion-to-function paradigm. This suggests that, alongside structure[7,35], motion can also serve as an indicator of both function and identity. This development opens avenues for employing motion as a biomarker and for label-minimal analyses—effectively substituting fluorescent labels with temporal diffusional analysis. Such a shift could simplify experimental design and reduce preparation time, reduce phototoxicity or potentially enrich experiments by reallocating redundant fluorescent markers for other applications.

To simplify the implementation of DeepSPT across laboratories and expand our user base to a broader audience, we have provided a standalone, intuitive GUI allowing the execution of all functionalities of the DeepSPT from outputting the analyzed data, diffusional features, segmentation and classification to publication-quality figures. The GUI also allows users to conveniently utilize any of the modules individually or all of them combined. The interactive built-in trajectory viewer allows visualization of the trajectories and diffusional segmentation in 2D and 3D. The open-source implementation allows specialized users to optimize and augment the pipeline for their specific needs.

Widespread implementation of DeepSPT across laboratories could facilitate the creation of comprehensive libraries detailing characteristic movements of cells, subcellular structures and biomolecules. An open-source diffusional library of this kind would offer a new instrument for the scientific community, aiding in the exploration of 4D cell biology through temporal diffusional behavior.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41592-025-02665-8.

References

1. Cocucci, E., Aguet, F., Boulant, S. & Kirchhausen, T. The first five seconds in the life of a clathrin-coated pit. *Cell* **150**, 495–507 (2012).
2. Johnson, C., Exell, J., Lin, Y., Aguilar, J. & Welsher, K. D. Capturing the start point of the virus-cell interaction with high-speed 3D single-virus tracking. *Nat. Methods* **19**, 1642–1652 (2022).
3. Liu, T.-L. et al. Observing the cell in its native state: Imaging subcellular dynamics in multicellular organisms. *Science* **360**, eaaq1392 (2018).
4. Thomsen, R. P. et al. A large size-selective DNA nanopore with sensing applications. *Nat. Commun.* **10**, 5655 (2019).
5. Aguet, F. et al. Membrane dynamics of dividing cells imaged by lattice light-sheet microscopy. *Mol. Biol. Cell* **27**, 3418–3435 (2016).
6. Moses, M. E. et al. Single-molecule study of *Thermomyces lanuginosus* lipase in a detergency application system reveals diffusion pattern remodeling by surfactants and calcium. *ACS Appl. Mater. Interfaces* **13**, 33704–33712 (2021).
7. Jensen, S. B. et al. Biased cytochrome P450-mediated metabolism via small-molecule ligands binding P450 oxidoreductase. *Nat. Commun.* **12**, 2260 (2021).
8. Gabriele, M. et al. Dynamics of CTCF- and cohesin-mediated chromatin looping revealed by live-cell imaging. *Science* **376**, 496–501 (2022).
9. Jaqaman, K. et al. Robust single-particle tracking in live-cell time-lapse sequences. *Nat. Methods* **5**, 695–702 (2008).
10. Wan, F. et al. Ultrasmall TPGS-PLGA hybrid nanoparticles for site-specific delivery of antibiotics into *Pseudomonas aeruginosa* biofilms in lungs. *ACS Appl. Mater. Interfaces* **12**, 380–389 (2020).
11. Gal, N., Lechtman-Goldstein, D. & Weihs, D. Particle tracking in living cells: a review of the mean square displacement method and beyond. *Rheol. Acta* **52**, 425–443 (2013).
12. Arcizet, D., Meier, B., Sackmann, E., Rädler, J. O. & Heinrich, D. Temporal analysis of active and passive transport in living cells. *Phys. Rev. Lett.* **101**, 248103 (2008).
13. Pinholt, H. D., Bohr, S. S.-R., Iversen, J. F., Boomsma, W. & Hatzakis, N. S. Single-particle diffusional fingerprinting: a machine-learning framework for quantitative analysis of heterogeneous diffusion. *Proc. Natl Acad. Sci. USA* **118**, e2104624118 (2021).
14. Kowalek, P., Loch-Olszewska, H. & Szwabiński, J. Classification of diffusion modes in single-particle tracking data: feature-based versus deep-learning approach. *Phys. Rev. E* **100**, 032410 (2019).
15. Benning, N. A. et al. Dimensional reduction for single-molecule imaging of DNA and nucleosome condensation by polyamines, HP1α and Ki-67. *J. Phys. Chem. B* **127**, 1922–1931 (2023).
16. Vega, A. R., Freeman, S. A., Grinstein, S. & Jaqaman, K. Multistep track segmentation and motion classification for transient mobility analysis. *Biophys. J.* **114**, 1018–1025 (2018).
17. Monnier, N. et al. Inferring transient particle transport dynamics in live cells. *Nat. Methods* **12**, 838–840 (2015).
18. Persson, F., Lindén, M., Unoson, C. & Elf, J. Extracting intracellular diffusive states and transition rates from single-molecule tracking data. *Nat. Methods* **10**, 265–269 (2013).
19. Chen, Z., Geffroy, L. & Biteen, J. S. NOBIAS: analyzing anomalous diffusion in single-molecule tracks with nonparametric Bayesian inference. *Front. Bioinform.* **1**, 742073 (2021).
20. Arts, M., Smal, I., Paul, M. W., Wyman, C. & Meijering, E. Particle mobility analysis using deep learning and the moment scaling spectrum. *Sci. Rep.* **9**, 17160 (2019).
21. Vink, J. N. A., Brouns, S. J. J. & Hohlbein, J. Extracting transition rates in particle tracking using analytical diffusion distribution analysis. *Biophys. J.* **119**, 1970–1983 (2020).
22. Martens, K. J. A. et al. Visualisation of dCas9 target search in vivo using an open-microscopy framework. *Nat. Commun.* **10**, 3552 (2019).
23. Karslake, J. D. et al. SMAUG: analyzing single-molecule tracks with nonparametric Bayesian statistics. *Methods* **193**, 16–26 (2021).
24. Simon, F., Tinevez, J.-Y. & van Teeffelen, S. ExTrack characterizes transition kinetics and diffusion in noisy single-particle tracks. *J. Cell Biol.* **222**, e202208059 (2023).
25. Momboisse, F. et al. Tracking receptor motions at the plasma membrane reveals distinct effects of ligands on CCR5 dynamics depending on its dimerization status. *eLife* **11**, e76281 (2022).
26. Hansen, A. S. et al. Robust model-based analysis of single-particle tracking experiments with Spot-On. *eLife* **7**, e33125 (2018).
27. Martens, K. J. A., Turkowyd, B., Hohlbein, J. & Endesfelder, U. Temporal analysis of relative distances (TARDIS) is a robust, parameter-free alternative to single-particle tracking. *Nat. Methods* https://doi.org/10.1038/s41592-023-02149-7 (2024).
28. Muñoz-Gil, G. et al. Objective comparison of methods to decode anomalous diffusion. *Nat. Commun.* **12**, 6253 (2021).
29. You, B. & Yang, G. Attention-based LSTM for motion switching detection of particles in living cells. In *2021 International Joint Conference on Neural Networks (IJCNN)* 1–6 (IEEE, 2021); https://doi.org/10.1109/IJCNN52387.2021.9533629

30. Dosset, P. et al. Automatic detection of diffusion modes within biological membranes using back-propagation neural network. *BMC Bioinformatics* **17**, 197 (2016).
31. Wagner, T., Kroll, A., Haramagatti, C. R., Lipinski, H.-G. & Wiemann, M. Classification and segmentation of nanoparticle diffusion trajectories in cellular micro environments. *PLoS ONE* **12**, e0170165 (2017).
32. Granik, N. et al. Single-particle diffusion characterization by deep learning. *Biophys. J.* **117**, 185–192 (2019).
33. Simon, F., Weiss, L. E. & van Teeffelen, S. A guide to single-particle tracking. *Nat. Rev. Methods Prim.* **4**, 66 (2024).
34. Qu, X. et al. Semantic segmentation of anomalous diffusion using deep convolutional networks. *Phys. Rev. Res.* **6**, 013054 (2024).
35. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
36. Weigert, M. et al. Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nat. Methods* **15**, 1090–1097 (2018).
37. Thomsen, J. et al. DeepFRET, a software for rapid and automated single-molecule FRET data classification using deep learning. *eLife* **9**, e60404 (2020).
38. Malle, M. G. et al. Single-particle combinatorial multiplexed liposome fusion mediated by DNA. *Nat. Chem.* **14**, 558–565 (2022).
39. Levet, F. et al. SR-Tesseler: a method to segment and quantify localization-based super-resolution microscopy data. *Nat. Methods* **12**, 1065–1071 (2015).
40. Kim, H. K. et al. SpCas9 activity prediction by DeepSpCas9, a deep learning-based model with high generalization performance. *Sci. Adv.* **5**, eaax9249 (2019).
41. Wong, F. et al. Discovery of a structural class of antibiotics with explainable deep learning. *Nature* **626**, 177–185 (2024).
42. Dunn, K. W., Kamocka, M. M. & McDonald, J. H. A practical guide to evaluating colocalization in biological microscopy. *Am. J. Physiol., Cell Physiol.* **300**, C723–C742 (2011).
43. Merino Urteaga, R. & Ha, T. Mind your tag in single-molecule measurements. *Cell Rep. Methods* **3**, 100623 (2023).
44. Yin, X.-X., Sun, L., Fu, Y., Lu, R. & Zhang, Y. U-Net-based medical image segmentation. *J. Healthc. Eng.* **2022**, 4189781 (2022).
45. Ruthardt, N., Lamb, D. C. & Bräuchle, C. Single-particle tracking as a quantitative microscopy-based approach to unravel cell entry mechanisms of viruses and pharmaceutical nanoparticles. *Mol. Ther.* **19**, 1199–1211 (2011).
46. Guo, C., Pleiss, G., Sun, Y. & Weinberger, K. Q. On calibration of modern neural networks. In *Proc. 4th International Conference on Machine Learning* 1321–1330 (JMLR, 2017).
47. Michalet, X. Mean square displacement analysis of single-particle trajectories with localization error: Brownian motion in an isotropic medium. *Phys. Rev. E* **82**, 041914 (2010).
48. Slator, P. J., Cairo, C. W. & Burroughs, N. J. Detection of diffusion heterogeneity in single particle tracking trajectories using a hidden Markov model with measurement noise propagation. *PLoS ONE* **10**, e0140759 (2015).
49. Abdelhakim, A. H. et al. Structural correlates of rotavirus cell entry. *PLoS Pathog.* **10**, e1004355 (2014).
50. Salgado, E. N., Garcia Rodriguez, B., Narayanaswamy, N., Krishnan, Y. & Harrison, S. C. Visualization of calcium ion loss from rotavirus during cell entry. *J. Virol.* **92**, e01327-18 (2018).
51. Aoki, S. T. et al. Cross-linking of rotavirus outer capsid protein VP7 by antibodies or disulfides inhibits viral entry. *J. Virol.* **85**, 10509–10517 (2011).
52. Rink, J., Ghigo, E., Kalaidzidis, Y. & Zerial, M. Rab conversion as a mechanism of progression from early to late endosomes. *Cell* **122**, 735–749 (2005).
53. Piper, R. C. & Katzmann, D. J. Biogenesis and function of multivesicular bodies. *Annu. Rev. Cell Dev. Biol.* **23**, 519–547 (2007).
54. Collinet, C. et al. Systems survey of endocytosis by multiparametric image analysis. *Nature* **464**, 243–249 (2010).
55. Cocucci, E., Gaudin, R. & Kirchhausen, T. Dynamin recruitment and membrane scission at the neck of a clathrin-coated pit. *Mol. Biol. Cell* **25**, 3595–3609 (2014).
56. Allan, D., Caswell, T., Keim, N. & van der Wel, C. Trackpy: Trackpy v0.3.2. *Zenodo* https://doi.org/10.5281/zenodo.60550 (2016).
57. Mizrak, A. et al. Single-molecule analysis of protein targeting from the endoplasmic reticulum to lipid droplets. Preprint at *bioRxiv* https://doi.org/10.1101/2024.08.27.610018 (2024).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

DeepSPT's diffusional fingerprinting module

Encompassing and expanding on the work in ref. 13, a recent study providing a set of diffusional metrics to transform single-particle trajectories into fixed-length representations of interpretable features functioning as unique identifiers, that is, diffusional fingerprints, this work both extends the number of descriptive features of diffusional behavior from 17 to 40 and, importantly, provides temporal features and sequential representations to enable time-resolved predictions (Supplementary Table 3). This is outlined in the following sections.

Confinement radius and directed velocity

As DeepSPT allows accurate diffusional behavior segmentation, there is the possibility of using MSD equations adapted for the specific motion types such as $\mathrm{MSD} = r^2(1 - A_1 e^{\frac{-2A_2 dDt}{r^2}}) + \mathrm{offset}$, where r is the confinement radius, A_1 and A_2 are shape parameters, d is the number of dimensions, D is the diffusion coefficient, t is time and $\mathrm{MSD} = 2dDt^2 + v^2t^2 + \mathrm{offset}$ for confined and directed velocity (v), respectively[31]. The offset in MSD analysis corresponds to the constant contribution to MSD from localization error. Thus, the velocity term of directed motion and the confinement radius for confined motion can be extracted for whole trajectories or for subtracks.

Directionality analysis using vector DPs

The dot product (DP) between two vectors informs of the angle between them and the product of their magnitudes, specifically for normalized vectors the DP returns the cosine of the angle (−1, 1) with 1 for parallel vectors, 0 for orthogonal and −1 for anti-parallel. For a given trajectory, the two vectors $((p_{i0}, p_{i+1})$ and $(p_{i+1}, p_{i+2}))$ formed by the three consecutive coordinates (p) at times (i) in the trajectory; p_0, p_1, p_2 can be used to compute the rolling DP along the trajectory, to investigate persistence in directionality. As performing a rolling DP calculation requires constructing two vectors using three consecutive coordinates, the resulting rolling DP vector is two elements shorter than the original trajectory. For matching a rolling DP vector to the length of the original trajectory, two zeros are padded to start the DP vector. Three strategies are used to aggregate the per-frame DP series into a single value to complement the diffusional fingerprinting: (1) Averaging (MeanDP) to show any average tendency in directionality, where ~0 is to be expected for normal diffusion, >0 for directed and <0 for subdiffusive; (2) persistence (persistDP) to investigate whether consecutive vectors in general tend to have direction persistence, specifically, the percentage of successive occurrences both being positive or negative is computed; and (3) sign analysis (AvgSignDP) counts the percentage of vector DP signs being positive.

Additional step length descriptive statistics

Step lengths contain much information on single-particle tracks, here we add additional descriptive statistics such as minimum (MinSL); maximum (MaxSL); broadness of the step length distribution (BroadSL), that is, MaxSL − MinSL; and the coefficient of variation of the step length distribution that is the ratio between the s.d. and mean, which measures the variability of the distribution in relation to the mean. The arrested fraction and fast-moving fraction are system specific but computed as the percentage of steps under 0.1 µm and above 0.4 µm, respectively. The calculation of the instantaneous diffusion coefficient was implemented using the MSD between adjacent positions, $\mathrm{MSD}/2dt$.

Volume/area of trajectories

The area or volume (referred to as volume for consistency) of 2D or 3D trajectories is the volume of the convex hull enclosing the trajectory x, y, (z) coordinates in 2D or 3D, respectively. The volume of a trajectory or a subtrack is a direct reflection of the trajectory shape and holds information on the amount of volume explored by the trajectory and indirectly the morphology of the explored region. Therefore, the volume may be used to identify restricted particles versus more freely moving particles and provide a metric for the volume of confinement, and is computed using the Python package SciPy[58].

Temporal features

To inform a classifier of temporal variation in trajectories, the temporal segmentation was condensed into unique features. Four features constructed from the percentage of time spent as either normal diffusion, directed motion, confined diffusion or subdiffusive, and a feature from the number of changes in diffusion. To inform on the history of diffusional changes, the sequence of diffusional behaviors was encoded as normal diffusion, 0; directed motion, 1; confined diffusion, 4 and subdiffusive, 6; for example, 0146 for sampling each behavior once starting with normal diffusion. Encoding values were purposely chosen to encode similar motion types to similar values while encoding unique distances between values for each motion type. Six features were constructed from this encoding: mean and median informing on most likely motion type, maximum and minimum informing on sampled motion, and s.d. and median of distances between adjacent sequence values to inform on the changes and similarities of sampled motion types.

DeepSPT's temporal segmentation module

The temporal segmentation module within DeepSPT consists of an ensemble of three U-Nets[59] providing end-to-end transformation from raw trajectories to trajectory segmentation. We chose U-Nets based on our hypothesis that their focus on local receptive fields, shared feature maps, hierarchical feature composition and translational invariance in their inductive biases make them particularly well-suited for modeling the properties of diffusion. Each U-Net model is trained on a dataset of 300,000 simulated trajectories as described in the 'Stochastic simulation of diffusion' and 'Generating heterogeneous diffusion' sections with 80% heterogeneous motion and 20% homogeneous. The end-to-end architecture of the well-known U-Net can be seen as a downsampling by an encoder network, an upsampling by a decoder network followed by a classifier. Individual U-Nets contain 1D convolutional layers and max pooling in the encoding network and 1D convolutional layers and nearest neighbor upsampling in the decoding network, with the encoder and decoder connected by skip connections. The encoding and decoding are directly followed by a series of convolutional layers before an ensemble average SoftMax output. The SoftMax outputs from each model in the ensemble are combined by averaging to produce the final DeepSPT prediction. Specific hyperparameters were found by Optuna's tree-structured Parzen estimator[60], and the best set of hyperparameters, not including ensemble size, were selected based on the performance on a test set only used for the hyperparameter search. For 2D temporal segmentation, the selected hyperparameters included four downsampling and four upsampling steps. The encoder comprises two layers and the decoder one layer, with three bottom layers and four output layers. A kernel size of 7 with a dilation rate of 2 is used for the encoder, bottom layers and decoder, while the output layers employ a kernel size of 3 without dilation. The channel dimension following the input features is set to 130, with a factor of 2 applied for channel multiplication during downsampling and division during upsampling. For 3D temporal segmentation, the hyperparameters include three downsampling and three upsampling steps. The encoder consists of three layers, the decoder of four layers, with four bottom layers and two output layers. A kernel size of 5 with a dilation rate of 2 is applied to the encoder, bottom layers and decoder, while the output layers use a kernel size of 3 without dilation. The channel dimension following the input features is set to 48, with a factor of 2 applied for channel multiplication during downsampling and division during upsampling. The model is written in Pytorch. See our open-source implementation on GitHub for more detail.

Overview on the training and evaluation of DeepSPT modules

The training data of the DeepSPT temporal segmentation module (M1) consists of three independent datasets of simulated diffusion each comprising 300,000 trajectories, where 80% are heterogeneous and 20% are homogeneous diffusion. The construction of these trajectories is described in 'Stochastic simulation of diffusion' and 'Generating heterogeneous diffusion' sections. The training of M1 is by the Adam optimizer and a cross-entropy loss. The accuracy evaluation of M1 is done using withheld test sets. The DeepSPT diffusional fingerprinting module (M2) is an extensive set of diffusional heuristics derived directly from any data, thus does not require training. The descriptive power of the diffusional fingerprinting features is evaluated by classification performance and feature importance evaluation. The training data of the DeepSPT task-specific classifier (M3) module is task specific and training is performed directly on experimental data and evaluation is by cross-validation.

DeepSPT's temporal segmentation module for the 2021 AnDi challenge data

The principles and architecture of DeepSPT when retrained for the AnDi challenge data remain the same as explained in 'DeepSPT's temporal segmentation module' section. Each U-Net model is trained on a dataset of 400,000 simulated trajectories as described in the 'Simulated test set for evaluation of temporal segmentation for AnDi challenge data' section both for AnDi task 3 and for the highly heterogeneous trajectories. Again, specific hyperparameters were found by Optuna's tree-structured Parzen estimator[60] using an independent validation set.

Temperature scaling

Neural networks have been shown to be overly confident even in erroneous predictions[46]. Such overconfidence can be mitigated by uncertainty calibration, so confidence estimates resemble more the actual ground truth correctness likelihood. Temperature scaling is one such method, which has been found effective despite its simplicity[46]. Temperature scaling introduces a constant into the last layer before SoftMax of a classifier and this constant is tuned to minimize negative log likelihood as recent work shows that this allows approximation of the actual posterior probability distribution[46]. Measures of uncertainty calibration include expected calibration error, which is effectively the one-norm error between perfect calibration and the actual calibration; sharpness, which measures the distance between the maximum class probability scores for the k classes, and 1 as the perfect classifier would have a class probability score of 1 for correct prediction; and last, the negative log likelihood (NLL), here NLL is reported as the NLL improvement relative to random predictions for a more intuitive metric.

Stochastic simulation of diffusion

Tracks are generated with a stochastic diffusion coefficient log-uniformly sampled between 0.0001 and 0.5 μm^2 s^{-1}. Owing to the scale invariance property of diffusion, the sampling of D from a broad spectrum equates sampling with varying sized time steps (t), thus allows the model to learn the characteristics of diffusional behavior both for varying D and t. Simulated tracks were generated starting in $x = y = 0$ with lengths uniformly drawn between 5 frames and 600 frames challenging the model to pick up regularities even in shorter traces. Normal, directed and subdiffusion were simulated following Pinholt et al.[13] and Kowalek et al.[14]. The different parameters for the simulation of these diffusion types were chosen similarly to Wagner et al.[31], Pinholt et al.[13] and Kowalek et al.[14]. Except for three parameters that were made even broader distributed: α, an anomalous exponent term measuring motion persistence 0–0.7; the step length to localization error ratio defined as $Q = \sqrt{\frac{D\times t}{\sigma_{\mathrm{noise}}}}$ and $Q_{\mathrm{directed}} = \sqrt{\frac{D\times t+v^2t^2}{\sigma_{\mathrm{noise}}}}$ with Q and Q-directed uniformly sampled between 1 and 16; and last, the extent to which active motion affects the diffusion, $R = \frac{v^2t}{4D}$ was generated uniformly between 5 and 25. Confined motion also differs from previous work as the confinement radius in this work is independent of track duration, reflecting the case where a radius of reflecting boundary area does not grow as the experiment progresses. Instead, we simulated confinements irrespective of the longevity of the tracks. The area or volume of confinement is defined by an ellipse in 2D and an ellipsoid in 3D, allowing any orientation, with the semi-major and semi-minor axes uniformly sampled between 5 nm and 250 nm. For the 3D case, the two semi-minor axes were chosen to be equilength, thus producing a large range of confinement areas/volumes in any given orientation.

Generating heterogeneous diffusion

Trajectories with heterogeneous diffusional behavior were simulated as homogeneous tracks with the addition of sampling random change points in diffusional behaviors of up to four states. Therefore, a given trajectory was separated into random subtraces with a required minimum length of five frames, thus, the length of the trajectory must be larger than the product of change points and minimum length. Sampling change points randomly inside a trajectory was purposely chosen instead of having the states follow a user-defined Markov model to ensure DeepSPT's decision-making is not influenced by learning an underlying Markov model that does not necessarily resemble anything found in nature, but rather keeping DeepSPT fully agnostic.

Simulated test set for evaluation of temporal segmentation

Evaluation was performed on 20,000 withheld simulated trajectories, 80% heterogeneous and 20% homogeneous motion, with all associated diffusional parameters broadly distributed as described under 'Stochastic simulation of diffusion' section.

Optimizing benchmark LSTM performance

To improve the performance of the attention-based LSTM used for benchmarking, the LSTM is retrained using hyperparameters of the original publication[29] and retrained on the 300.000 simulated trajectories used to train one of three individual U-nets in DeepSPT.

Moving simulated diffusion to 3D

The work of Wagner et al.[31], Pinholt et al.[13] and Kowalek et al.[14] is exclusively focused on diffusion in 2D, whereas due to our 3D live-cell lattice light-sheet experiments, we are required to extend previous work to 3D. Simulating normal diffusion and subdiffusive motion easily extends to 2D and 3D cases as axes of diffusion are independent. In Wagner et al.[31], Pinholt et al.[13] and Kowalek et al.[14], Didected motion is simulated using the cosine and sine to ensure directionality in the x and y direction, respectively, with the velocity as a magnitude term, which we extend to the 3D case by considering the unit sphere instead, thus the added velocities become $\mathrm{d}x = v \times \mathrm{d}t \times \sin(\phi) \times \cos(\theta)$, $\mathrm{d}y = v \times \mathrm{d}t \times \sin(\phi) \times \sin(\theta)$ and $\mathrm{d}z = v \times \mathrm{d}t \times \cos(\phi)$, where θ is the polar angle and ϕ is the azimuthal angle.

Simulation of two populations of heterogeneous diffusion with diffusional properties with overlapping distributions

Two populations of 500 trajectories, each with track durations uniformly sampled between 150 and 200 time points, were constructed using the aforementioned simulation framework of heterogeneous motion of the four diffusion types (Methods). Both populations had the step length to localization error ratio uniformly sampled between 6 and 16. Population 1 had active motion ratios uniformly sampled between 5 and 12 and subdiffusive motion with α uniformly sampled between 0.3 and 0.6. Population 2 had active motion ratios uniformly sampled between 8 and 15 and subdiffusive motion with α uniformly sampled between 0.4 and 0.7. Otherwise, populations were constructed identically at increments of diffusion coefficients. Diffusion coefficients are log-uniformly sampled between 0.004 μm^2 s^{-1} and 0.0008 μm^2 s^{-1} separating in increments of 0.005 μm^2 s^{-1}. After each stochastic simulation of trajectories, the distributions of instantaneous diffusion coefficients were computed as described (Methods) and overlap in the two distributions computed as the histogram intersection-normalized total tracks in one population. To build a test set for diffusional change point prediction with ground truth, trajectories from these two simulated populations are combined into 5,000 trajectories with a single change point. These trajectories, sampling both behaviors, can start as either population with change points randomly distributed yet a minimum of five time points from either end of the trajectory. The two populations chosen to be combined had approximately 75% overlap in diffusion coefficient.

Simulated test set for evaluation of temporal segmentation for AnDi challenge data

Two test sets were constructed based on the 2021 AnDi challenge. First, trajectories were simulated using the 2021 AnDi challenge task 3 open-source framework directly, which constructs trajectories with a duration of 200 time points, each consisting of two segments of anomalous diffusion randomly selected between the five diffusion behaviors in the AnDi challenge[28]. The five diffusion behaviors in the AnDi challenge include annealed transient time motion, continuous-time random walk, fractional Brownian motion, Lévy walk and scaled Brownian motion[28]. Second, trajectories were simulated by combining the anomalous diffusion behaviors from the AnDi challenge open-source framework into heterogeneous trajectories sampling multiple diffusional behaviors with multiple change points. Specifically, for each track of 200 time points, (1) a random number of segments was chosen by sampling uniformly between 3 and 6; (2) each segment's length was uniformly sampled between 5 and 200, while ensuring the sum of segment durations totals 200; (3) each segment samples one of the five diffusional behaviors uniformly while ensuring that neighboring segments can not express the same behavior, as resampling behavior type generates fewer, longer lasting segments, which can simplify analysis; (4) trajectories are generated for each motion type using the AnDi challenge open-source framework; (5) following Muñoz-Gil et al.[28], trajectories are standardized to ensure unitary s.d. of displacements over time and then scaled by a random factor drawn from the unit normal distribution; (6) trajectories are displaced at each time point independently by an addend randomly sampled from a normal distribution to mimic localization error. This normal distribution has mean zero and a s.d. corresponding to 50% of the average step length in each dimension (x, y and z).

Effects of tracking errors on DeepSPT temporal segmentation. Evaluating the effect of tracking errors on DeepSPT's temporal segmentation accuracy is performed by simulating a population of 1,000 trajectories for 200 frames with broadly distributed diffusional parameters inside a box of varying dimensions and tracking using Trackpy. The diffusional properties of the population are defined as per 'Stochastic simulation of diffusion' section, apart from a maximum of 200 frames and diffusion coefficient ranging from 0.01 to 0.05 $\mu m^2 s^{-1}$. The initial positions of the trajectories are randomly sampled with replacement within a box. Any track leaving the box is removed. In addition, 10,000 localizations not belonging to any trajectory are randomly distributed within the box and uniformly across frames producing on expectation of 50 false positive detections per frame. These trajectories and the false positives are then treated as individually detected localizations and tracked using Trackpy with a search range of 1.5 μm and no memory. The box dimensions are chosen as 20,000, 2,000, 1,000, 500, 200, 150 and 100 μm, with tracking errors increasing with decreasing box dimensions. Three simulations per dimension are performed each with an independent set of 1,000 trajectories.

Time-resolved task-specific downstream classifier using temporal and diffusional features

Segmentation of rotavirus trajectories into 'before uncoating' and 'after uncoating' was performed using a sequence-to-sequence model trained on time-resolved diffusional features computed using the temporal segmentation and diffusional fingerprinting module in windows. Raw trajectories can be seen as a time series with three parallel channels (xyz) per time point; these were transformed into time series of identical length but now 40 channels (one per feature in the temporal segmentation and diffusional fingerprinting modules) utilizing a window of 31 frames centered on each time point in a given trajectory. The ground truth of the uncoating time point for dual-labeled rotavirus was constructed based on the loss of colocalization between VP4 and VP7. For single-labeled rotavirus, labels are based on manual annotation of the endpoint of intensity drop following loss of VP7 signal (Supplementary Fig. 17). Both cases produce binary target time series, which are filtered if loss of colocalization is in the first or last frame. The sequence-to-sequence model architecture consists of a bidirectional five-layer gated recurrent unit followed by a fully connected feedforward layer. The output length of the gated recurrent unit is twice the input length due to its bidirectionality, thus split in half and combined by summation to match the input length before the fully connected layer. Before training, trajectories are evaluated for similarity by root mean squared distance and trajectories with a root mean squared distance less than 0.6 μm are grouped as connected components in a graph. Model training is performed using a tenfold grouped cross-validation (validation and test size are 10% each) to ensure similar trajectories are in the same fold while saving the model with highest average recall on validation set.

Task-specific downstream classifier for prediction exclusively from diffusional characteristics

In all cases, classifiers receive a fixed-length representation of trajectories and outputs a probability estimate per class. Fixed-length representations are constructed from raw trajectories by using the temporal behavior segmentation and diffusional fingerprinting modules, totaling 40 descriptive diffusional features (Methods). Filtering using a confidence threshold on outputted probability estimates is performed by requiring estimated probabilities to be larger than the given threshold, otherwise trajectories are considered to be predicted as 'unknown'. Prediction of two simulated populations (Fig. 2e) consists of a simple logistic regression model from Scikit-learn[61] using 'lbfgs' as solver, with an allowed number of iterations at 10,000 evaluated in a fivefold stratified cross-validation with a test size of 10%. Prediction of EEA1-positive and NPC1-positive compartments (Fig. 4) is performed using a simple multilayer perceptron model consisting of an input layer (size 40) and output layer (size 2) with SoftMax activation. Training was performed with random oversampling of the minority class to mitigate majority class bias and evaluation in a tenfold stratified cross-validation. Prediction of EEA1-positive and NPC1-positive compartments using viral cargo (Fig. 4) is performed using the same multilayer perceptron model trained in a 90%/10% train/validation split with minority class oversampling, saving the model with highest validation accuracy. Prediction of cellular localization of AP2 (Fig. 4) is performed exactly as for EEA1-positive and NPC1-positive compartments.

Temporally persistent distance-based colocalization analysis

Colocalization is defined based on temporally consistent proximity between trajectories across acquisition channels. For each trajectory of interest in each imaging channel, the lock-step Euclidean distances are computed to trajectories in the secondary channel of interest. A minimum number of consecutive frames within a user-defined search distance threshold is required to be defined as a colocalizing segment. To mitigate spurious peaks in distance between two trajectories interrupting a true colocalization segment, a certain number of frames is allowed above the given distance threshold ('forgiveness'), a certain number of frames and colocalizing segments on either side will be linked. To simultaneously increase certainty in registered colocalization segments and mitigate registering transient, spurious colocalization, multiple filters are added on top of 'distthreshold', 'min_coloc' and the 'foregiveness': minimum total colocalization length, minimum average lock-step Euclidean distance and a minimum Pearson correlation between individual coordinates axes were enforced.

Colocalizing rotavirus and endosomes were identified using a minimum number of consecutive frames of 5, a search distance threshold of 750 nm purposely sat high to account for any interchannel aberration, a forgiveness of 5 frames, a Pearson correlation threshold of 0.8, a minimum total colocalizing length of 5 frames and a minimum average lock-step Euclidean distance of 750 nm.

Colocalizing rotavirus VP4 to VP7 signal was done by initially correcting chromatic aberrations by identifying long-lived colocalization, computing *xyz* chromatic shift between VP4 and VP7 signal and shifting all VP7 tracks by their average *xyz* offset. Initial parameters were the minimum number of consecutive frames of 5, search distance threshold of 750 nm, forgiveness of 3 frames, Pearson correlation threshold of 0.9, minimum total colocalizing length of 20 frames and a minimum average lock-step Euclidean distance of 500 nm. Following correction of chromatic offset, colocalizing rotavirus VP4 to VP7 signal was performed using the minimum number of consecutive frames of 3, a search distance threshold of 400 nm, forgiveness of 2 frames, Pearson correlation threshold of 0.9, minimum total colocalizing length of 20 frames and a minimum average lock-step Euclidean distance of 600 nm.

Inferring AP2 position relative to coverslip by 3D plane fitting

AP2 coordinates in a 3D space were obtained by LLSM[3,62]. These coordinates were rotated 30° around the y axis of the LLSM imaging direction[3,62] to account for the detection angle of the LLSM by the DP of coordinates with the standard rotation matrix $M = ((\cos(\theta), 0, \sin(\theta)), (0, 1, 0), (-\sin(\theta), 0, \cos(\theta)))$, where θ is the rotation angle in radians. Rotated coordinates point the cell's ventral side in positive z-direction. Utilizing AP2 generally localizes to the plasma membrane, rotated xy coordinates were binned (all bins left inclusive) in a grid size of 5 µm and for each bin the lowest z coordinate was extracted representing the most dorsal AP2 coordinates. To account for outliers in the dorsal z coordinates, the Mahalanobis distance (using mean and covariance of all dorsal z coordinates) was calculated for each dorsal z coordinate, filtering coordinates with distance of 1.8 or above. The resulting dorsal z coordinates were used to fit parameters of a 3D plane by minimizing the sum squared distance between dorsal z coordinates and the plane, resulting in an inferred coverslip position. All AP2 coordinates had their distance calculated to the resulting plane.

Statistical tests

The comparison of results in Fig. 2e and Fig. 4i were performed using a two-sided Welch's t-test to evaluate the null hypothesis that the conditions in question have equal means. The Welch's t-test was chosen due to its strength as a location test and its robustness to populations with unequal variance and sample sizes. Welch's t-test was implemented in SciPy[58], utilizing the Welch–Satterthwaite equation to compute degrees of freedom (Supplementary Tables 1 and 2).

Viral and endosomal labeling for LLSM imaging

Cells with gene-edited early endosomal antigen 1 with mScarlett (EEA1-mScarlett) and a Halo-tagged version of the cholesterol transporter Niemann Pick C1 with JFX646 (NPC1-Halo-JFX646) were thawed samples from frozen aliquots generated in the Kirchhausen laboratory by Kang et al.[63]. The cells were originally from the American Type Culture Collection (CRL-8621). Cells with the clathrin adapter complex, AP2 gene edited at its sigma subunit with eGFP were samples thawed from Cocucci et al.[55]. For imaging, SVG-A cells were plated onto coverslips with a diameter of 5 mm inside a 35 mm culture dish at ~50% confluency the day before each experiment. Cells were incubated with 10 µl of labeled rotavirus particles at ~40 µg ml^{-1} and a multiplicty of infection of ~10 for 10 min before being moved directly to the microscope. Cells were imaged in phenol red-free media (FluoroBrite DMEM, 25 mM HEPES and 1% PenStrip) and a soluble fluorescent dye was added to the media (either Alexa Fluor647 or Alexa Fluor488 carboxylic acid). Experiments without rotavirus use FluoroBrite DMEM, 25 mM HEPES and 1% PenStrip with 5% FBS. For virus labeling, the triple-layer particles (TLPs) were diluted to 0.4 mg ml^{-1} in a total volume of 50 µl using HNC (20 mM HEPES pH 8.0, 100 mM NaCl and 1 mM $CaCl_2$) and 5.5 µl of 1 M $NaHCO_3$ (pH 8.3) was added. This solution was mixed with 0.5 µl of 0.76 mg ml^{-1} Atto488 NHS ester. The reaction proceeded at room temperature for 1 h before quenching it with 5 µl of 1 M Tris pH 8.0. The labeled TLPs were then buffer exchanged into a solution containing 20 mM Tris pH 8.0, 100 mM NaCl and 1 mM $CaCl_2$ using a Zeba Spin Desalting Column. For recoated particle formation and labeling, TLPs, DLPs, VP7 and VP4 were purified[51]. VP7 and VP4 were expressed in Sf9 cells infected with a baculovirus vector. VP7 was purified by successive affinity chromatography on concanavalin A and monoclonal antibody (mAb159), specific for the VP7 trimer[51] (elution by EDTA). Purified VP7 was desalted into a solution containing 2 mM HEPES (pH 7.5), 10 mM NaCl and 0.1 mM $CaCl_2$ (0.1 HNC). For VP4, collected cells were lysed by freeze-thawing and clarified by centrifugation at 2,900*g* after the addition of a completely EDTA-free protease inhibitor (Roche). VP4 was precipitated by the addition of ammonium sulfate to 30% saturation, pelleted, and resuspended in a solution containing 20 mM Tris (pH 8.0) and 1 mM EDTA, and then loaded onto a HiTrap Q column (GE Healthcare), and eluted in a gradient of 10 to 150 mM NaCl. Pooled fractions containing VP4 were dialyzed overnight in a 20 mM Tris (pH 8.0), 100 mM NaCl and 1 mM EDTA buffer.

VP7 and DLP were labeled as previously described[50]. VP7 was brought to 1.7 mg ml^{-1} in a total volume of 100 µl using 0.1 HNC and 11.1 µl of 1 M $NaHCO_3$ (pH 8.3) was added. This solution was mixed into 0.71 µl of 0.76 mg ml^{-1} Atto565-NHS ester. The reaction proceeded at room temperature for 1 h before quenching with 10 µl of 1 M Tris (pH 8.0). The labeled VP7 was then desalted into a solution containing 2 mM Tris (pH 8.0), 10 mM NaCl and 0.1 mM $CaCl_2$ (0.1 TNC). Then, 50 µg DLP was brought to a volume of 100 µl in HN, to which 11.1 µl 1 M $NaHCO_3$ (pH 8.3) was added. This solution was then added to 1.5 µl of 0.5 mg ml^{-1} of Atto647N-NHS ester. The reaction proceeded 1 h and room temperature before quenching with 10 µl of 1 M Tris (pH 8.0). The sample was then desalted through a 0.5 ml Zeba Spin desalting column into a solution containing 20 mM Tris (pH 8.0) and 100 mM NaCl (TN). We distributed 45 µg of DLPs in HNE equally among five 1.5 ml conical tubes (0.5 µl per tube). We first added 1 M sodium acetate pH 5.2 to a final concentration of 100 mM and then added 82 µl VP4 (stored at 1.8 mg ml^{-1}) to a final concentration of 0.9 mg ml^{-1} in the final reaction volume, resulting in a 33-fold excess of VP4 monomer over a total of 180 sites on DLPs. A 0.1 mg ml^{-1} aprotinin solution was added to the samples to a final concentration of 0.2 µg ml^{-1} followed by incubation at 37 °C for 1 h. Required amounts of VP7 (7.14 µl stored at 1.26 mg ml^{-1} in HNE) to achieve a 2.3-fold excess of VP7 monomer over a total of 780 sites on DLPs were premixed with 0.1 volumes of TC buffer (20 mM Tris, 10 mM $CaCl_2$, pH 8.0) and 0.1 volumes of 1 M sodium acetate pH 5.2 for 15 min before adding them to the DLP-VP4 mixture. Samples were incubated for 1 h at room temperature and then quenched by the addition of 0.1 volumes of 1 M Tris pH 8.0. Recoated particles from the five tubes were combined, and TNC was added to a final volume of 2.5 ml. Recombinant TLPs (rcTLPs) were separated from excess VP4 and VP7 by ultracentrifugation at 4 °C in a Beckman Coulter rotor TLS 55 at 215,000*g* for 1 h. We removed 2.0 ml of the supernatant, returned the volume to 2.5 ml with TNC and pelleted again. Supernatant was carefully removed so that 100–200 µl remained. The rcTLP pellets were resuspended in the remaining buffer and stored at 4 °C.

Insulin labeling for SDCM imaging

HI was labeled with Atto-655-NHS ester. HI Atto-655-NHS (LysB29Atto-655-HI) ester was prepared following previous publications[64,65].

In short, HI (21 mg, 0.0036 mmol, 3 equiv.) was dissolved in 0.1 M Tris buffer (0.2 ml) with pH adjusted to 10.5 for complete dissolution, Atto-655-NHS ester (1.0 mg, 0.00122 mmol, 1.0 equiv.) in DMF (0.3 ml) was added by drops over 5 min to the HI solution, followed by stirring for 15 min. The reaction was monitored by liquid chromatography–mass spectrometry. Subsequently, the reaction mixture was diluted by 2.0 ml of milliQ water and pH adjusted to 7.8. The product was isolated using reversed-phase high-performance flash chromatography, using a Biotage SNAP ultracolumn (C18, 30 g, 25 µm). CH_3CN/H_2O mixed with 0.1% formic acid was used as eluents at a linear gradient of 5–50% CH3CN over 20 min at a flow rate of 25 ml min^{-1}. Each fraction was analyzed through liquid chromatography–mass spectrometry. Monosubstituted products were collected separately, CH_3CN was removed at reduced pressure using a rotary evaporator, followed by lyophilization (LysB29Atto-655-HI yield: 5.6 mg, 79%).

LLSM imaging, experimental setup and SPT

Rotavirus, EEA1-mScarlett. NPC1-Halo-JFX646 and AP2 tracking experiments were carried out using an in-house built LLSM as in previously published work[3,5]. The LLSM imaging ran in sample scan mode with 0.25 µm spacing between each plane along the *z*-imaging axis producing molecular videos consisting of 3D volumes using a dithered multi-Bessel lattice light-sheet illumination. The exposure times and frame rates are listed per experiment in Supplementary Table 4. The resulting *z*-stacks were de-skewed, followed by detection and linking across frames of punctuate light point sources producing sets of *xyzt* trajectories using an automated tracking algorithm implemented in MATLAB based on u-track[9], involving least-squares numerical fitting of a 3D Gaussian, as previously described[5,9,66]. The used settings are identical to the original settings from u-track[9] with the exception of the search radius lower limit set to 3 pixels and search radius upper limit set to 6 pixels. To assess localization error in the LLSM, we generated simulated point spread functions (PSFs) replicating the experimental PSF in *x*, *y* and *z* dimensions. This was achieved by fitting a Gaussian to an experimental bead measured within the lattice using the same sample scan settings as those used to collect the experimental data. The simulated PSFs were overlaid with background noise identical to that observed in experiments, and 100 frames of the *z*-stack were generated with varying background intensities to reproduce realistic noise fluctuations. This process was repeated for five different PSF intensities. Finally, the simulated data were analyzed using the point source detection method[9] and compared against the known ground truth.

SDCM imaging and SPT

An inverted spinning disk confocal microscope (SDCM) (Olympus SpinSR10, Olympus) was used for all 2D imaging of insulin. The SDCM uses an oil immersion 60× objective (Olympus) and numerical aperture of 1.4 connected to a complementary metal oxide semiconductor (CMOS) camera (photometrics PRIME 95B) with an effective pixel size of 183 nm × 183 nm. Before imaging ~10,000–20,000 HeLa cells were grown at 37 °C, 5% CO_2 in Ibidi IbiTreat 8-well plates for 2 days. LysoTracker Green DND-26 using commercial stock concentration diluted 1:20,000 was added to incubate for 1 h (37 °C, 5% CO_2) before imaging. Last, for imaging insulin, 0.05 mg ml^{-1} LysB29Atto-655-HI was added to the HeLa cells to incubate for 1 h (37 °C, 5% CO_2) before imaging. Before any SDCM imaging experiments, wells were washed three times with fresh, preheated 10 mM HEPES in HBSS buffer. For simultaneous SPT of insulin and compartments, dual imaging was performed using lasers of 640 nm and 488 nm. SDCM live-cell imaging was performed with 30.4 ms exposure and 3 EM gain in SDCM streaming settings, resulting in 36 ms between frames including lag time. The insulin was recorded with laser power 100% with a 640 nm laser, and compartments were recorded with a laser power of 10% for LysoTracker Green DND-26 for a total of 2,000 frames at 37 °C. For tracking of LysB29Atto-655-HI in-house tracking scripts[6] were used based on Trackpy[56] with an object diameter of 9 pixels, search range of 5 pixels and gap closing of 1 frame with mean-multiplier manually evaluated between 0.6 and 1. Postprocessing of tracks was done using thresholds of eccentricity (ecc) <0.3, intensity >0 and duration >20, in addition a logistic regression model trained to differentiate detection made in videos with/without insulin using detection features directly from tracking script was used to further filter detections. To assess localization error in the SDCM, a series of 100 fluorescence microscopy images, each capturing diffraction-limited particles, was analyzed using the Crocker–Grier algorithm to track particle spatiotemporal localizations. For each particle detected over time, a 2D Gaussian function was fitted to the PSF using the centroid position obtained from the Crocker–Grier algorithm as the initial guess. The fitting was performed within a region of interest of 10 × 10 pixels. The standard error, or error on the mean localization, represents the localization error, and is influenced by the signal-to-noise ratio. Assuming a global noise term, this error is directly dependent on the brightness of the background-subtracted particle signal. Consequently, the localization error is not uniform across particles and should be considered relative to particle brightness. To assess this, we categorized the particles into brightness quantiles and reported both the overall localization error and the error within each quantile.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All data are available via the University of Copenhagen repository at https://erda.ku.dk/archives/804ea1ea88f340b79ada3e57141a6d6e/published-archive.html (ref. 67). All biological materials have no restrictions and are available upon reasonable request. We refer to the original publications for the additional experimental data used in this work: refs. 10,26,68.

Code availability

A minimal repository of code is available at https://erda.ku.dk/archives/752e4b0695c0dd16ec3c1a130f6ac70b/published-archive.html (ref. 69). The repository of code and models is available at https://erda.ku.dk/archives/4c5adaaacc5c867f6450bcf89ec55a45/published-archive.html (ref. 70). Both are under a CC BY-NC-ND 4.0 DEED license. In addition, upon publication, the code will also be freely available on GitHub.

References

58. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).
59. Ronneberger, O., Fischer, P. & Brox, T. U-Net: convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)* Vol. 9351 (eds Navab, N., Hornegger, J., Wells, W. M. & Frangi, A. F.) 234–241 (Springer, 2015).
60. Akiba, T., Sano, S., Yanase, T., Ohta, T. & Koyama, M. Optuna: a next-generation hyperparameter optimization framework. In *Proc. 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (eds Teredesai, V. K. et al.) 2623–2631 (ACM Press, 2019); https://doi.org/10.1145/3292500.3330701
61. Pedregosa, F. et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
62. Chen, B.-C. et al. Lattice light-sheet microscopy: imaging molecules to embryos at high spatiotemporal resolution. *Science* **346**, 1257998 (2014).
63. Kang, Y.-L. et al. Inhibition of PIKfyve kinase prevents infection by Zaire ebolavirus and SARS-CoV-2. *Proc. Natl Acad. Sci. USA* **117**, 20803–20813 (2020).

64. Bohr, F. et al. Enhanced hexamerization of insulin via assembly pathway rerouting revealed by single particle studies. *Commun. Biol.* **6**, 178 (2023).
65. Østergaard, M., Mishra, N. K. & Jensen, K. J. The ABC of insulin: the organic chemistry of a small protein. *Chem. Eur. J.* **26**, 8341–8357 (2020).
66. He, K. et al. Dynamics of phosphoinositide conversion in clathrin-mediated endocytic traffic. *Nature* **552**, 410–414 (2017).
67. Kæstel-Hansen, J. DeepSPT data and models. *University of Copenhagen* https://doi.org/10.17894/ucph.75da99a5-f7f1-44e7-bb6e-3fcc97bf0a15 (2024).
68. Bohr, S. S.-R. et al. Direct observation of *Thermomyces lanuginosus* lipase diffusional states by single particle tracking and their remodeling by mutations and inhibition. *Sci. Rep.* **9**, 16169 (2019).
69. Kæstel-Hansen, J. DeepSPT code. *University of Copenhagen* https://doi.org/10.17894/ucph.927973c8-6821-49c8-abae-6cd5996f1c47 (2024).
70. Kæstel-Hansen, J. DeepSPT code and models. *University of Copenhagen* https://doi.org/10.17894/ucph.25800387-29f5-4815-ae49-9a9d4d063bc4 (2024).

Acknowledgements

We thank members of our laboratories for help and encouragement. We thank S. C. Harrison for the fruitful discussions on the rotavirus data interpretation. We thank N. Kumar Mishra and K. J. Jensen for discussion concerning the preparation and use of insulin for our experiments. This work was funded by the Villum Foundation by being part of BioNEC (grant 18333) to J.K.-H. and N.S.H., the Novo Nordisk foundation challenge center for Optimised Oligo escape (NNF23OC0081287) (N.S.H. PI and T.K., co-PI), the center for 4D cellular dynamics (NNF22OC0075851). N.S.H. is affiliated with The Novo Nordisk Foundation Center for Protein Research funded by a generous donation from the Novo Nordisk Foundation (grant no. NNF14CC0001). J.K.-H., A.J.N., S.V.B. and N.S.H. are members of the Integrative Structural Biology Cluster at the University of Copenhagen. VILLUM FONDEN (40516) for W.B. and N.S.H. The Novo Nordisk Foundation Center for Basic Machine Learning Research in Life Science (NNF20OC0062606) for W.B. T.K. acknowledges support from NIH Maximizing Investigators' Research Award GM130386, NIH grant AI163019, 1R01, IONIS Pharmaceuticals. M.d.S was supported by NIH/NCI CA13202 grant to S. C. Harrison.

Author contributions

J.K.-H., N.S.H. and T.K. wrote the paper with feedback from all the authors. J.K.-H. performed all computational work with input from T.K., W.B. and N.S.H. M.d.S. performed all rotavirus preparation. A.S., R.F.B.D.C.C. and G.S. performed LLSM imaging with the presence of J.K.-H. in the laboratory of T.K. at Harvard Medical School. A.J.N. and S.V.B. performed SDCM imaging and insulin assays and K.T. wrote the software package in the laboratory of N.S.H. at the University of Copenhagen. J.K.-H. and N.S.H. conceived the project idea. N.S.H. had the overall project management with tight interactions with T.K.

Competing interests

The authors declare no competing interests.

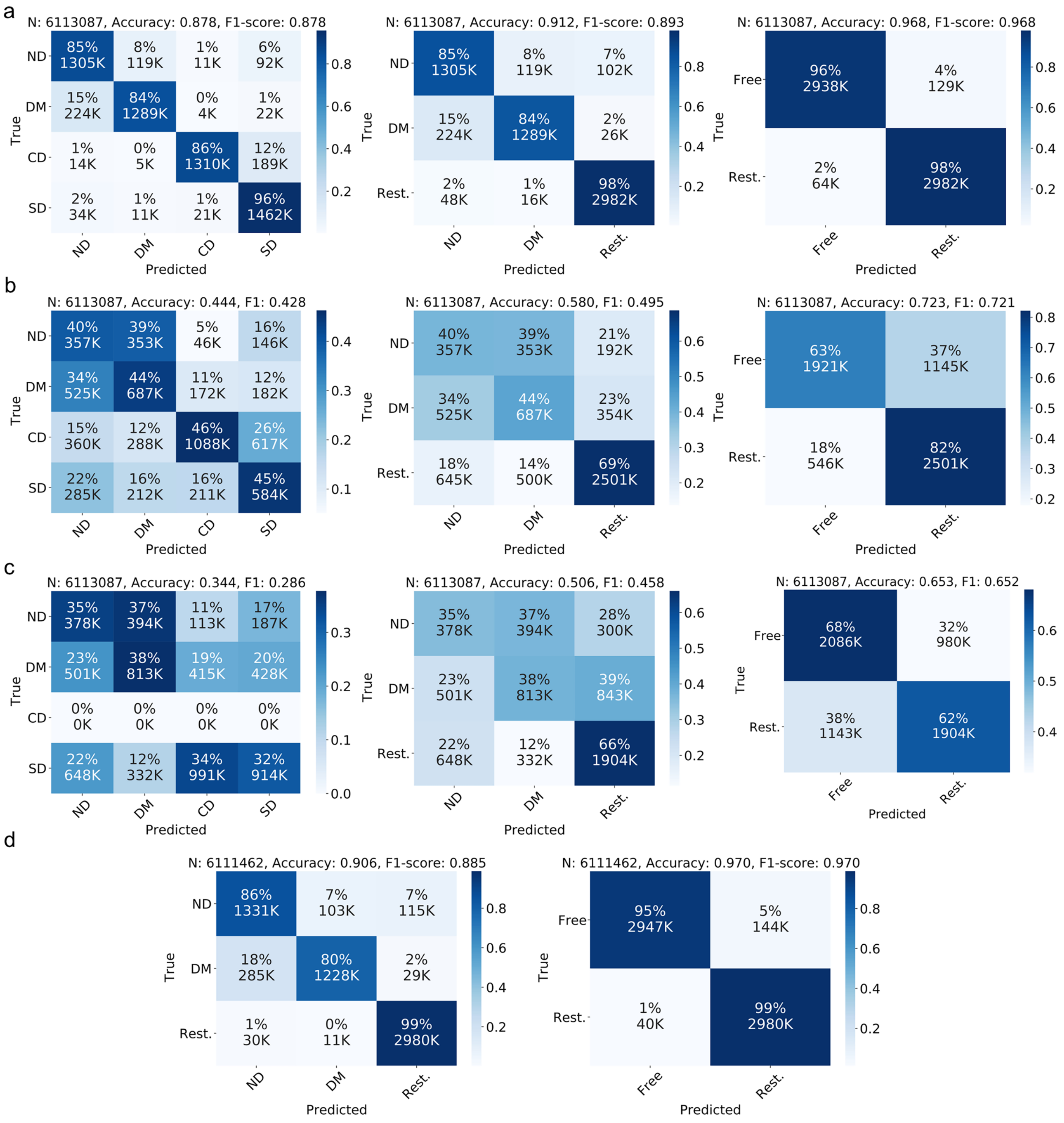
Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41592-025-02665-8.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41592-025-02665-8.

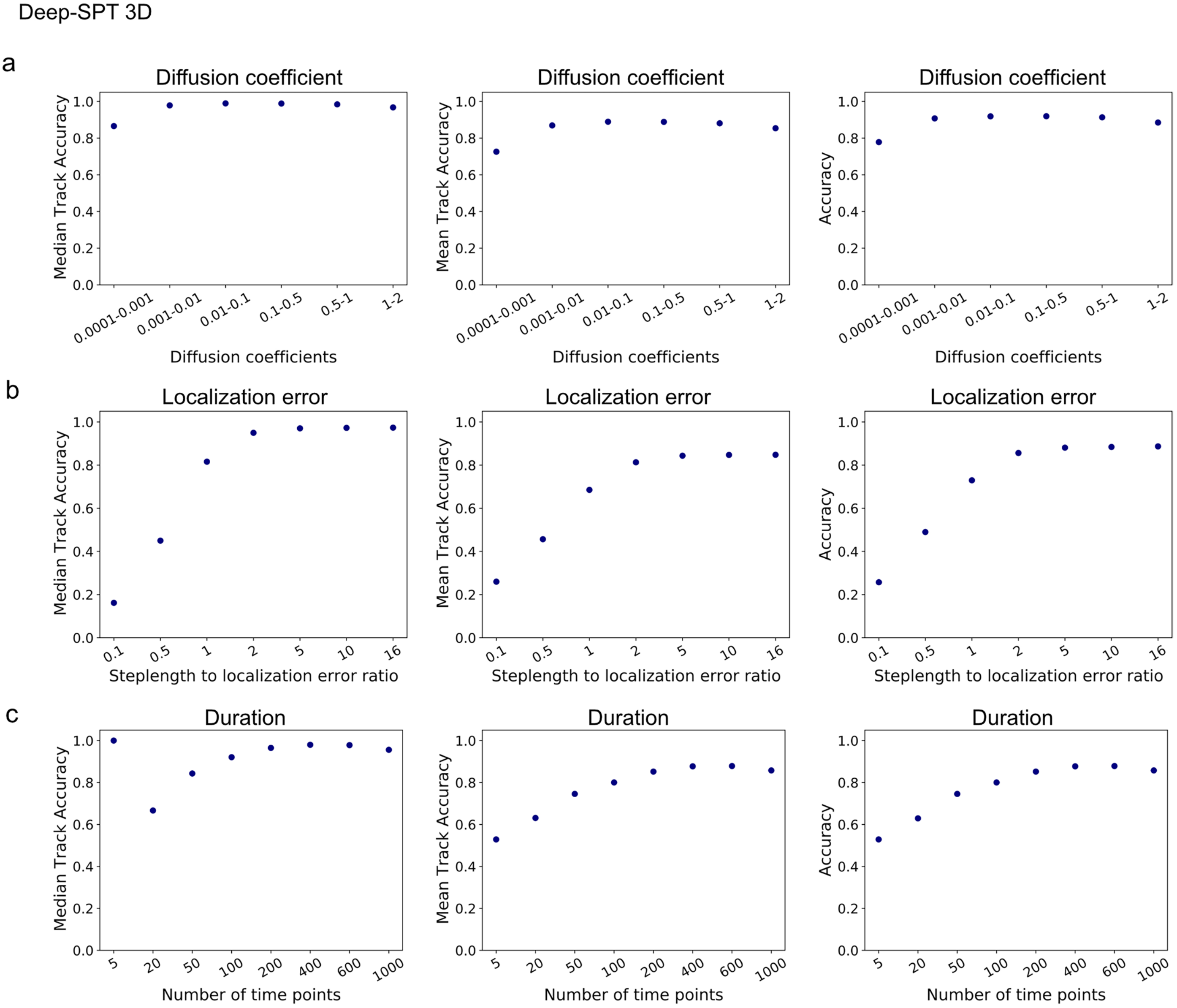
Correspondence and requests for materials should be addressed to Tomas Kirchhausen or Nikos S. Hatzakis.

Peer review information *Nature Methods* thanks Koen Martens and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Rita Strack, in collaboration with the *Nature Methods* team. Peer reviewer reports are available.

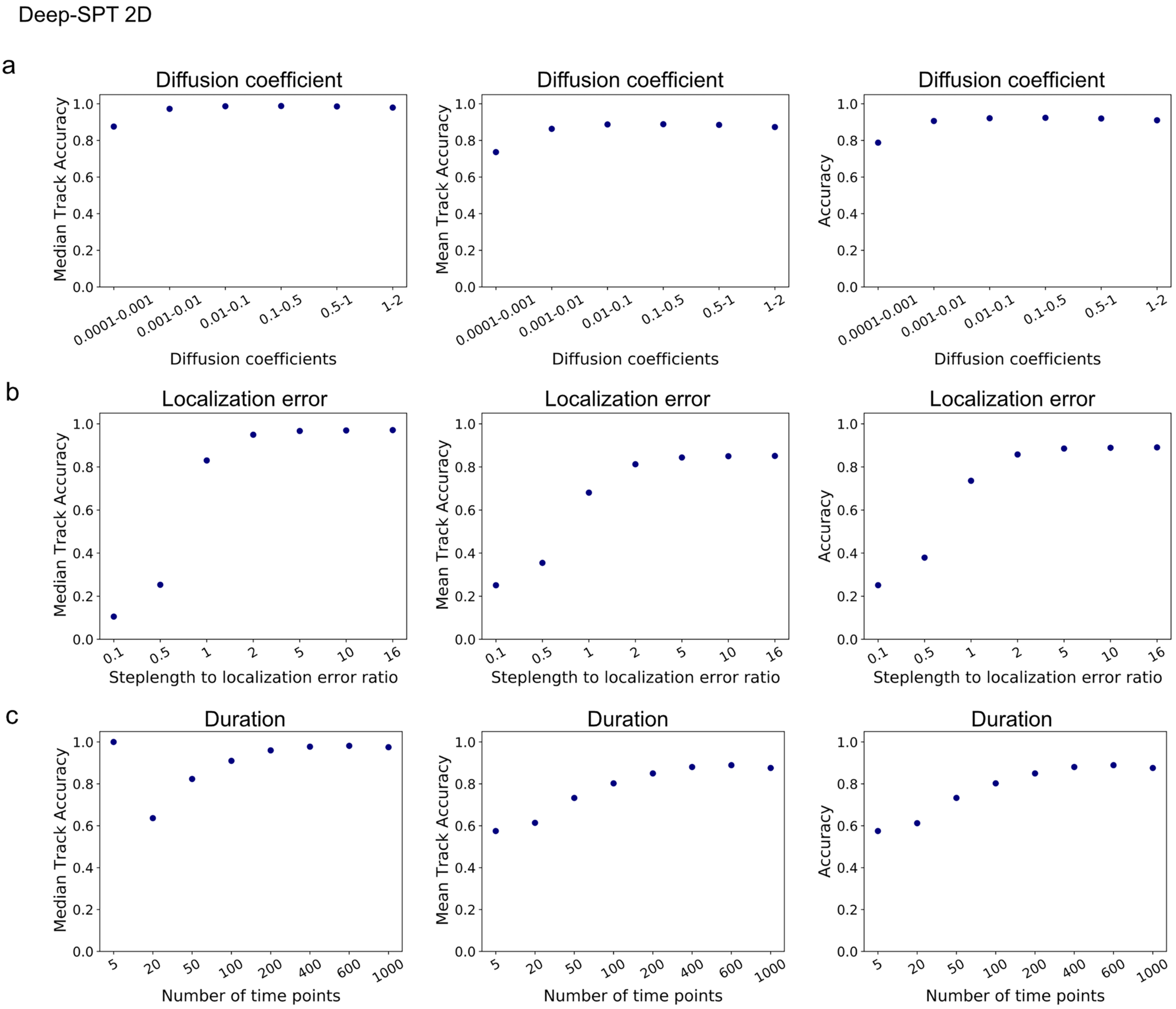


Extended Data Fig. 1 | Comparison of classification accuracy of DeepSPT with benchmarks. Comparison of classification accuracy of DeepSPT, LSTM-based model[29], and rolling MSD for various number of diffusion types. **a**, Confusion matrices of DeepSPT predictions per frame on 2D simulated test set trajectories (N=20000 tracks), Data displayed for either 4 diffusional states, or 3 states normal, directed and confined/subdiffusive, or 2 states normal/directed, confined/subdiffusive. **b**, Confusion matrices of attention BiLSTM[29] predictions per frame on 2D simulated test set trajectories. Data for for 4, 3, 2 diffusional states as above **c**, Confusion matrices of rolling MSD[12,47] predictions per frame on 2D simulated test set trajectories (predicts three classes as it bases predictions on the alpha exponent in a MSD fit based on (subdiffusive alpha)<(normal alpha)<(directed alpha)) Data for for 4, 3, 2 diffusional states as above. **d**, DeepSPT on 3D test set trajectories when combining classes into three and two respectively.

https://doi.org/10.1038/s41592-025-02665-8

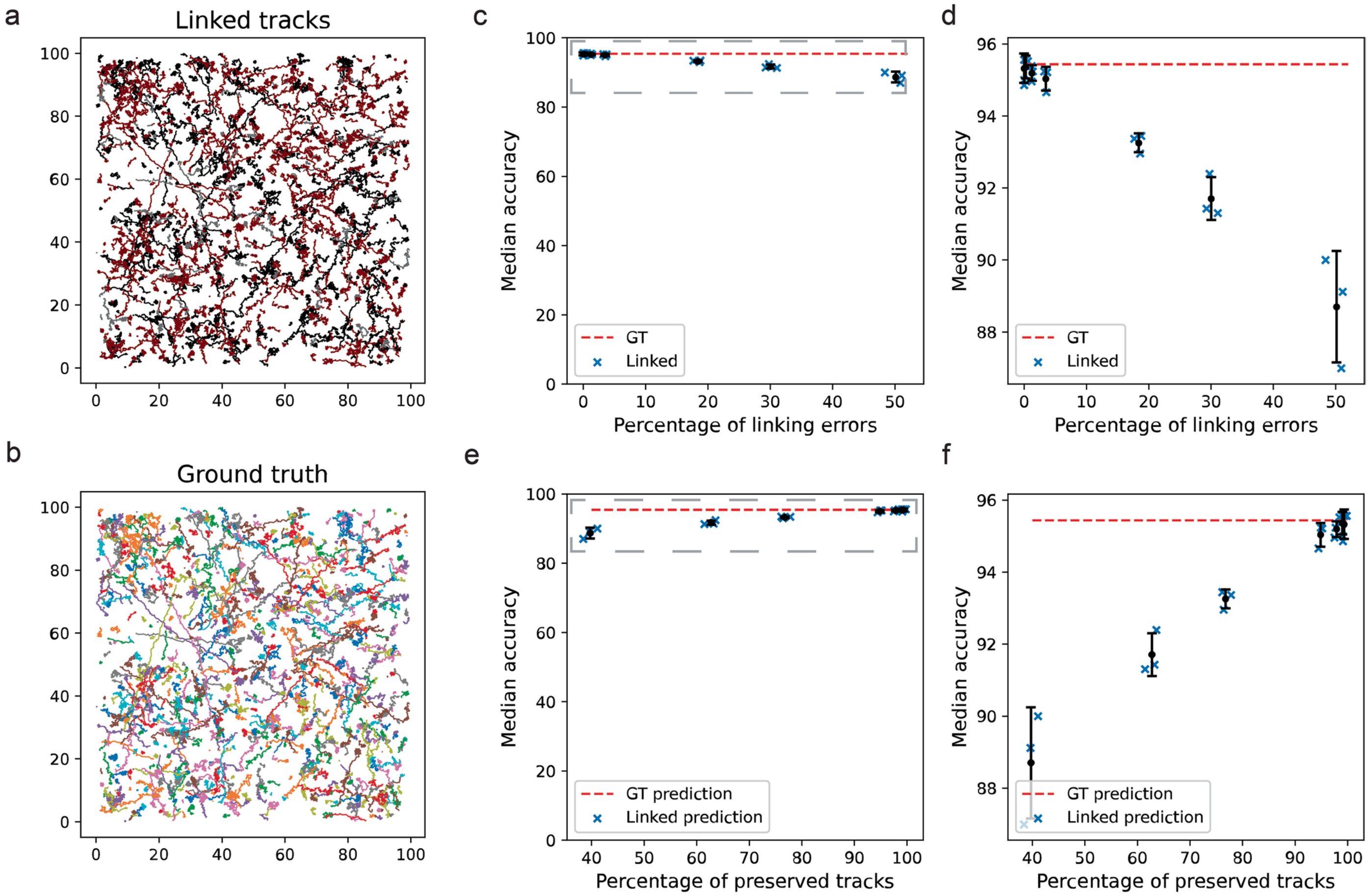


Extended Data Fig. 2 | Investigating DeepSPT's robustness to simulation parameters for the heterogeneous diffusing test set in 3D. Investigation of model generalizability and limitations for key simulation parameters for traces containing all 4 diffusional types at random (see Methods for test set elaboration): **a**, Varying ranges of diffusion coefficients (D) for simulated trajectories (varying D is equivalent to varying temporal resolution for observation of a diffusing particle due to scale invariance of diffusion). **b**, step length to localization error ratio, i.e ratio of contribution to displacements from actual diffusion and localization error respectively where a value of 1 signifies an equal contribution and >1 signifies actual diffusion being larger. **c**, Duration of track. Median accuracy and mean accuracy are track-level metrics providing descriptive statistics for the distribution track-level accuracies for each test set track (N=20000). Flattened accuracy measures the accuracy for all frame-level predictions inside each trajectory pooled together.

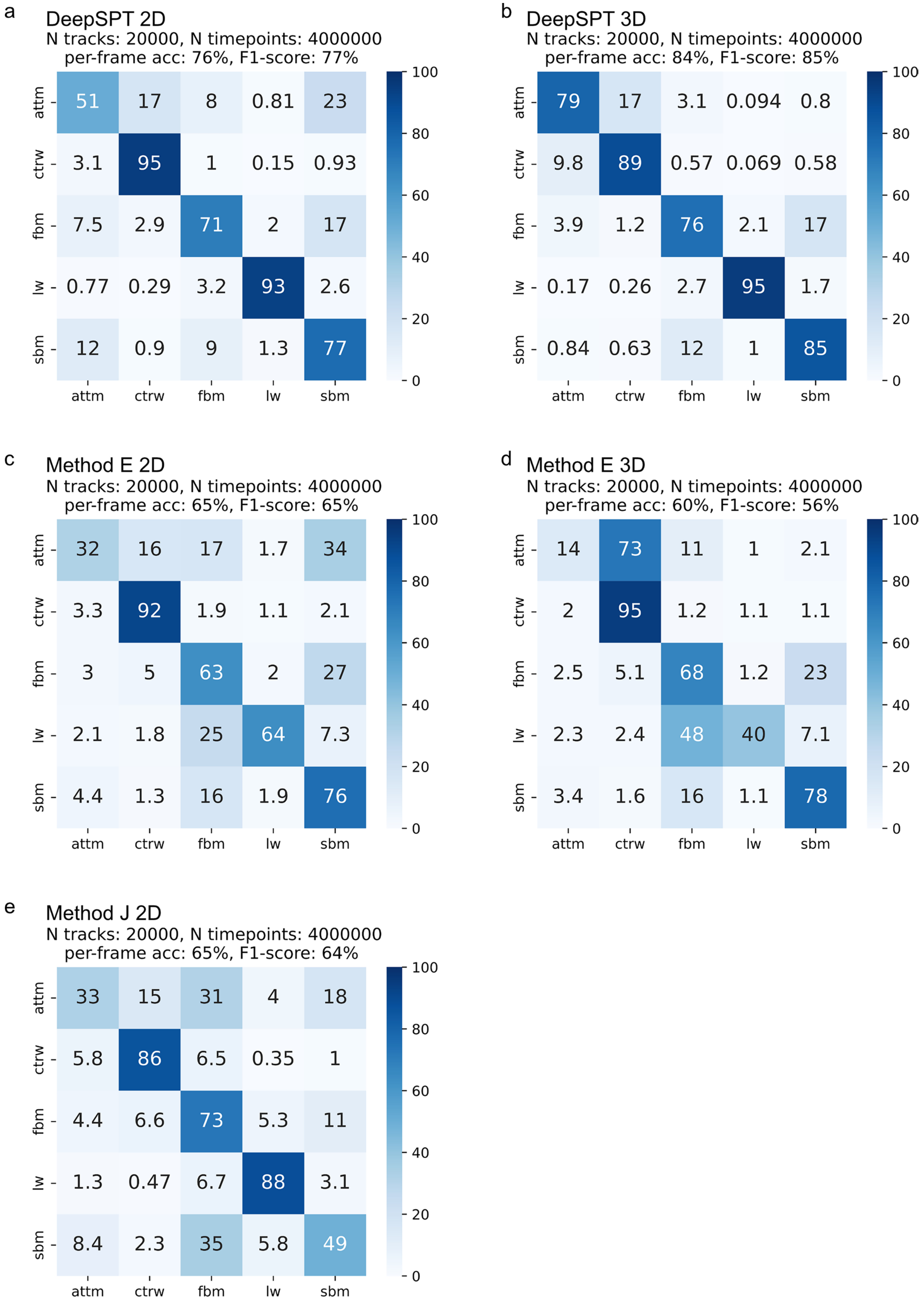


Extended Data Fig. 3 | Investigating DeepSPT's robustness to simulation parameters for the heterogeneous diffusing test set in 2D. Investigation of model generalizability and limitations for key simulation parameters for traces containing all 4 diffusional types at random (see Methods for test set elaboration): **a**, Varying ranges of diffusion coefficients (D) for simulated trajectories (varying D is equivalent to varying temporal resolution for observation of a diffusing particle due to scale invariance of diffusion). **b**, step length to localization error ratio, i.e ratio of contribution to displacements from actual diffusion and localization error respectively where a value of 1 signifies an equal contribution and >1 signifies actual diffusion being larger. **c**, Duration of track. Median accuracy and mean accuracy are track-level metrics providing descriptive statistics for the distribution track-level accuracies for each test set track (N=20000). Flattened accuracy measures the accuracy for all frame-level predictions inside each trajectory pooled together.

https://doi.org/10.1038/s41592-025-02665-8

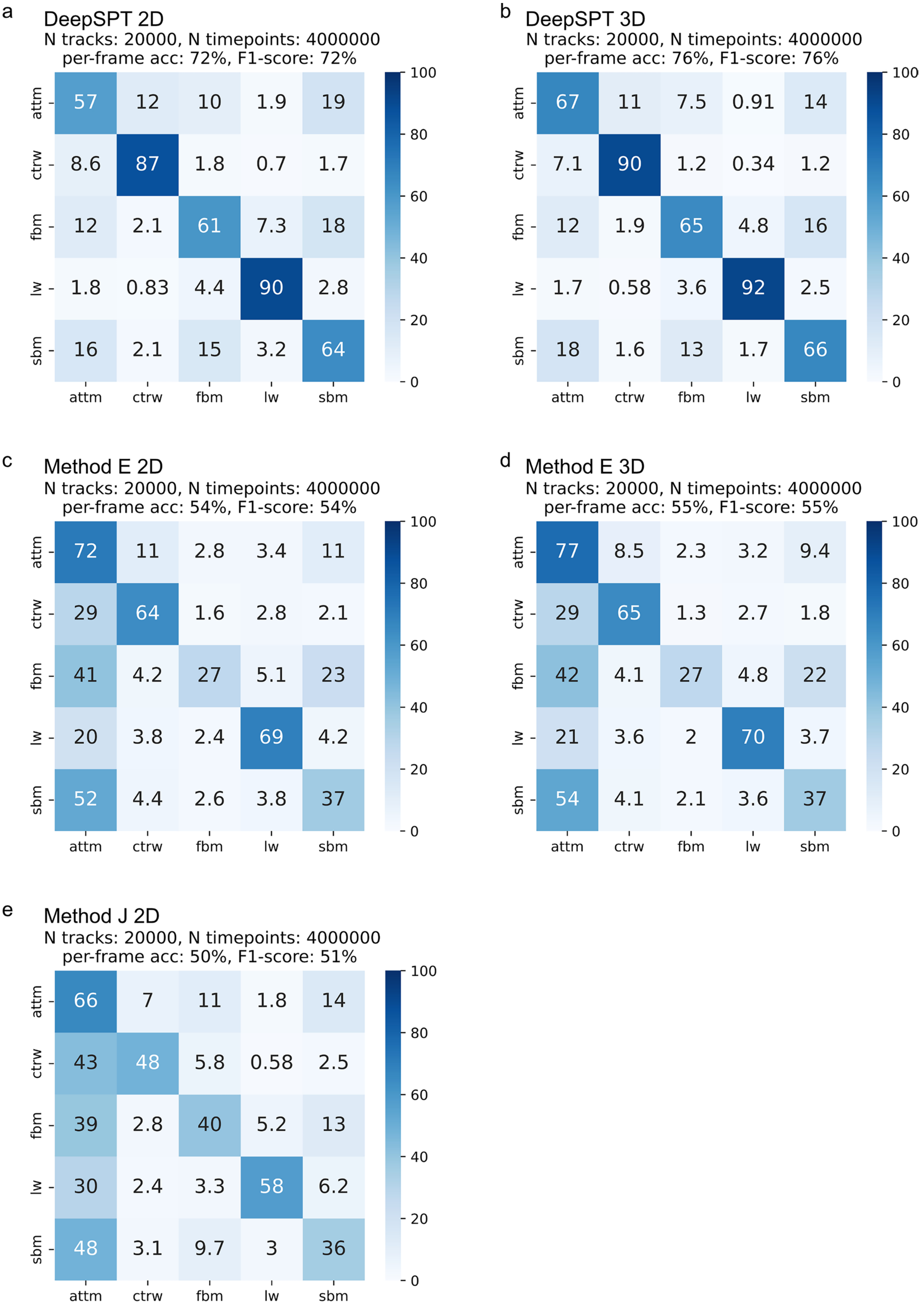


Extended Data Fig. 4 | Evaluation of the effects of tracking errors on DeepSPT temporal segmentation. 1000 trajectories are simulated within a box of varying dimensions and tracked using Trackpy[56] (see Methods). The trajectories obtained by Trackpy are analyzed using DeepSPT's temporal segmentation module. **a**, Example of trajectories produced by Trackpy showing tracks with one or more linking errors (red), correct/preserved tracks (black), and tracks without linking errors but still differ from the simulated tracks (grey). **b**, Example of simulated ground truth tracks (colored by ID). **c**, The median accuracy of DeepSPT in prediction diffusional behavior per time point per track for varying simulated box dimensions, thus varying degrees of linking errors. Median accuracy refers to the median of the distribution of correctly predicted time points in each individual track. Each blue cross represents an independent set of 1000 trajectories. Three sets are simulated per simulated box dimension with black dots and error bars representing the mean and standard deviations across each of the three sets. Red line represents the median accuracy of DeepSPT directly on simulated trajectories. **d**, Zoom-in, showing the dashed grey box in (**c**). **e**, The median accuracy of DeepSPT in prediction diffusional behavior per time point per track versus various degrees of preserved tracks, that is number of tracks perfectly tracked. As in (**c**) each blue cross represents an independent set of 1000 trajectories with three sets per simulated box dimension with black dots and error bars representing the mean and standard deviations across each of the three sets. Red line represents the median accuracy of DeepSPT directly on simulated trajectories. **f**, Zoom-in, showing the dashed grey box in (**e**).



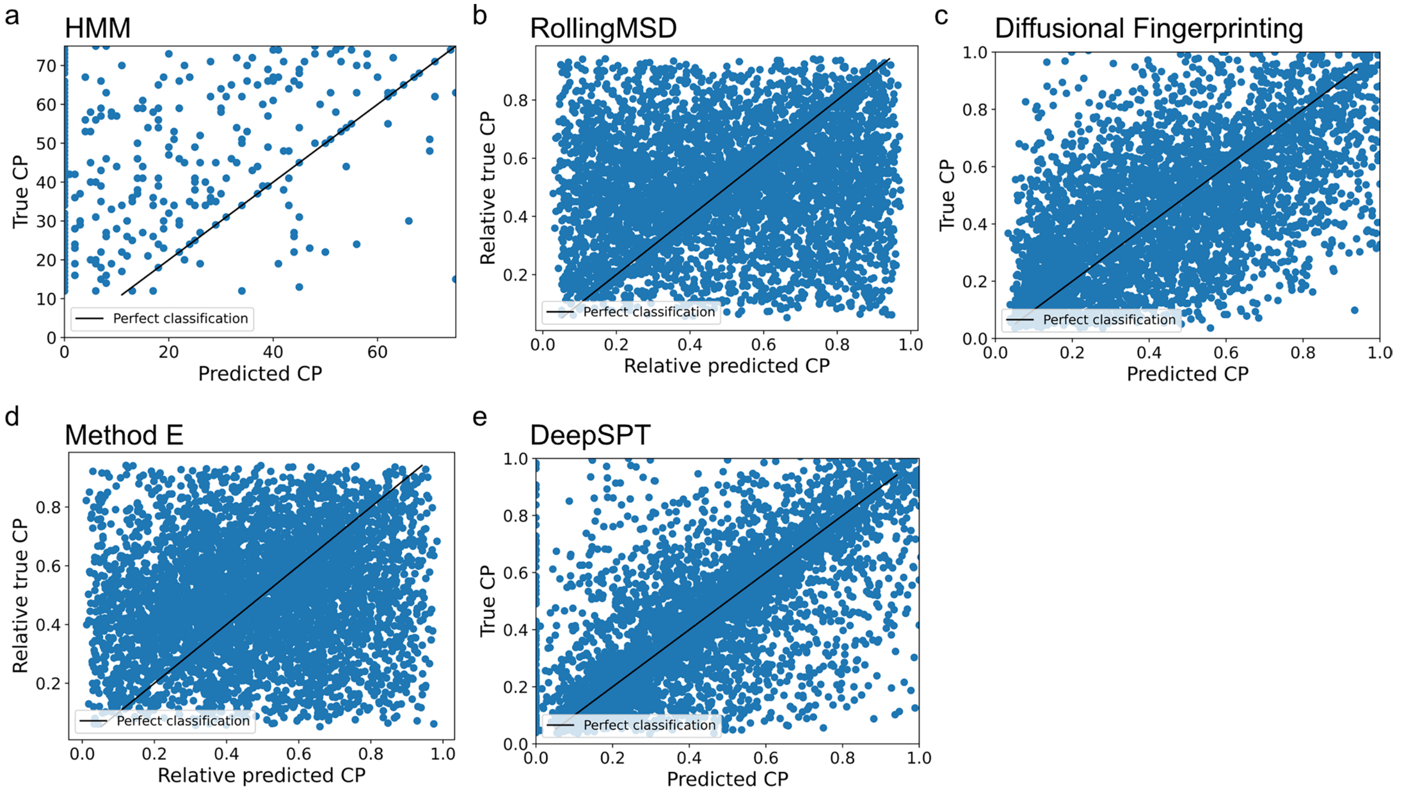
Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | Evaluation of DeepSPT and two AnDi challenge models on AnDi challenge task 3. Confusion matrix for all individual time point predictions within the 20000 2D (**a, c, e**) and 3D (**b, d**) test set trajectories simulated using the 2021 AnDi challenge task 3 open-source framework totalling 4 million predictions. See Muñoz-gil et al.[28] for further test set specification. Diagonal entries are correct predictions and off-diagonal indicates confused classes. Each entry reports the percentage of predictions normalized to the actual number of true labels in the given class. **a**, Confusion matrix for DeepSPT on 2D trajectories. **b**, Confusion matrix for DeepSPT on 3D trajectories. **c**, Confusion matrix for Method E on 2D trajectories. **d**, Confusion matrix for Method E on 3D trajectories. **e**, Confusion matrix for Method J on 2D trajectories.



Extended Data Fig. 6 | See next page for caption.

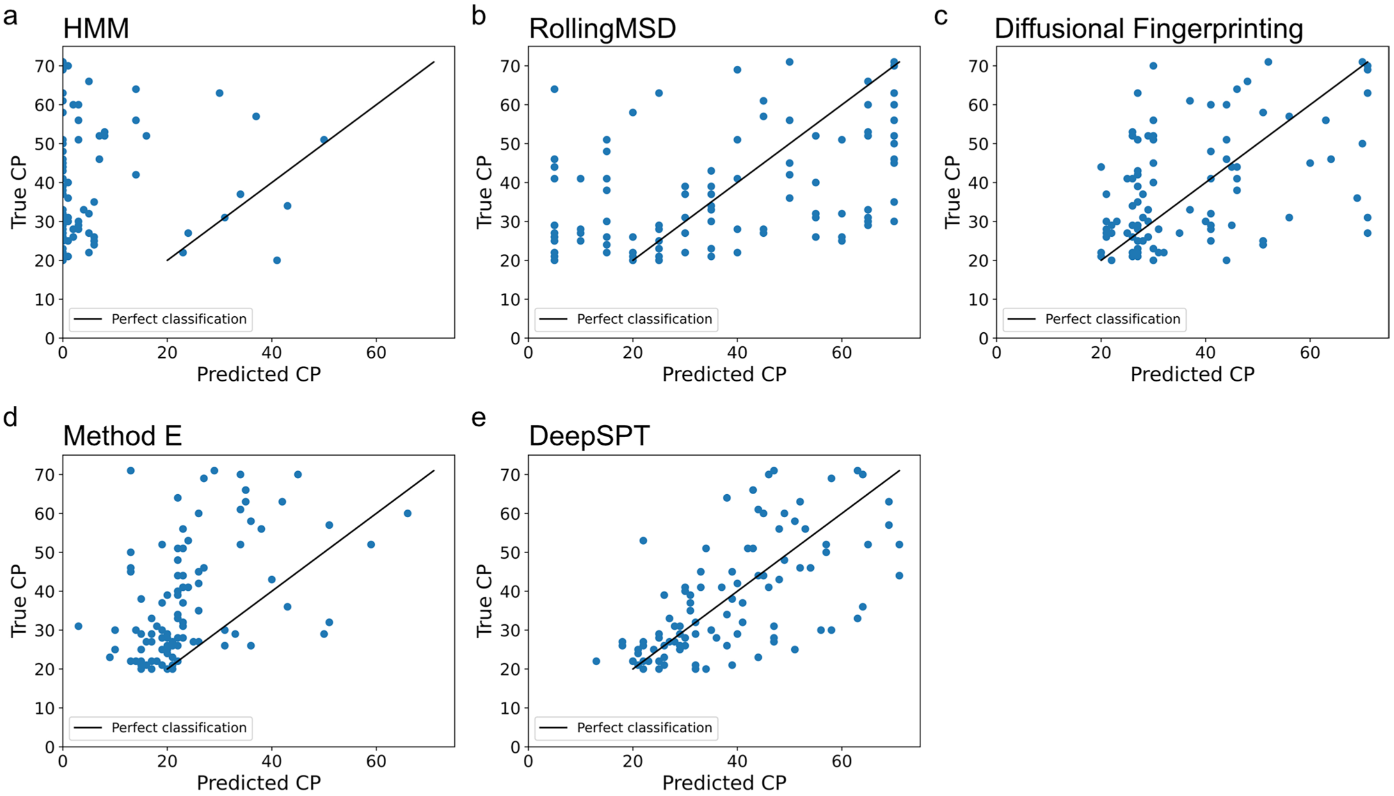
Extended Data Fig. 6 | Evaluation DeepSPT and two AnDi challenge models on traces with multiple changes between diffusion behaviors from the AnDi challenge. Confusion matrix for all individual time point predictions within the 20000 2D (**a, c, e**) and 3D (**b, d**) test set trajectories simulated by combining the anomalous diffusion behaviors from the 2021 AnDi challenge open-source framework[28] into heterogeneous trajectories sampling multiple diffusional behaviors with multiple change points totalling 4 million predictions. Diagonal entries are correct predictions and off-diagonal indicates confused classes. Each entry reports the percentage of predictions normalized to the actual number of true labels in the given class. **a**, Confusion matrix for DeepSPT on 2D trajectories. **b**, Confusion matrix for DeepSPT on 3D trajectories. **c**, Confusion matrix for Method E on 2D trajectories. **d**, Confusion matrix for Method E on 3D trajectories. **e**, Confusion matrix for Method J on 2D trajectories.



f

	HMM	RollingMSD	Diffusional Fingerprinting	Method E	DeepSPT
Median frame error	57	31	30	30	**17**
Mean frame error	72.3	43.6	37.2	37.2	**31.1**
Mean squared error	8372.6	3660.6	2306.2	2306.2	**2207.8**
R^2	-1.91	-0.24	0.22	0.21	**0.25**

Extended Data Fig. 7 | Temporal segmentation and changepoint prediction of simulated trajectories for DeepSPT and benchmark approaches. **a**-**e**, Predicted changepoints (CP) versus true changepoints. Trajectories are constructed by combining two populations with overlapping diffusional properties into individual tracks with one changepoint (see Methods). Black line represents perfect classification. **a**, HMM-bayes[17], 1556 data points compared to 5000 for other approaches due to computational time restraints as HMM-bayes requires several minutes per track. **b**, Rolling MSD. **c**, Original diffusional fingerprinting (Pinholt et al.[13]) **d**, Method E from the 2021 AnDi Challenge. **e**, DeepSPT. **f**, Table of classification metrics for the temporal segmentation of trajectory timepoints as post- and pre-uncoating and changepoint prediction. Median accuracy measures the median accuracy per trajectory (N=100). F1-score measures the F1-score of all individual timepoint predictions. Median frame error measures the median of absolute distances between predicted and true changepoints across all trajectories. Mean frame error measures the mean of absolute distances between predicted and true changepoints across all trajectories. Mean squared error measures the mean of the second norm distances between predicted and true changepoints across all trajectories. R2 measures the coefficient of determination.

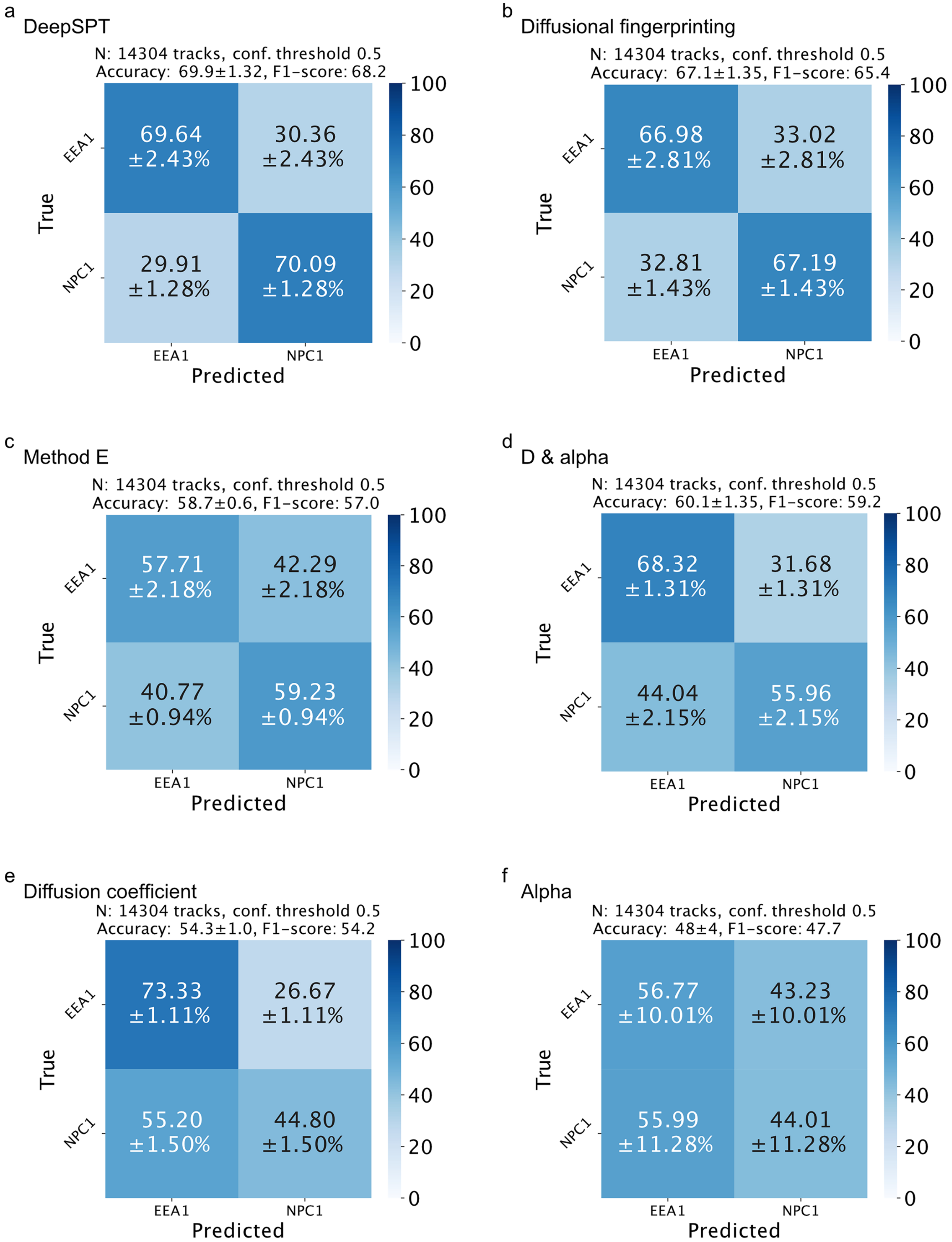


f

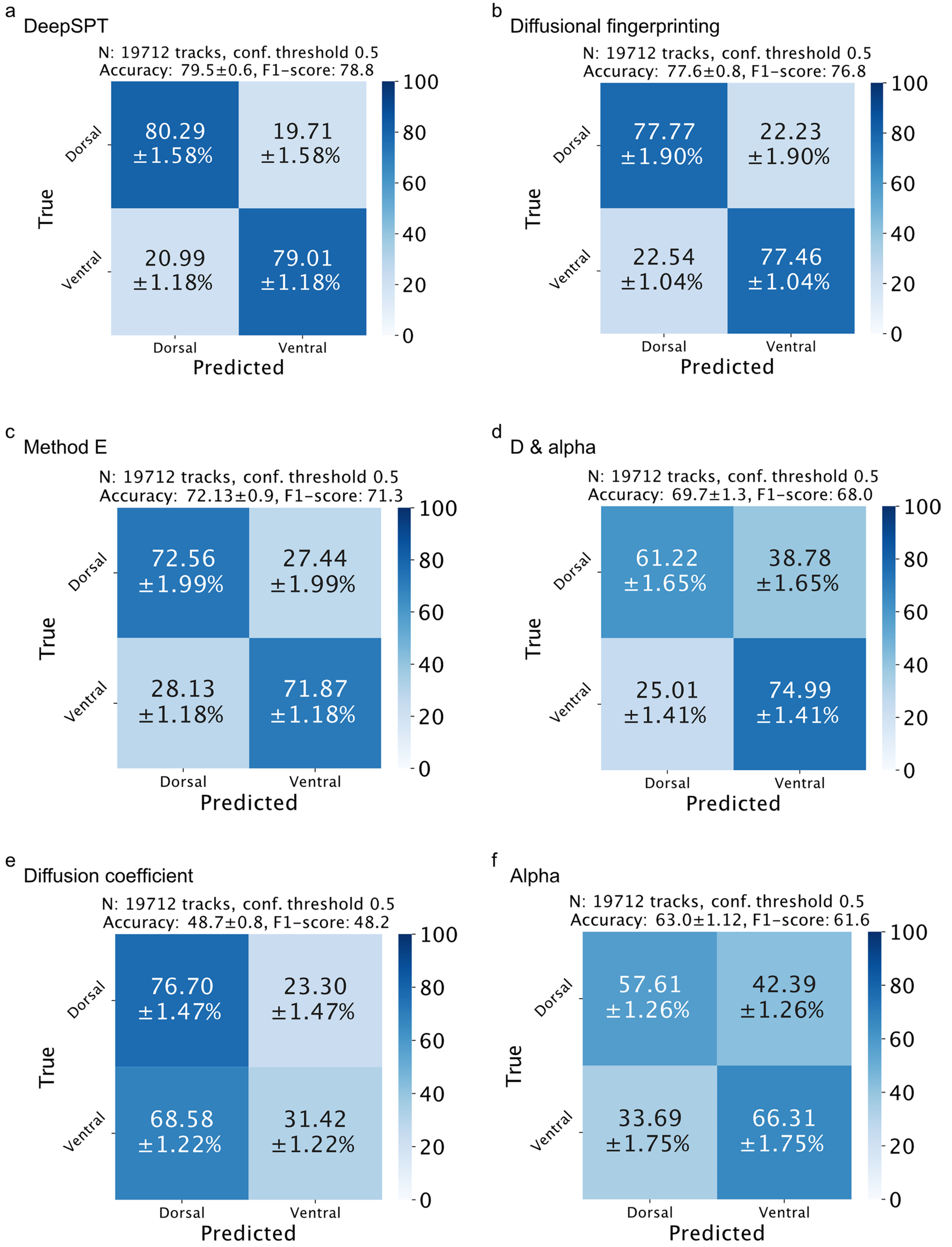
	HMM	RollingMSD	Diffusional Fingerprinting	Method E	DeepSPT
Median accuracy	47%	74%	83%	74%	**88%**
F1-score	59%	60%	73%	71%	**77%**
Median frame error	29	15	8	12.5	**6**
Mean frame error	33.3	16.5	11.2	15.2	**8.9**
Mean squared error	1355.6	438.8	224.4	365.5	**143.7**
R^2	-5.41	-1.07	-0.06	-0.72	**0.32**

Extended Data Fig. 8 | Temporal segmentation and changepoint prediction of 3D lattice light sheet rotavirus trajectory timepoints as post- and pre-uncoating for DeepSPT and benchmark approaches. a-e, Predicted changepoints (CP) versus true changepoints. Black line represents perfect classification. **a**, HMM-bayes[17] **b**, Rolling MSD. **c**, Original diffusional fingerprinting (Pinholt et al.[13]) **d**, Method E from the 2021 AnDi Challenge. **e**, DeepSPT. **f**, Table of classification metrics for the temporal segmentation of trajectory timepoints as post- and pre-uncoating and changepoint prediction. Median accuracy measures the median accuracy per trajectory per time point (N=100), that is all time points before a changepoint is defined as predicted 'before', while time points after are defined as predicted 'after'. These predictions are compared to the ground truth. F1-score measures the F1-score of all individual time point predictions as for median accuracy. Median frame error measures the median of absolute distances between predicted and true changepoints across all trajectories. Mean frame error measures the mean of absolute distances between predicted and true changepoints across all trajectories. Mean squared error measures the mean of the second norm distances between predicted and true changepoints across all trajectories. R2 measures the coefficient of determination.

https://doi.org/10.1038/s41592-025-02665-8



Extended Data Fig. 9 | Benchmark EEA1- versus NPC1- positive endosome predictions. Confusion matrices displaying classification performance on outputting NPC1 and EAA1 identity based solely on diffusion. **a**, DeepSPT. **b**, Original diffusional fingerprinting (Pinholt et al.[13]). **c**, Method E of the ANDI challenge. **d**, Utilizing the anomalous diffusion exponent (alpha) and diffusion coefficient. **e**, Only variations in diffusion coefficient. **f**, Only variations in alpha.



Extended Data Fig. 10 | Benchmark dorsal versus ventral AP2 predictions. Confusion matrices displaying classification performance of outputting whether AP2 is dorsal or ventral based solely on diffusion. **a**, DeepSPT. **b**, Original diffusional fingerprinting (Pinholt et al.[13]). **c**, Method E of the ANDI challenge. **d**, Utilizing the anomalous diffusion exponent (alpha) and diffusion coefficient. **e**, Only variations in diffusion coefficient. **f**, Only variations in alpha.

nature portfolio

Corresponding author(s): Nikos S. Hatzakis

Last updated by author(s): Jan 19, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	All Lattice Light Sheet Microscopy utilizes custom Labview-software developed by the Betzig Lab, HHMI Janelia. All Spinning disk confocal microscopy experiments utilized CellSens Dimension 2.3.
Data analysis	Minimal repository of code can be found at: https://erda.ku.dk/archives/752e4b0695c0dd16ec3c1a130f6ac70b/published-archive.html Repository of code and models can be found at: https://erda.ku.dk/archives/4c5adaaacc5c867f6450bcf89ec55a45/published-archive.html. Both under license CC BY-NC-ND 4.0 DEED. In addition, code is also freely available on GitHub https://github.com/JKaestelHansen/DeepSPT. Code uses Python (3.8.16), Adam optimizer (from Pytorch/torch 1.13.1), NumPy (1.23.5), Pandas (1.5.3), Optuna (3.2.0), Scipy (1.10.0), Trackpy (0.5.0), Matplotlib (3.6.3)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All data can be found in University of Copenhagen repository: https://erda.ku.dk/archives/804ea1ea88f340b79ada3e57141a6d6e/published-archive.html.(DOI in manuscript, Kæstel-Hansen 2024) All biological materials have no restrictions and are available upon reasonable request. We refer to the original publications for the additional experimental data used in this work; Wan et al., Bohr et al., and Hansen et al. as referenced in the manuscript.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample size calculation. For each of the experiments the sample size was chosen to provide training and test set sizes representative of variability and sample sizes seen in typical single particle tracking experiments. Simulations were created to provide sample sizes well above sample sizes seen in typical single particle tracking experiments.
Data exclusions	No data were excluded from the manuscript.
Replication	For all experiments we provide the code and data to retrain all used models and reproduce the findings. We perform K-fold cross-validation when possible or provide independent test sets. Microscopy experiments were successfully performed multiple independent times and for multiple replicates as indicated in the manuscript.
Randomization	For each experiment, the held-out test set, validation set, and training sets were randomly selected from the data. Multiple cross-validation folds were sampled as indicated in the main text.
Blinding	Blinding was not relevant to our study as we report analysis software as main finding. Yet, for each experiment, final model evaluation and reported results are based on held-out test sets or independent data which have no overlap with data used for training, thus model training is blind to the test set .

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Monoclonal antibody (mAb159) for rotavirus VP7 purification by affinity chromatography. These antibodies used were used in published work by Aoki et al, cited in main text and below, and graciously donated from Greenberg's lab.
Validation	Validation: Aoki ST, Trask SD, Coulson BS, Greenberg HB, Dormitzer PR, Harrison SC. Cross-linking of rotavirus outer capsid protein VP7 by antibodies or disulfides inhibits viral entry. J Virol. 2011 Oct;85(20):10509-17. doi: 10.1128/JVI.00234-11. Epub 2011 Aug 17. PMID: 21849465; PMCID: PMC3187514.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	HeLa cells are from ATCC (CCL-2), SVG-A cells are from ATTC (CRL-8621), sf9 cells are from Thermo Fisher Scientific (catalog # 11496015)
Authentication	All cell lines are authenticated by their commercial vendor. HeLa cells by STR profiling, karyotyping, and morphology. SVG-A by morphology and STR profiling. sf9 cells by isozyme and karyotype analysis.
Mycoplasma contamination	All cell lines are tested for mycoplasma by vendor. SVG-A cell stock are regularly tested for mycoplasma contamination (negative). HeLa cells were not tested for mycoplasma. The sf9 cells were not tested for mycoplasma but rather new vials directly from vendor (negative mycoplasma test by vendor) were thawed for each use.
Commonly misidentified lines (See ICLAC register)	To the best of our knowledge none of the used cell lines are part of the commonly misidentified lines.