How To Make Your Cell Tracker Say "I dunno!"

Richard D. Paul[1,2] Johannes Seiffarth[1,3] David Rügamer[2,4] Katharina Nöh[1] Hanno Scharr[1]
[1] Forschungszentrum Jülich [2] LMU Munich [3] RWTH Aachen University
[4] Munich Center for Machine Learning
{r.paul,j.seiffarth,h.scharr,k.noeh}@fz-juelich.de david.ruegamer@stat.uni-muenchen.de

Abstract

Cell tracking is a key computational task in live-cell microscopy, but fully automated analysis of high-throughput imaging requires reliable and, thus, uncertainty-aware data analysis tools, as the amount of data recorded within a single experiment exceeds what humans are able to overlook. We here propose and benchmark various methods to reason about and quantify uncertainty in linear assignment-based cell tracking algorithms. Our methods take inspiration from statistics and machine learning, leveraging two perspectives on the cell tracking problem explored throughout this work: Considering it as a Bayesian inference problem and as a classification problem. Our methods admit a framework-like character in that they equip any frame-to-frame tracking method with uncertainty quantification. We demonstrate this by applying it to various existing tracking algorithms including the recently presented Transformer-based trackers. We demonstrate empirically that our methods yield useful and well-calibrated tracking uncertainties.

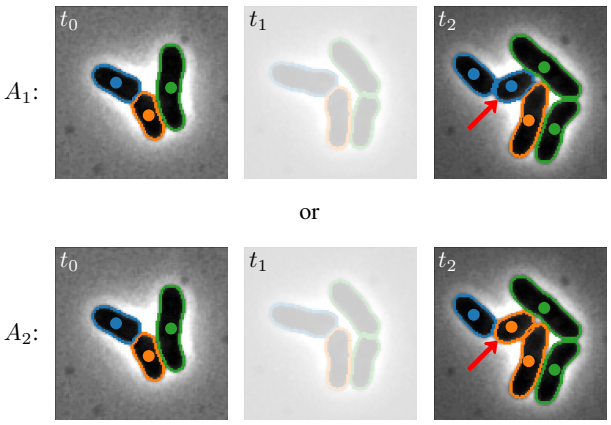

Figure 1. Example of ambiguity in two different assignment solutions A_1 and A_2 caused by the similar appearance of cells and missing frames. Assignments are color-coded, *i.e.* cells in frame t_2 with the same color as those in t_0 are considered daughters of the latter. Standard tracking methods yield point estimates, *i.e.* they will choose either solution, but not report any uncertainty caused by the ambiguity in choosing the correct mother for the cell marked by the red arrow.

1. Introduction

Uncertainty-aware cell tracking is a key requirement for fully automated data analysis of high throughput live-cell microscopy (LCM) data, where images often contain hundreds or thousands of almost indistinguishably looking, moving, growing and dividing cells and where temporal resolution of the time-lapses is limited by biological and technical considerations such as phototoxicity [38] or camera movement speed in multi-colony setups [34]. LSCM enables researchers to analyze cellular behavior beyond the population level, revealing dynamics, development and multi-species interactions [4, 5, 8]. The recent advances in computer vision, mainly driven by the success of deep learning and ever-improving computational resources, form a substantial pillar of modern LCM analyses, as the amount of data collected in a single experiment exceeds what humans are able to overlook [17, 34, 44].

Commonly, a data analysis pipeline for LCM data consists of a two-step procedure referred to as the *Tracking-by-Detection* (TbD) paradigm [10, 12, 22, 32, 33]. Cells first get detected and segmented using a deep learning model [7, 21, 31, 33, 36, 40] before being tracked across a sequence of images. Up until recently, tracking was most often solved using hand-crafted heuristics like nearest neighbor associations or linear assignments on overlap- or distance-based cost matrices. More recent methods replace such hand-crafted cost functions by ones learned from data [1, 12], having shown great generalization performance.

The strong reliance on computational tools demands for high reliability and trustworthiness, which is improved by uncertainty-aware analyses aiming to reliably estimate the confidence in their own predictions. While over-confidence in deep neural networks encountering distribution shifts is a commonly known issue in the deep learning community

[14, 19], uncertainty estimation as a remedy has – so far – attracted only little attention in cell tracking.

In this paper, we strive to make cell tracking more reliable by complementing it with uncertainty estimation. To this end, we explore two perspectives on the tracking problem, considering it as a Bayesian inference and as a classification problem. Our Bayesian perspective gives rise to the *cell tracking posterior*, motivating sampling and approximate sampling methods for uncertainty quantification, which however come at increased computational costs. The classification perspective provides less costly, nevertheless useful alternatives to quantify uncertainty. Moreover, the classification perspective also provides tools to evaluate and calibrate uncertainty estimates using known techniques such as *temperature scaling* [14]. Notably, the methods under consideration are applicable to the large family of linear assignment-based tracking methods, which are embedded naturally within our framework. We demonstrate this by testing our methods on various tracking algorithms, namely distance- [6, 10], overlap- [10], and activity-based cell tracking [32] as well as a novel, recently presented Transformer-based cell tracking algorithm [12]. Our results show that, indeed, vanilla cell tracking algorithms tend to be overconfident, especially at low temporal resolution, though their uncertainties still prove useful in that they positively correlate with the tracking performance. Moreover, we show that the methods developed in Section 2 improve the calibration of uncertainty estimation in cell tracking. Most notably, temperature scaling [14] turns out to be a cheap, but useful method to greatly improve calibration of tracking uncertainty if ground truth tracking data is available.

2. Probabilistic Cell Tracking

A major contributor to uncertainty in cell tracking are low frame rates at which microscopy images are often acquired (cf. Figure 1). The frame rate is often limited by technical or biological factors, like *e.g.* camera and stage movement speed [34] or phototoxicity [38]. For fast growing organisms, especially microbes, this becomes an issue when the dynamics within the population become too fast to allow for unambiguous tracking [28, 37]. We sketch the issue by means of an illustrative example in Figure 1, where one of the daughter cells (indicated by red arrows) is located such that a solely position-based association to the mother cell becomes ambiguous. In such a case, an off-the-shelf assignment solver will yield either of the two solutions, possibly choosing the wrong one, but remaining unaware of the alternative solution, and hence being unreliable. An uncertainty-aware cell tracking algorithm on the other hand should be able to express its uncertainty about the generated solution, *e.g.* in terms of a confidence score, indicating the ambiguity of the given problem. In the following, we introduce and discuss techniques for estimating the uncertainty in cell tracking.

2.1. Cell Tracking as Bayesian Inference Problem

Our consideration of the cell tracking problem starts with two consecutive frames of a microscopy image sequence with detections $\mathcal{X}_t, \mathcal{X}_{t'}$ at times $t < t'$ such as exemplarily shown in Figure 2a. At the core of most cell tracking algorithm lies the choice of a cost function $w : \mathcal{X}_t \times \mathcal{X}_{t'} \to \mathbb{R}_+$. Assuming x_i, x'_j denote the positions of the cells in frame t and t', respectively, *i.e.* $x_i, x'_j \in \mathbb{R}^d$ with $d = 2$ or $d = 3$ typically, the commonly used cost function

$$w_{\text{L2}}(x_i, x'_j) = \frac{\lambda}{2} \| x_i - x'_j \|_2^2 \tag{1}$$

is equivalent to assuming that cells behave in a Brownian fashion [6], *i.e.*

$$x'_j = x_i + \epsilon, \quad \text{for } \epsilon \sim \mathcal{N}(0, \lambda^{-1} I) \tag{2}$$

with λ being a hyperparameter representing the movement speed. From this we derive the likelihood $\mathbb{P}(x'_j \mid x_i)$ of observing cell x_i from the frame at time t at the position x'_j in the next frame at time t'.

Recently, cost functions have also been implemented by means of neural networks [1, 12], allowing to learn them from data rather than handcrafting them. Many of those techniques follow a simple paradigm, where a neural network is used as a feature extractor $f_\theta : \mathcal{X} \to \mathbb{R}^d$ with learnable parameters θ, that maps arbitrary cell detections and its features $\mathcal{X}$ to some continuous latent space [1, 12]. The neural network-based cost function is then computed as

$$w_{\text{NN}}(x_i, x'_j) = \frac{1}{2} \| f_\theta(x_i) - f_\theta(x'_j) \|_2^2. \tag{3}$$

In the case of multiple cells $x_1, \dots, x_m \in \mathcal{X}_t$ and $x'_1, \dots, x'_n \in \mathcal{X}_{t'}$, we aim to find an *assignment* $A \in \mathcal{P}(\mathcal{X}_t \times \mathcal{X}_{t'})$, such that every mother cell is represented at most twice and every daughter cell is represented at most once. We denote the subset of all assignments that adhere to those constraints as $\mathcal{A}^* \subset \mathcal{P}(\mathcal{X}_t \times \mathcal{X}_{t'})$, the *biologically feasible assignments* (*cf*. Figure 2b). The common approach to solve for a single assignment is to optimize the joint likelihood of observing cells $\mathcal{X}_{t'}$ given $\mathcal{X}_t$ and an assignment A, *i.e.*

$$A_{\text{opt}} := \arg\max_{A \in \mathcal{A}^*} \mathbb{P}(\mathcal{X}_{t'} \mid \mathcal{X}_t, A) \tag{4}$$

$$= \arg\max_{A \in \mathcal{A}^*} \sum_{(x_i, x'_j) \in A} -w(x_i, x'_j) - m w_a - n w_d \tag{5}$$

where m and n are the numbers of appearing and disappearing cells, $-w_a$ and $-w_d$ are the respective log probabilities of the events, and $w(x_i, x'_j)$ is some chosen cost

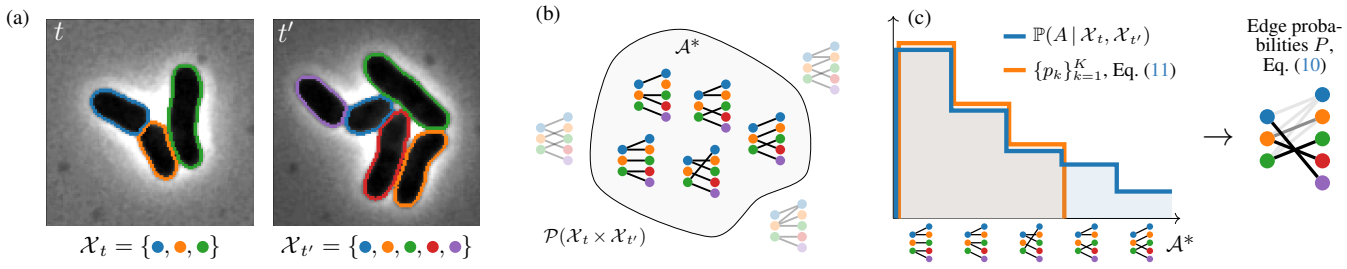

Figure 2. Schematic presentation of our Bayesian perspective and inference method for uncertainty-aware cell tracking. (a) depicts the detections $\mathcal{X}_t, \mathcal{X}_{t'}$ of two consecutive frames, which are the input to tracking methods from the TbD paradigm. (b) The set of biologically feasible assignments $\mathcal{A}^*$ forms a subset within the set of all possible many-to-many assignments. (c) Sorting assignment solutions by their posterior density we approximate the full posterior by a set of the K most plausible solutions. Edge probabilities are estimated as self-normalized importance-weighted average (*cf*. Equation (11)).

function like those from Equation (1) & (3) or Appendix B. A detailed derivation of Equation (4) & (5) is given in Appendix A. Conveniently, Equation (5) is the solution of a linear assignment problem (LAP).

Considering that the structure of the solution $A \in \mathcal{A}^*$, *i.e.* its biological feasibility, is known *a priori* without the need for any actual observations, the formulation from Equation (4) lends itself nicely to a Bayesian interpretation with A_{opt} being the *maximum a posteriori* (MAP) estimate of the posterior distribution of assignments

$$\mathbb{P}(A \mid \mathcal{X}_{t'}, \mathcal{X}_t) = \frac{\mathbb{P}(\mathcal{X}_{t'} \mid \mathcal{X}_t, A)\,\mathbb{P}(A)}{\sum_{A \in \mathcal{A}} \mathbb{P}(\mathcal{X}_{t'} \mid \mathcal{X}_t, A)\,\mathbb{P}(A)}, \tag{6}$$

where we choose the uniform distribution over valid assignments $\mathcal{A}^*$ as prior distribution, $\mathbb{P}(A) := \mathbb{1}_{\{A \in \mathcal{A}^*\}} / |\,\mathcal{A}^*\,|$. Given this *cell tracking posterior* distribution, we obtain the predictive distribution of the event that x_i is the mother of x'_j

$$P_{ij} := \mathbb{P}((x_i, x'_j) \in A \mid \mathcal{X}_t, \mathcal{X}_{t'}) \tag{7}$$

$$= \sum_{A \in \mathcal{A}^*} \mathbb{1}_{\{(x_i, x'_j) \in A\}}\, \mathbb{P}(A \mid \mathcal{X}_t, \mathcal{X}_{t'}) \tag{8}$$

as the weighted frequency of x_i being the mother of x'_j among all possible solutions, where we drop the explicit dependence on the t-th and t'-th frame for brevity in P_{ij}. Based on the close connection between the LAP and that of finding maximal weight matchings in bipartite graphs, we refer to a daughter-mother pair (x_i, x'_j) as an *edge* (*cf*. Figure 2).

2.1.1. Bayesian Inference on Assignments

The Bayesian formulation of cell tracking given in Equation (6) suggests to sample said posterior instead of just obtaining a single solution maximizing its density. This idea is also known as *joint probabilistic data association* [9]. Given a set of K assignments $A_1, \ldots, A_K \sim \mathbb{P}(A \mid \mathcal{X}_t, \mathcal{X}_{t'})$ sampled from the posterior distribution in Equation (6), we may estimate the predictive posterior from Equation (7) as

$$P_{ij} \approx \frac{1}{K} \sum_{k=1}^{K} \mathbb{1}_{\{(x_i, x'_j) \in A_k\}}. \tag{9}$$

Few works have considered the arising sampling problem thus far within the context of cell tracking. The close connection of solving Equation (5) and bipartite maximum weight matching observed in [32] suggests to use Markov chain Monte Carlo algorithms for sampling assignments. Alternatively, as $\mathcal{A}^*$ is a finite set, one can solve Equation (7) exactly by means of enumerating the solution set $\mathcal{A}^*$. The size of $\mathcal{A}^*$, however, is at least $\mathcal{O}(N!)$ with $N = \min\{|\,\mathcal{X}_t\,|, |\,\mathcal{X}_{t'}\,|\}$. Nevertheless, enumeration of the top-K best solutions to an LAP is possible [29] and offered by commercial LAP solvers like `Gurobi`. This allows to estimate Equation (7) by means of the self-normalized importance-weighted estimator [39]

$$P_{ij} \approx \sum_{k=1}^{K} \mathbb{1}_{\{(x_i, x'_j) \in A_k\}} \cdot p_k, \tag{10}$$

with weight

$$p_k := \frac{\mathbb{P}(\mathcal{X}_{t'} \mid \mathcal{X}_t, A_k)\,\mathbb{P}(A_k)}{\sum_{l=1}^{K} \mathbb{P}(\mathcal{X}_{t'} \mid \mathcal{X}_t, A_l)\,\mathbb{P}(A_l)}, \tag{11}$$

for $k = 1, \ldots, K$ and where we now assumed $A_1, \ldots, A_K$ to be the top-K best solutions. We sketch this method in Figure 2c.

2.1.2. Feature Perturbations

An alternative approach to create a set of K candidate solutions, which mitigates the problem of sampling or enumerating the set of valid assignments, is *feature perturbation*. Inspired by test-time augmentation [18, 27, 42, 43], we resample cellular features $\hat{x} \sim \mathcal{Q}(\hat{x} \mid x)$ from some noise

distribution $\mathcal{Q}$ conditional on the actual observations. A simple example of a noise distribution for positional features is a Gaussian perturbation centered on the cell detections, *i.e.* $\hat{x} \sim x + \epsilon$ with $\epsilon \sim \mathcal{N}(0, \gamma I)$. Given K sets of perturbed features $\hat{\mathcal{X}}_t^{(1)}, \dots, \hat{\mathcal{X}}_t^{(K)}$ and $\hat{\mathcal{X}}_{t'}^{(1)}, \dots, \hat{\mathcal{X}}_{t'}^{(K)}$ for each frame t, t', *i.e.* were the features of each cell have been resampled from the noise distribution, we solve the corresponding assignments using the perturbed features to obtain an ensemble of K assignment solutions $A_1, \dots, A_K$. Similar to the approach from Section 2.1.1, we use that ensemble of solutions to construct a Monte Carlo estimator

$$P_{ij} \approx \frac{1}{K} \sum_{k=1}^{K} \mathbb{1}_{\{(x_i, x'_j) \in \hat{A}_k\}}, \tag{12}$$

with $\hat{A}_k := \arg\max_{A \in \mathcal{A}^*} \mathbb{P}(\hat{\mathcal{X}}_{t'}^{(k)} \mid \hat{\mathcal{X}}_t^{(k)}, A)$ to obtain a confidence score for any particular mother-daughter relationship. We sketch this method in Section 2.1.3. Note that this approach requires the solution of K assignment problems. As a heuristic alternative, we also investigate performing the aggregation over the K perturbed samples before solving the assignment, *i.e.* we solve the LAP for the mean cost function

$$\hat{w}(x_i, x'_j) = \frac{1}{K} \sum_{(\hat{x}_i, \hat{x}'_j) \in \hat{\mathcal{X}}} w(\hat{x}_i, \hat{x}'_j) \tag{13}$$

only once. Since for this approach we cannot count the frequency of given edges in the perturbed solutions as in (12) before, we apply the classification-based approach introduced in Section 2.2 to get probabilistic confidence scores.

2.1.3. Bayesian Neural Perturbations

A key question in the feature perturbation approach is the specific choice of noise distribution $\mathcal{Q}$ and its parameters, *e.g.* the variance γ in our previous example of a Gaussian perturbation of the positions. A particularly interesting approach to this question can be derived from a probabilistic perspective at the feature extraction. More specifically, if we have access to a stochastic feature extractor, we may use its predictive distribution as our noise distribution $\mathcal{Q}$. For this, we consider the feature extractor to be a neural network $f_\theta : \mathcal{I} \to \mathbb{R}^d$ mapping raw images from the image space $\mathcal{I}$ to some latent space $\mathbb{R}^d$. The neural network in this case may either be a segmentation model [21, 31, 33, 40, 41] or the neural network from the earlier discussed cost function in Equation (3).

One conceptually attractive framework to obtain probabilistic predictions from the neural feature extractor is that of *Bayesian neural networks* [24]. Given some loss function $\ell_\mathcal{D}(\theta)$ over some dataset $\mathcal{D}$with which the neural network is trained, we consider the posterior distribution

$$\pi(\theta \mid \mathcal{D}) \propto \mathcal{L}(\mathcal{D} \mid \theta) \cdot p(\theta) \tag{14}$$

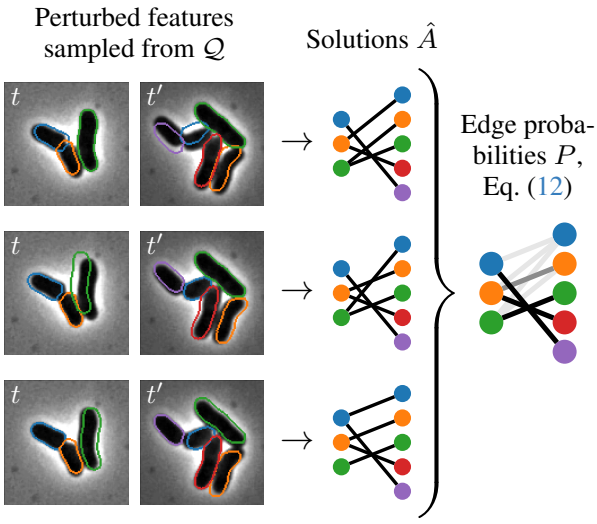

Figure 3. Schematic presentation of the feature perturbation approach. As an example, we perturb segmentation masks by shifting them randomly, which defines a particular noise distribution $\mathcal{Q}$. For each perturbed pair of frames we compute the corresponding assignment solutions $\hat{A}$. The posterior edge probability is approximated as the frequency at which an edge appears in the perturbed solutions.

of the neural network parameters θ. We do so by reinterpreting the loss as the negative log-likelihood $\ell_\mathcal{D}(\theta) \propto -\log \mathcal{L}(\mathcal{D} \mid \theta)$ and choosing a prior distribution $p(\theta)$ over neural network parameters. Since the posterior distribution $\pi(\theta \mid \mathcal{D})$ of neural networks is generally intractable in practice, one has to use an approximation instead. Typical approximation techniques are mean-field variational inference [3, 13], Monte Carlo Dropout [11] or Laplace approximations [20, 23]. Given the resulting approximate, but tractable surrogate posterior $q(\theta)$, the predictive distribution of the Bayesian neural network is approximated by generating an ensemble $\{f_{\theta_k}\}_{k=1}^K$ with $\theta_k \sim q(\theta)$ for $k = 1, \dots, K\}$ of differently parametrized neural networks sampled from $q(\theta)$.

For feature perturbation in cell tracking, we use the ensemble to generate a discrete distribution of features using the neural network ensemble. For the neural network-based cost function (*cf*. Equation (3)), the varying parametrization of the neural networks results in feature perturbation in the latent space onto which the neural network encodes the incoming features.

2.2. Cell Tracking as Classification Task

Given that sampling or enumeration of solutions to the LAP problem from Equation (5) is costly and can, thus, become infeasible, we search for alternative approaches to perform uncertainty quantification in tracking. In this

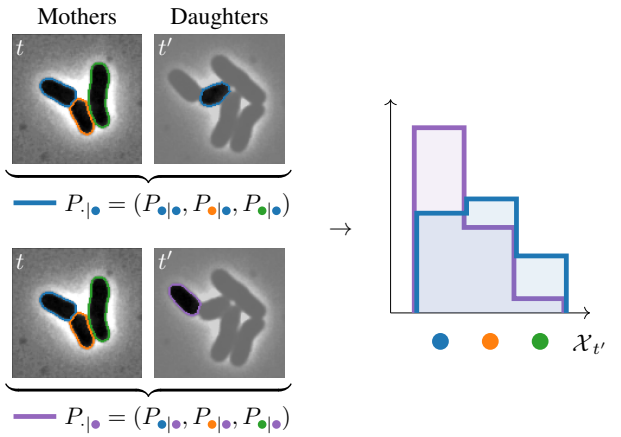

Figure 4. Two example DBMC class distributions. The blue cell from frame t' is located between the blue and orange cells from frame t, thus assuming distance-based tracking this results in high probability to be daughter of either of the two cells. In contrast, the purple cell is almost in the exact some spot as the blue cell is in frame t, resulting in high probability only for that mother.

section, we view the tracking problem as a classification problem, which recently has been leveraged to learn cost functions $w : \mathcal{X}_t \times \mathcal{X}_{t'} \to \mathbb{R}$ by means of neural networks [1, 12]. We consider the classification perspective to leverage known techniques for uncertainty quantification and calibration for the cell tracking problem.

2.2.1. Daughter-based Mother Classification

Cell tracking can be viewed as a multi-class classification problem by considering the problem of choosing the right mother for each individual daughter, which we refer to as *daughter-based mother classification* (DBMC). That is, for your favorite daughter $x'_j \in \mathcal{X}_{t'}$, the *classes* are just the mothers $\mathcal{X}_t$ from the preceding frame. The class – or rather mother – probabilities are obtained by performing daughter-wise softmax normalization of the predicted costs

$$P_{i|j} = \mathbb{P}((x_i, x'_j) \in A \,|\, x'_j, \mathcal{X}_t, \mathcal{X}_{t'}) \tag{15}$$

$$= \frac{\exp\{-w(x_i, x'_j)\}}{\sum_{x \in \mathcal{X}_t} \exp\{-w(x, x'_j)\}}. \tag{16}$$

From a generative perspective, this distribution could be used to sample daughter-mother associations. By construction, this approach guarantees that every daughter does indeed choose only at most one mother, but it does not guarantee that every mother has at most two daughters, as more than just two daughters might choose the same mother.

The DBMC perspective enables straightforward setup of training pipelines for neural network and other parameters involved in the design of the cost function w, as has been done in [12]. Moreover, as we elaborate below, we can interpret the softmax probabilities $P_{i|j}$ from Equation (15) as confidence scores, enabling their usage for uncertainty estimation.

In practice, using the *parental softmax* recently proposed by Gallusser and Weigert [12] is recommended to handle appearing cells, that *e.g.* enter the image across its border. We discuss the statistical assumptions introduced from parental softmax in Appendix C.

Finally, the DBMC approach is also compatible to the Bayesian approaches from Section 2.1 by simply column-normalizing the edge probabilities P_{ij} from Equation (10) and Equation (12), *i.e.*

$$P_{i|j} = \frac{P_{ij}}{\sum_k P_{kj}}. \tag{17}$$

2.2.2. Calibration

A typical measure to assess the reliability of uncertainty-aware predictions is *calibration*. A classifier producing a prediction y with attached confidence score $s \in [0, 1]$ is said to be *perfectly calibrated*, if the confidence score matches the empirical probability of the result to be correct [14], *i.e.* the probability to have the correct result y^* should be exactly s, *i.e.*

$$\mathbb{P}(y = y^* \mid s) = s. \tag{18}$$

The lack of calibration is quantified by considering the *expected calibration error* (ECE) [14]

$$\mathbb{E}\left[\left|\, \mathbb{P}(y = y^* \mid s) - s \right|\right]. \tag{19}$$

In practice, calibration has to be assessed by binning predictions based on the confidence score as presented in [14].

In order to consider calibration in cell tracking, we leverage the introduced DBMC perspective. That is, given a predicted daughter-mother pair $(x_i, x'_j) \in A$, the confidence on this decision is given as the mother probability $P_{i|j}$. Note that this is readily applicable not only to the DBMC approach based on the cost matrix as in Equation (15), but also to the Bayesian treatments based on the column-normalization from Equation (17). However, since the Bayesian treatments generate distributions of solutions, rather than a single solution, it may remain ambiguous what we consider a 'prediction'. A natural candidate in this setting is to consider the MAP estimate A_{opt} from Equation (4).

2.2.3. Temperature Scaling

A widely used post-hoc technique to improve calibration of classifiers is *temperature scaling* [14], which introduces a temperature parameter $\tau > 0$, *s.t.*

$$P_{i|j\tau} = \frac{P_{ij}^{\tau}}{\sum_k P_{kj}^{\tau}}. \tag{20}$$

The exponent τ is computed by optimizing the cross entropy loss between predicted and true class distribution [14].

It is worth noting that temperature scaling for the distance-based cost function w_{L2} from Equation (1) effectively scales the variance of the Brownian motion (*cf*. Appendix B). For other cost functions that are based on a p-norm distance between possibly latent cell features, as *e.g.* in Equation (3), it also scales the variance of the implicitly assumed generalized normal likelihood function $\mathbb{P}(x'_j \mid x_i)$ of said latent features (*cf*. Appendix B).

2.2.4. Quantifying Tracking Uncertainty

In the DBMC approach, we use the conditional probability $P_{i|j}$ as the probability for x_i being the mother of a given cell x'_j among all candidates $\mathcal{X}_t$. Given an assignment solution A, then the probability for any edge – a daughter-mother pair – to be correct is just the probability $P_{i|j}$ and thus a low probability is linked to high uncertainty.

An alternative approach to quantify the uncertainty within DBMC is to consider the *daughter-wise entropy*

$$\mathcal{H}_j = -\sum_i P_{i|j} \log P_{i|j} \tag{21}$$

of the conditional distribution $P_{i|j}$. High entropy means that the distribution is closer to uniformity and, thus, the decision is ambiguous [15]. However, since entropy is not a probability, it does not fit into the earlier discussed calibration framework. Nevertheless, entropy can be used as a decision criterion to distinguish possibly erroneous decisions from correct ones.

3. Related Work

This work is not the first to address cell tracking in terms of probabilistic methods. Already the seminal work by Crocker and Grier [6] phrased tracking as a maximum likelihood problem assuming Brownian particle motion. In this work, we challenge this viewpoint, arguing that the commonly used methods are rather to be seen as inferring a Bayesian MAP estimate, which not only nicely accommodates biological feasibility within the prior distribution, but also lays the foundation to argue about uncertainty in the tracking solution using the arising posterior distribution.

Work by Theorell et al. [37] and Kaiser et al. [16] is close to ours in that they also model tracking uncertainty by keeping track of multiple possible tracking hypotheses. Nevertheless, our work differs from theirs substantially in that we provide a very general framework for probabilistic tracking under a Markov assumption, *i.e.* independence between the tracking solutions from frame t to $t+1$ and those from $t+1$ to $t+2$. In contrast, Theorell et al. [37] and Kaiser et al. [16] specifically overcome the Markov assumption by informing the cost function for frame t to $t+1$ based on the set of tracking hypotheses for all frames up to time t.

Concurrently to our work, Betjes et al. [2] presented a method for *"cell tracking with accurate error prediction"* where link likelihoods predicted by a neural network are calibrated by enumerating alternative assignment solutions within a small neighborhood of adjacent cells and applying temperature scaling. Our work generalizes their approach to a wider family of tracking algorithms based on our plug-and-play-like framework. Moreover, we evaluate uncertainty estimation based on the DBMC approach, whereas Betjes et al. [2] do so by treating tracking as a binary edge classification problem. Finally, their local enumeration algorithm may be a promising alternative to the global enumeration method used within this work in Section 2.1.1.

Approaching cell tracking as a multi-class classification problem has been considered before [12]. We build upon that idea to quantify tracking uncertainty and calibration, which to the best of our knowledge has not yet been considered. In addition, we give some statistical considerations on the parental softmax introduced in [12].

Table 1. Overview of our methods for probabilistic cell tracking: Softmax'ing (SM), Feature Perturbation (FP), Feature Perturbation with Assignment (FP+A) and Assignment Sampling (AS). For all methods, we also test using temperature scaling, denoted by a trailing +TS in the abbreviation. To refer to temperature scaling in general, we use TS.

Method	P_{ij}	*cf*. Eq.
SM	$\exp\{-w(x_i, x'_j)\}$	(15)
FP	$\exp\{-\hat{w}(x_i, x'_j)\}$	(13)
FP+A	$K^{-1} \sum_{k=1}^{K} \mathbb{1}_{\{(x_i, x'_j) \in \hat{A}_k\}}$	(12)
AS	$\sum_{k=1}^{K} \mathbb{1}_{\{(x_i, x'_j) \in A_k\}} \cdot p_k$	(10)
SM+TS	$\exp\{-\tau w(x_i, x'_j)\}$	(15), (20)
FP+TS	$\exp\{-\tau \hat{w}(x_i, x'_j)\}$	(13), (20)
FP+A+TS	$\left(K^{-1} \sum_{k=1}^{K} \mathbb{1}_{\{(x_i, x'_j) \in \hat{A}_k\}}\right)^{\tau}$	(12), (20)
AS+TS	$\left(\sum_{k=1}^{K} \mathbb{1}_{\{(x_i, x'_j) \in A_k\}} \cdot p_k\right)^{\tau}$	(10), (20)

4. Experimental Evaluation & Discussion

Our methods for complementing cell tracking with uncertainty quantification work as a framework, which accommodates existing linear assignment-based TbD algorithms by simply specifying the corresponding weight function w. In the following experimental evaluation, we consider activity-based [32], distance-based [6, 10], overlap-based [10], and Transformer-based tracking [12], each defining a different cost function w (*cf*. Appendix B). We test the different methods introduced in Section 2 and summarized in Table 1.

For the feature perturbation-based methods, we use Gaussian perturbations of the cell centers with variance $\gamma = 0.1$ for the distance- and activity-based algorithms, a randomized inflation or deflation of the cell area by thick-

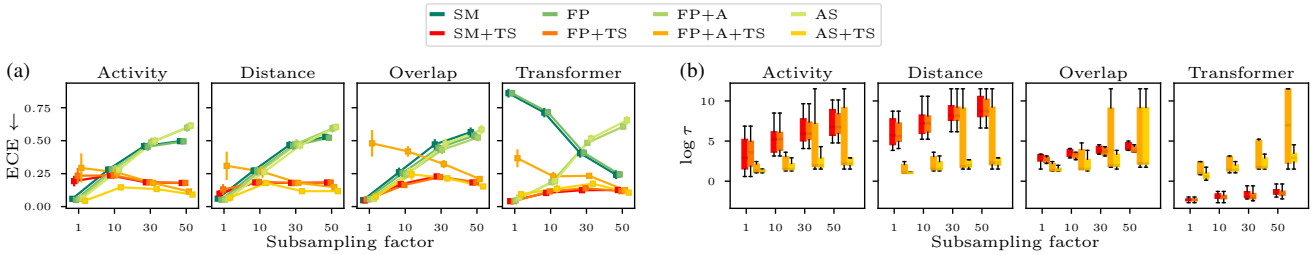

Figure 5. (a) Mean ECE and standard error thereof across datasets at decreasing temporal resolution. (b) Box plots showing the distributions of optimized temperatures at decreasing temporal resolution. For non-tempered methods, the temperature is $\log \tau = 0$, implicitly.

ening or thinning the cell borders for the overlap-based algorithm, and a Bayesian version of the Transformer-based algorithm using MC Dropout [11]. For the latter, we test dropout rates 0.5, 0.25, and 0.125. For the assignment sampling and feature perturbation strategy we take $K = 10$ samples.

The datasets under consideration are 2D datasets from the Cell Tracking Challenge [25, 26] and a publicly available dataset with five fully annotated time-lapse sequences of microbial colonies [34]. We used the ground truth segmentations as detections for the tracking, thus assuming perfect segmentation, isolating the effect of uncertainty intrinsic to the tracking problem. Since each dataset consisted of at least two sequences, we split all datasets such that one sequence was used for temperature scaling (where applicable) and the remaining sequences were used to evaluate our methods. Moreover, to simulate low temporal resolution, we also subsample the sequences by only considering every 10th, 30th, and 50th image from a time-lapse sequence.

4.1. Calibration

We focus our evaluation on the question of whether our approaches to uncertainty-aware tracking are capable of reporting 'useful' uncertainties. Calibration (*cf*. Section 2.2.2) tests for a one-to-one correspondence between estimated confidence and empirically achieved average accuracy. Thus, for a well-calibrated predictor, the estimated confidence is a predictor of the chance for the result to be correct. In Figure 5a, we depict the expected calibration error (ECE) averaged across the used datasets of our methods for various tracking algorithms and for varying temporal resolution. Most notably, we observe that temperature scaling (TS ■) almost always improves the ECE of any of our uncertainty-aware tracking methods, especially at the lowest temporal resolutions tested. The AS+TS ■ method yields particular good calibration across all tested algorithm and temporal resolutions.

For all non-tempered methods, except for SM ■ and FP ■ using the Transformer-based tracking, we do observe declines in calibration as the temporal resolution decreases. A possible explanation for the described decline in calibration, visible by an increase in ECE, may be our earlier consideration that the temperature represents the average squared displacement across time, as we discussed for the distance-based tracking (*cf*. Section 2.2.3 & Appendix B.1). Hence, as the temporal resolution decreases, the average squared displacement between images increases and thus the implicitly assumed temperature of $\tau = 1$ does not match to the increasing squared displacement.

We verify this explanation by comparing the optimized temperature to the mean squared displacements observed at given temporal resolution in Figure 9, which, indeed, suggests a positive correlation between the two. This observation further implies that improved calibration is achieved if the Brownian motion model underlying the distance-based tracking also models the motion's variance correctly, which is usually neglected. Moreover, for the other tracking algorithms, where the temperature parameter does not admit a physical interpretation, we depict the distribution of optimized temperatures across datasets against the decrease in temporal resolution in Figure 5b. As expected, the temperature raises as temporal resolution decreases, suggesting that less confident predictions improve calibration in the low resolution setting. This observation is in line with our intuitive expectations, as even humans struggle with correctly tracking cells when temporal resolution is low, due to the possibly large changes inbetween frames.

Interestingly, the increase in temperature is least notable for the SM ■ and FP ■ methods when using Transformer-based tracking. The same setups are the only non-tempered ones to achieve improved calibration as the temporal resolution declines (*cf*. Figure 5a). A possible explanation for this might be that the Transformer-based approach computes its cost matrices for frame-to-frame associations as the off-diagonal elements of a multi-frame block-wise cost matrix (*cf*. Appendix B) of encoded features. Those features have been encoded via an attention mechanism and, thus, already interacted with each other across multiple frames. This allows the Transformer to include multi-frame information into its feature encoding and finally the cost matrix, thus,

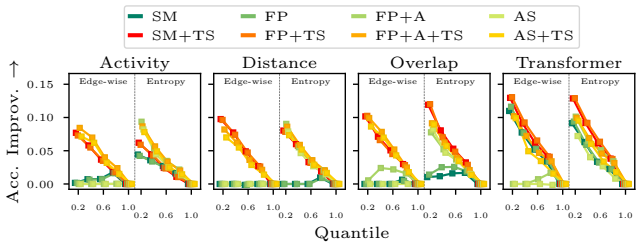

Figure 6. Mean accuracy improvement across tested datasets and varying temporal resolution (higher is better) using either the *edge-wise* probability or the daughter-wise *entropy* as a criterion to sparsify the prediction, *i.e.* omit predictions, which exceed a certain sparsification theshold computed as quantiles over the sparsification criterion.

possibly calibrating the costs already based on the multiple, available frames.

4.2. Sparsification

Another perspective on the 'usefulness' of uncertainty estimates is given by *sparsification plots*. Intuitively speaking, useful uncertainties are negatively correlated with the performance of a model, that is, high uncertainty corresponds to erroneous decisions, whereas low uncertainty corresponds to correct ones. Sparsification plots analyze this correlation by gradually removing the most uncertain decisions, while observing the performance of the remaining more certain ones, which should ideally increase. This type of analysis has particular relevance for practical applications of uncertainty estimation, as it allows to estimate prediction improvements and required effort within a human-in-the-loop approach, where the most uncertain predictions are corrected manually by an expert.

In Figure 6 we present such sparsification plots, where we show the average improvement in accuracy across different subsampling factors for our various uncertainty estimation methods and tracking algorithms. We refer to Figure 10 for a more fine-grained presentation of the sparsification results. Moreover, we measure uncertainty using either edge-wise probabilities or the daughter-wise entropy as discussed in Section 2.2.4.

We note that in comparison to ECE (*cf*. Figure 5a), TS is less critical to achieve accuracy improvement, especially when using the daughter-wise entropy for uncertainty estimation. In particular, the FP+A and AS methods exhibit only slight increases in accuracy improvement when using TS (*cf*. FP+A+TS & AS+TS) on top. However, for the edge-wise uncertainty estimation, TS is crucial for both methods and across all tested tracking algorithms in order to gain any accuracy improvements at all. For the SM and FP methods on the other hand, TS seems to be only crucial for the overlap-based tracking. In most other cases, TS does still yield accuracy improvements, but not as distinct. This seems to be mostly independent of the uncertainty measure in use.

In summary, entropy-based uncertainty estimation is less reliant on TS, and, thus, on annotated training data. A possible explanation for this might be that the edge-wise uncertainty estimation before TS does not incorporate any information on the class probabilities for the alternative mothers at choice except of a normalization factor, whereas computing the entropy across those class probabilities does. Nevertheless, TS does still mostly yield improved sparsification performance and should be, thus, used, if annotated training data is available.

5. Conclusion

In this work, we presented methods to quantify uncertainty in cell tracking caused by the ambiguity arising from low temporal resolution and high similarity of the objects which are to be tracked. To this end, we leveraged the two perspectives – as a Bayesian inference and as a multi-class classification problem – on cell tracking, which motivate the presented methods explored in this work. Based on the classification perspective, we presented approaches to quantify the reliability of our uncertainty-aware tracking methods in terms of calibration and sparsification.

Our experimental results indicate that tracking algorithms are not *per se* well-calibrated, but calibration is achievable using temperature scaling if annotated tracking data is available. Nevertheless, we notice that using our proposed daughter-wise entropy approach for quantifying uncertainty, even the vanilla tracking algorithms can produce 'useful' uncertainties, *i.e.* uncertainties that positively correlate with model performance. However, in particular our approximate Bayesian methods FP+A & AS show improved performance in terms of sparsification and are less reliant on temperature scaling and, thus, annotated tracking data, though this comes at increased computational costs (*cf*. Figure 7). We give a tabular overview of pros and cons of the different approaches in Table 2 in the appendix.

5.1. Future Work

Although our experimental results indicate that our methods do already provide useful uncertainty estimation, we hope to further inspire the development of uncertainty-aware tracking algorithms and incorporation of quantitative evaluation on the usefulness of the estimated uncertainties.

Moreover, tuning hyperparameters of our methods, such as the noise distribution in the feature perturbation approaches, which were fixed within this work, motivate the development of automated hyperparameter tuning pipelines for our uncertainty-aware tracking methods.

References

[1] Tal Ben-Haim and Tammy Riklin Raviv. Graph Neural Network for Cell Tracking in Microscopy Videos. In *Computer Vision – ECCV 2022*, pages 610–626, Cham, 2022. Springer Nature Switzerland. 1, 2, 5

[2] Max A Betjes, Sander Tans, and Jeroen Van Zon. Cell tracking with accurate error prediction, 2024. bioRxiv:2024.10.11.617799. 6

[3] Charles Blundell, Julien Cornebise, Koray Kavukcuoglu, and Daan Wierstra. Weight Uncertainty in Neural Network. In *Proceedings of the 32nd International Conference on Machine Learning*, pages 1613–1622. PMLR, 2015. ISSN: 1938-7228. 4

[4] Luisa Blöbaum, Luca Torello Pianale, Lisbeth Olsson, and Alexander Grünberger. Quantifying microbial robustness in dynamic environments using microfluidic single-cell cultivation. *Microbial Cell Factories*, 23(1):44, 2024. 1

[5] Alina Burmeister, Qiratt Akhtar, Lina Hollmann, Niklas Tenhaef, Fabienne Hilgers, Fabian Hogenkamp, Sascha Sokolowsky, Jan Marienhagen, Stephan Noack, Dietrich Kohlheyer, and Alexander Grünberger. (Optochemical) Control of Synthetic Microbial Coculture Interactions on a Microcolony Level. *ACS Synthetic Biology*, 10(6):1308–1319, 2021. Publisher: American Chemical Society. 1

[6] John C. Crocker and David G. Grier. Methods of Digital Video Microscopy for Colloidal Studies. *Journal of Colloid and Interface Science*, 179(1):298–310, 1996. 2, 6

[7] Kevin J. Cutler, Carsen Stringer, Teresa W. Lo, Luca Rappez, Nicholas Stroustrup, S. Brook Peterson, Paul A. Wiggins, and Joseph D. Mougous. Omnipose: a high-precision morphology-independent solution for bacterial cell segmentation. *Nature Methods*, 19(11):1438–1448, 2022. Publisher: Nature Publishing Group. 1

[8] Niklas Fante, Christian K. Desiderato, Christian U. Riedel, and Alexander Grünberger. Time-resolved cell-to-cell heterogeneity of Listeria innocua after nisin exposure. *Frontiers in Bioengineering and Biotechnology*, 12, 2024. Publisher: Frontiers. 1

[9] T. Fortmann, Y. Bar-Shalom, and M. Scheffe. Sonar tracking of multiple targets using joint probabilistic data association. *IEEE Journal of Oceanic Engineering*, 8(3):173–184, 1983. Conference Name: IEEE Journal of Oceanic Engineering. 3

[10] Yohsuke T Fukai and Kyogo Kawaguchi. LapTrack: linear assignment particle tracking with tunable metrics. *Bioinformatics*, 39(1):btac799, 2023. 1, 2, 6

[11] Yarin Gal and Zoubin Ghahramani. Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. In *Proceedings of The 33rd International Conference on Machine Learning*, pages 1050–1059. PMLR, 2016. ISSN: 1938-7228. 4, 7, 12

[12] Benjamin Gallusser and Martin Weigert. TRACKASTRA: Transformer-Based Cell Tracking for Live-Cell Microscopy. In *Computer Vision – ECCV 2024*, pages 467–484, Cham, 2025. Springer Nature Switzerland. 1, 2, 5, 6, 11, 12

[13] Alex Graves. Practical Variational Inference for Neural Networks. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2011. 4

[14] Chuan Guo, Geoff Pleiss, Yu Sun, and Kilian Q. Weinberger. On Calibration of Modern Neural Networks. In *Proceedings of the 34th International Conference on Machine Learning*, pages 1321–1330. PMLR, 2017. ISSN: 2640-3498. 2, 5

[15] Eyke Hüllermeier and Willem Waegeman. Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods. *Machine Learning*, 110(3):457–506, 2021. 6

[16] Timo Kaiser, Maximilian Schier, and Bodo Rosenhahn. Cell Tracking according to Biological Needs - Strong Mitosis-aware Multi-Hypothesis Tracker with Aleatoric Uncertainty. *IEEE Transactions on Medical Imaging*, pages 1–1, 2025. 6

[17] Keitaro Kasahara, Markus Leygeber, Johannes Seiffarth, Karina Ruzaeva, Thomas Drepper, Katharina Nöh, and Dietrich Kohlheyer. Enabling oxygen-controlled microfluidic cultures for spatiotemporal microbial single-cell analysis. *Frontiers in Microbiology*, 14, 2023. Publisher: Frontiers. 1

[18] Masanari Kimura. Understanding Test-Time Augmentation. In *Neural Information Processing: 28th International Conference, ICONIP 2021, Sanur, Bali, Indonesia, December 8–12, 2021, Proceedings, Part I*, pages 558–569, Berlin, Heidelberg, 2021. Springer-Verlag. 3

[19] Agustinus Kristiadi, Matthias Hein, and Philipp Hennig. Being Bayesian, Even Just a Bit, Fixes Overconfidence in ReLU Networks. In *Proceedings of the 37th International Conference on Machine Learning*, pages 5436–5446. PMLR, 2020. ISSN: 2640-3498. 2

[20] Pierre Simon Laplace. Mémoire sur la probabilité des causes par les événements. In *Mémoires de Mathématique et Physique*. Académie Royale des Sciences, 1774. 4

[21] Gihun Lee, SangMook Kim, Joonkee Kim, and Se-Young Yun. MEDIAR: Harmony of Data-Centric and Model-Centric for Multi-Modality Microscopy. In *Proceedings of The Cell Segmentation Challenge in Multi-modality High-Resolution Microscopy Images*, pages 1–16. PMLR, 2023. ISSN: 2640-3498. 1, 4

[22] Katharina Löffler, Tim Scherr, and Ralf Mikut. A graph-based cell tracking algorithm with few manually tunable parameters and automated segmentation error correction. *PLOS ONE*, 16(9):0249257, 2021. 1

[23] David J. C. MacKay. Bayesian Interpolation. In *Maximum Entropy and Bayesian Methods: Seattle, 1991*, pages 39–66. Springer Netherlands, Dordrecht, 1992. 4

[24] David J C Mackay. Probable networks and plausible predictions — a review of practical Bayesian methods for supervised neural networks. *Network: Computation in Neural Systems*, 6(3):469–505, 1995. Publisher: Taylor & Francis _eprint: https://doi.org/10.1088/0954-898X_6_3_011. 4, 12

[25] Martin Maška, Vladimír Ulman, David Svoboda, Pavel Matula, Petr Matula, Cristina Ederra, Ainhoa Urbiola, Tomás España, Subramanian Venkatesan, Deepak M.W. Balak, Pavel Karas, Tereza Bolcková, Markéta Štreitová, Craig Carthel, Stefano Coraluppi, Nathalie Harder, Karl Rohr, Klas E. G. Magnusson, Joakim Jaldén, Helen M. Blau, Oleh Dzyubachyk, Pavel Křížek, Guy M. Hagen, David Pastor-Escuredo, Daniel Jimenez-Carretero, Maria J. Ledesma-Carbayo, Arrate Muñoz-Barrutia, Erik Meijering, Michal

Kozubek, and Carlos Ortiz-de Solorzano. A benchmark for comparison of cell tracking algorithms. *Bioinformatics*, 30 (11):1609–1617, 2014. 7, 14

[26] Martin Maška, Vladimír Ulman, Pablo Delgado-Rodriguez, Estibaliz Gómez-de Mariscal, Tereza Nečasová, Fidel A. Guerrero Peña, Tsang Ing Ren, Elliot M. Meyerowitz, Tim Scherr, Katharina Löffler, Ralf Mikut, Tianqi Guo, Yin Wang, Jan P. Allebach, Rina Bao, Noor M. Al-Shakarji, Gani Rahmon, Imad Eddine Toubal, Kannappan Palaniappan, Filip Lux, Petr Matula, Ko Sugawara, Klas E. G. Magnusson, Layton Aho, Andrew R. Cohen, Assaf Arbelle, Tal Ben-Haim, Tammy Riklin Raviv, Fabian Isensee, Paul F. Jäger, Klaus H. Maier-Hein, Yanming Zhu, Cristina Ederra, Ainhoa Urbiola, Erik Meijering, Alexandre Cunha, Arrate Muñoz-Barrutia, Michal Kozubek, and Carlos Ortiz-de Solórzano. The Cell Tracking Challenge: 10 years of objective benchmarking. *Nature Methods*, 20(7):1010–1020, 2023. Number: 7 Publisher: Nature Publishing Group. 7, 14

[27] Nikita Moshkov, Botond Mathe, Attila Kertesz-Farkas, Reka Hollandi, and Peter Horvath. Test-time augmentation for deep learning-based cell segmentation on microscopy images. *Scientific Reports*, 10(1):5068, 2020. Number: 1 Publisher: Nature Publishing Group. 3

[28] Richard D. Paul, Johannes Seiffarth, Hanno Scharr, and Katharina Nöh. Robust Approximate Characterization of Single-Cell Heterogeneity in Microbial Growth. In *2024 IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 1–5, 2024. ISSN: 1945-8452. 2

[29] Seyed Hamid Rezatofighi, Anton Milan, Zhen Zhang, Qinfeng Shi, Anthony Dick, and Ian Reid. Joint Probabilistic Data Association Revisited. In *2015 IEEE International Conference on Computer Vision (ICCV)*, pages 3047–3055, Santiago, Chile, 2015. IEEE. 3

[30] Hippolyt Ritter, Aleksandar Botev, and David Barber. A Scalable Laplace Approximation for Neural Networks. In *International Conference on Learning Representations*, 2018. 12

[31] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pages 234–241, Cham, 2015. Springer International Publishing. 1, 4

[32] Karina Ruzaeva, Jan-Christopher Cohrs, Keitaro Kasahara, Dietrich Kohlheyer, Katharina Nöh, and Benjamin Berkels. Cell tracking for live-cell microscopy using an activity-prioritized assignment strategy. In *2022 IEEE 5th International Conference on Image Processing Applications and Systems (IPAS)*, pages 1–7, 2022. 1, 2, 3, 6, 11

[33] Tim Scherr, Katharina Löffler, Moritz Böhland, and Ralf Mikut. Cell segmentation and tracking using CNN-based distance predictions and a graph-based matching strategy. *PLOS ONE*, 15(12):0243219, 2020. 1, 4

[34] Johannes Seiffarth, Luisa Blöbaum, Richard D. Paul, Nils Friederich, Angelo Jovin Yamachui Sitcheu, Ralf Mikut, Hanno Scharr, Alexander Grünberger, and Katharina Nöh. Tracking One-in-a-Million: Large-Scale Benchmark for Microbial Single-Cell Tracking with Experiment-Aware Robustness Metrics. In *Computer Vision – ECCV 2024 Workshops*, pages 318–334, Cham, 2025. Springer Nature Switzerland. 1, 2, 7, 14

[35] Yuesong Shen, Nico Daheim, Bai Cong, Peter Nickl, Gian Maria Marconi, Bazan Clement Emile Marcel Raoul, Rio Yokota, Iryna Gurevych, Daniel Cremers, Mohammad Emtiyaz Khan, and Thomas Möllenhoff. Variational Learning is Effective for Large Deep Networks. In *Proceedings of the 41st International Conference on Machine Learning*, pages 44665–44686. PMLR, 2024. ISSN: 2640-3498. 12

[36] Carsen Stringer, Tim Wang, Michalis Michaelos, and Marius Pachitariu. Cellpose: a generalist algorithm for cellular segmentation. *Nature Methods*, 18(1):100–106, 2021. 1

[37] Axel Theorell, Johannes Seiffarth, Alexander Grünberger, and Katharina Nöh. When a single lineage is not enough: Uncertainty-Aware Tracking for spatio-temporal live-cell image analysis. *Bioinformatics*, 35(7):1221–1228, 2019. 2, 6

[38] Jean-Yves Tinevez, Joe Dragavon, Lamya Baba-Aissa, Pascal Roux, Emmanuelle Perret, Astrid Canivet, Vincent Galy, and Spencer Shorte. Chapter fifteen - A Quantitative Method for Measuring Phototoxicity of a Live Cell Imaging Microscope. In *Methods in Enzymology*, pages 291–309. Academic Press, 2012. 1, 2

[39] Surya T. Tokdar and Robert E. Kass. Importance sampling: a review. *WIREs Computational Statistics*, 2(1):54–60, 2010. 3

[40] Eric Upschulte, Stefan Harmeling, Katrin Amunts, and Timo Dickscheid. Contour proposal networks for biomedical instance segmentation. *Medical Image Analysis*, 77:102371, 2022. 1, 4

[41] Eric Upschulte, Stefan Harmeling, Katrin Amunts, and Timo Dickscheid. Uncertainty-Aware Contour Proposal Networks for Cell Segmentation in Multi-Modality High-Resolution Microscopy Images. In *Proceedings of The Cell Segmentation Challenge in Multi-modality High-Resolution Microscopy Images*, pages 1–12. PMLR, 2023. ISSN: 2640-3498. 4

[42] Guotai Wang, Wenqi Li, Michael Aertsen, Jan Deprest, Sébastien Ourselin, and Tom Vercauteren. Aleatoric uncertainty estimation with test-time augmentation for medical image segmentation with convolutional neural networks. *Neurocomputing*, 338:34–45, 2019. 3

[43] Guotai Wang, Wenqi Li, Sébastien Ourselin, and Tom Vercauteren. Automatic Brain Tumor Segmentation Using Convolutional Neural Networks with Test-Time Augmentation. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pages 61–72, Cham, 2019. Springer International Publishing. 3

[44] Lennart Witting, Johannes Seiffarth, Birgit Stute, Tim Schulze, Jan Matthis Hofer, Katharina Nöh, Marion Eisenhut, Andreas P. M. Weber, Eric von Lieres, and Dietrich Kohlheyer. A microfluidic system for the cultivation of cyanobacteria with precise light intensity and CO 2 control: enabling growth data acquisition at single-cell resolution. *Lab on a Chip*, 25(3):319–329, 2025. Publisher: Royal Society of Chemistry. 1