Detecting and Tracking Hard-to-Detect Bacteria in Dense Porous Backgrounds

Anonymous CVPR submission

Paper ID *****

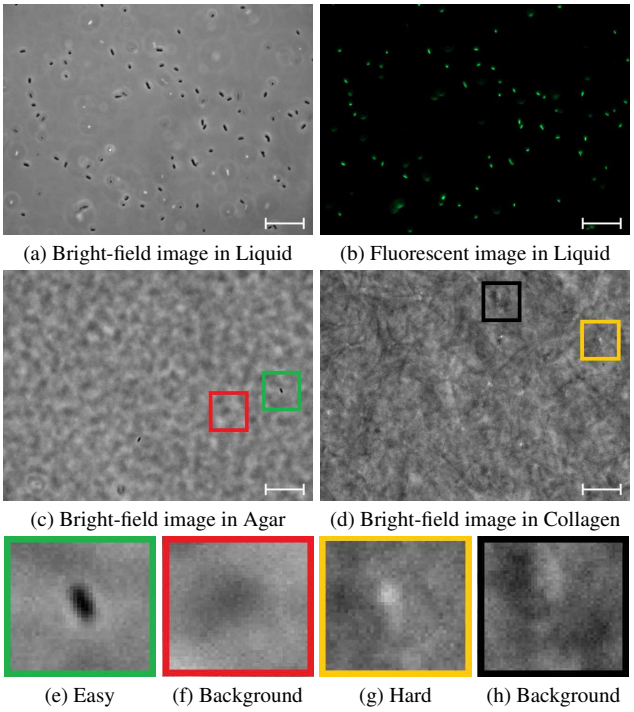
Abstract

Studying bacteria motility is crucial to understanding and controlling biomedical and ecological phenomena involving bacteria. Tracking bacteria in complex environments such as polysaccharides (agar) or protein (collagen) hydrogels is a challenging task due to the lack of visually distinguishable features between bacteria and surrounding environment, making state-of-the-art methods for tracking easily recognizable objects such as pedestrians and cars unsuitable for this application. We propose a novel pipeline for detecting and tracking bacteria in bright-field microscopy videos involving bacteria in complex backgrounds. Our pipeline uses motion-based features and combines multiple models for detecting bacteria of varying difficulty levels. We apply multiple filters to prune false positive detections, and then use the SORT tracking algorithm with interpolation in case of missing detections. Our results demonstrate that our pipeline can accurately track hard-to-detect bacteria, achieving a high precision and recall.

1. Introduction

An improved understanding of bacteria motility [19] is crucial to understanding and controlling biomedical and ecological phenomena involving bacteria. While automated tracking of bacteria has conventionally been done using fluorescent images (Fig. 1b), bacteria swimming speed (10-100 μm/s) necessitates high frame rate image acquisition, which is only attainable using gray-scale bright-field images (Fig. 1a), which have significantly lower contrast. Moreover, there has been a growing interest in studying bacterial motility in tissue-mimicking hydrogels or in porous media that resemble ecological settings. For example, agar is commonly used to culture bacteria [6], and collagen is the most abundant extracellular matrix protein in the body [1], making it relevant to use them as backgrounds while studying bacterial interactions with host cells. Hence, there is a need to solve the challenging problem of bacteria tracking in (gray-scale) bright-field images in more complex environments with textured backgrounds (Figs. 1c and 1d).

In this work, we explore the potential of using machine learning (ML) methods for multi-object tracking (MOT) for tracking bacterial cells in complex backgrounds. While existing work in MOT has primarily focused on detecting and tracking well-defined objects (e.g., pedestrians and cars), tracking bacteria is fundamentally more challenging for the following four reasons. ***(1) Lack of Distinguishable Features***: One of the major challenges in bacteria tracking is the visual ambiguity between the background (Figs. 1f and 1h) and the bacteria cells (Figs. 1e and 1g), making it difficult to distinguish between them. Furthermore, bacteria are often transparent or translucent, making it difficult to detect them using conventional imaging techniques. ***(2) 2D Imag-***



(a) Bright-field image in Liquid (b) Fluorescent image in Liquid (c) Bright-field image in Agar (d) Bright-field image in Collagen (e) Easy (f) Background (g) Hard (h) Background

Figure 1. While bacteria are easy to track in fluorescent (b) images, they are more difficult to detect in bright-field imagery available at high frame rates. Tracking bacteria in liquid media (a) is still easier than in realistic fibrous environments such as Agar (c) and Collagen (d). The visual ambiguity between the background (f and h) and bacteria (e and g) makes their detection challenging, which we focus in this work. (All scale bars are 20 μm.)

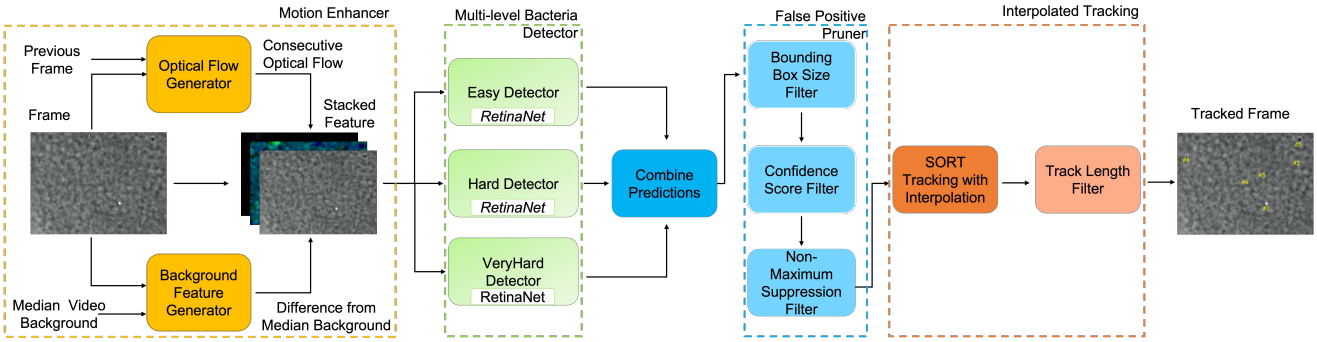


Figure 2. Overview of our proposed pipeline for tracking bacteria that consists of four modules: (1) Motion Enhancer, which adds motion features to the input frames, (2) Multi-level Bacteria Detector, which detects bacteria of three difficulty levels using RetinaNet models, (3) False Positive Pruning module, which filters detections to reduce false positives, and (4) Interpolated Tracker, which tracks bacteria over time using the SORT algorithm with interpolation.

ing of 3D Environments: While bacteria exhibit complex 3D motion, imaging is typically performed on a 2D focal plane, leading to restricted visibility. Consequently, bacteria may move in and out of the focal plane, resulting in halos or breaks in tracking. ***(3) Varying Difficulty Levels of Detecting Bacteria***: Because of a combination of (1) and (2), some bacteria are relatively easier to distinguish from the background (Fig. 1e) while others are hard to detect (Fig. 1g). To account for these differences, we use three different difficulty levels of detecting bacteria in our annotations: Easy, Hard, and Very Hard. Easy bacteria can be detected from a single frame, while the Hard bacteria require inspection of 2-3 consecutive frames to detect their motion. The Very Hard bacteria need the annotator to go over multiple frames, back and forth, to be able to detect them. ***(4) Erratic Movement Tracks of Bacteria***: Bacteria move in a variety of ways, including swimming, crawling, and tumbling, and their movements can be affected by various factors such as fluid flow, viscosity, and other environmental factors. The unpredictable nature of their movement may result in sudden changes in direction, speed, or orientation, making it difficult to track them consistently.

We propose the Multi-level Motion Enhanced Tracker (MMET) for tracking bacteria in bright-field microscopy videos by overcoming the challenge of tracking bacteria in cluttered and dynamic backgrounds (see Fig. 2). Specifically, we enhance the feature space using motion-based features, and then feed the stacked features to three object detection models to identify bacteria of varying difficulty levels using RetinaNet [15] as the backbone. We combine their predictions and apply multiple filters to prune out false positives. Finally, we employ the SORT algorithm [3] to track detected bacteria and use interpolation in case of missing detections due to bacteria movement in the Z dimension. Our results demonstrate that our pipeline can accurately predict and track hard-to-detect bacteria, achieving high precision and recall. Our proposed method not only represents a significant contribution to the field of biophysics and physical biology using computer vision, but also opens the potential of applying ML methods in applications like healthcare, environmental remediation, and chemical sensing where bacteria in porous media are relevant. [4]

2. Related Work

Tracking Bacteria: The most widely used software for manual bacteria tracking is ImageJ [22] and the available plugins such as Trackmate [24], MtrackJ [18], and CellProfiler3.0 [17]. Recently, an advanced version of Trackmate7 [8] was launched. It uses a tracking algorithm based on the LAP tracker (Linear Assignment Problem) [9], which assigns objects to tracks based on their similarity. CellProfiler4 [23] includes a module for tracking cells and other objects in microscopy images and videos and allows for the creation of custom analysis pipelines for specific applications. Other work [2, 7, 13, 26] track fluorescent image sequences in time-lapse images. But the field of fully automated tracking of bacteria at multi-cell level over a period of time to understand its motility patterns and behavior remains relatively unexplored.

Object Detection: There are two broad categories of methods for object detection: one-stage and two-stage. One-stage methods, such as YOLOv3 [20] and SSD [16], directly predict object bounding boxes and class labels from an input image. On the other hand, two-stage methods such as Faster R-CNN [21] and Mask R-CNN [10] first propose candidate object regions and then classify and refine them. We use RetinaNet [15], a one-stage method, as the base bacteria detector in our pipeline. Compared to other one-stage methods, RetinaNet achieves a better trade-off between speed and accuracy, making it suitable for detecting bacteria in cluttered and dynamic backgrounds.

Object Tracking: Several approaches have been proposed

to track detected objects including correlation filter-based methods [5], Kalman filter based methods [3] and deep learning-based methods [25, 27]. One of the most widely used tracking algorithms is the SORT algorithm [3]. It uses Kalman filter and provides state estimation and data association. SORT has been shown to perform well on various tracking tasks, including pedestrian tracking and vehicle tracking. However, the erratic motion of bacteria and missing detections on hard-to-detect bacteria may limit the performance of SORT, thus we propose a modification of SORT algorithm described in 3.

3. Multi-level Motion Enhanced Tracker

To address the challenges of tracking bacteria in complex environments, we propose a novel pipeline named Multi-level Motion Enhanced Tracker (MMET) that consists of four different modules: Motion Enhancer, Multi-level Bacteria Detector, False Positive Pruner and Interpolated Tracker. Before diving into each of the modules, we define the notations that we would be using in this paper. Let the input video with N frames be defined as $I^{1..T} = [I^1, I^2, ..., I^T]$, where $I^t \in \mathbb{R}^{C \times H \times W}$ denotes the t-th frame and C,H and W are the number of channels, height and width of the image respectively. Also note that we follow the approach of tracking-by-detection, i.e., tracking is done on top of predictions from the detection module.

Motion Enhancement: Object detectors designed for single-frame detection overlook the object's position in preceding or subsequent frames, which limits their effectiveness in tracking bacterias. However, to accurately detect hard and very hard bacteria, incorporating the concept of motion into the object detection model is crucial. Therefore, we propose two different feature engineering techniques in this module to capture motion features, which are then augmented with the image features to enhance detection accuracy.

Optical Flow Features: Optical flow [11] is a technique that estimates the motion of objects in a video sequence by analyzing the changes in pixel intensities between consecutive frames. We use the Lucas-Kanade method for optical flow computation which can be expressed mathematically as:

$$\partial I_x u + \partial I_y v + \partial I_t = 0 \tag{1}$$

where $u = \frac{dx}{dt}\big|_{t=t}$ and $v = \frac{dy}{dt}\big|_{t=t}$ represent the x and y components of the optical flow vector for the t-th frame, $\partial I_x = \frac{\partial I}{\partial x}$, $\partial I_y = \frac{\partial I}{\partial y}$, and $\partial I_t = \frac{\partial I}{\partial t}$ are the image gradients in the x, y, and time dimensions respectively. Solving this equation yields the optical flow vector $O = [u, v]$ for the t-th frame in the video.

Median Deviation Features: We define the median deviation as the pixel-wise difference between the image and the pixel-wise median of the video, which we use as another feature channel. Formally, we define median deviation $\Delta I = |I^t - \text{median}(I^{1..T})|$, where $\text{median}(I^{1..T})$ is the pixel-wise median for the video. The median deviation represents the difference between the pixel intensity at a point and the typical intensity at that point (which we assume is the intensity of the background media). In other words, it detects if a certain pixel is different from the background, i.e., if the pixel belongs to a bacteria we expect the median deviation to be positive. Finally, the input image features I^t are concatenated with the optical flow features O and the median deviation ΔI and used as inputs to the multi-level object detector.

Multi-level Bacteria Detection: For object detection, we use the RetinaNet [15] architecture, a state-of-the-art object detection method that has demonstrated excellent results in many computer vision applications. Training a single model on all of the different categories of bacteria is challenging as the visibility patterns and detection requirements vary significantly for each category. We propose a multi-level bacteria detection model where we train a different detector model for each category: Easy, Hard and Very Hard. Therefore, the detection models for each category of bacteria learn specific features and parameters tailored to their unique characteristics. This approach can help to enhance the detection accuracy of the object detection system for each category of bacteria and improve the overall performance of the system which we demonstrate empirically in Section 4.

False Positive Pruning: The multi-level bacteria detection module leads to a large number of false positives as well as duplicate predictions which need to be pruned without losing important predictions. We first combine all the detections from the 3 models and then filters them on two criteria: (1) pruning predictions that are be greater than the size of average bacterium (for our case, 22 pixels) and (2) removing predictions lower than a given confidence threshold. Next, we apply Non Maximum Suppression (NMS) [5], which is a technique used to eliminate redundant object detections by selecting the ones with the highest confidence score and discarding the others that overlap with them.

Interpolated Tracking: In the final step, we apply the Simple Online and Realtime Tracking (SORT) [3] algorithm to track the detected bacteria. The SORT algorithm uses a combination of Kalman filtering [12] and the Hungarian algorithm [14] to assign detected objects to existing tracks. However, to account for missing detections of Hard/Very-Hard bacteria, we modify the SORT algorithm. Specifically, we interpolate the missing detections by keeping the Kalman filter-based unmatched predictions for a given number of frames and drop the track post that threshold. The Kalman filter works by recursively updating estimates of the current state of a system based on the previous state and a set of measurements, while also taking into account the uncertainty of those measurements.

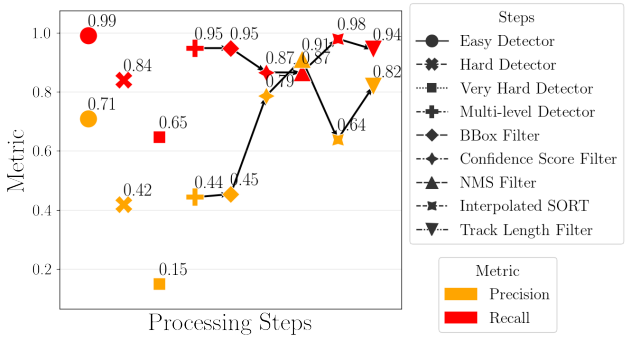


Figure 3. Precision and Recall values at each step of our pipeline for the Agar 0.2% medium.

Model	Task	Agar		Collagen	
		Precision	Recall	Precision	Recall
NM-MT	Detection	0.17	0.90	0.17	0.66
MMT	Detection	0.14	**0.98**	0.24	0.80
MMET (ours)	Detection	**0.45**	0.95	**0.29**	**0.82**
MMT	Tracking	0.05	**1.0**	0.14	**0.88**
MMT + FPP	Tracking	0.63	0.94	0.35	0.80
MMET (ours)	Tracking	**0.82**	0.94	**0.45**	0.84

Table 1. Comparing Test Precision and Recall for our model and baselines on agar and collagen datasets.

4. Experiments

Baselines: We compare our approach against two methods. (1) No Motion Monolithic Tracker (NM-MT) and (2) Motion-based Monolithic Tracker (MMT). NM-MT is trained only on images from bacteria videos without any motion features while MMT is trained on our motion enhanced features. Both these baselines use a single object detector (hence the name monolithic) as opposed to our multi-level approach. We apply Interpolated Tracking on both of them for tracking. We further enhance MMT by adding the False Positive Pruning module (FPP) before tracking.

In order to evaluate the effectiveness of our proposed pipeline for predicting and tracking bacteria, we conducted experiments on two different media: Agar 0.2% and Collagen. Our data consisted of 6 bacteria strains with a total of 4 videos at 60 FPS per strain in both agar and collagen media. We used 2 videos per strain for training, 1 for validation and 1 for testing. The limited size of the training sets is reflective of the paucity of labeled data encountered in many scientific applications, including biology. Each agar video had 100 frames while each collagen video had 150 frames. Each test set contains a range of highly motile to non-motile bacteria. For evaluation, we used precision and recall metrics.

Table 1 compares the results of our proposed method with the two baseline methods. We can see that using motion features improves the detection performance of MMT and MMET, especially in terms of recall. We can also observe the having a multi-level bacteria detector (MMET) leads to reduced False positives and hence higher precision. Comparing the tracking results for our model with the two baselines, we can see that the FPP module improves the precision of the baseline MMT by a significant amount. However, its precision is still lower than our proposed MMET model that uses a multi-level model instead of a monolithic model. Finally, while the baselines have a slightly higher recall on the detection task than our method, they suffer from extremely low precision, making them unfit for tracking. On the other hand, our approach can provide significantly bettter precision than baselines with little to no effect on the recall. Despite the preliminary nature of our evaluation on collagen, which involved only two difficulty levels of detecting bacteria instead of three, our proposed approach demonstrated better performance compared to the baselines. These initial results serve as a promising foundation for further investigations, including the use of 3 levels for collagen and more extensive hyper-parameter tuning.

Fig. 3 which shows the precision and recall values attained at each step of our pipeline for the agar test set. We can see that the precision significantly increase after each of the filters of the FPP module with a marginally small drop in recall. The Interpolated SORT module further boosts the recall, although at the cost of adding some false positives. After filtering out predicted tracks that are shorter than 55 frames in length, we are finally able to achieve a useful balance of precision and recall. We have further performed ablation studies to assess the impact of varying the bounding box size and applying NMS for removing redundant detections in the proposed pipeline for tracking bacteria.

5. Limitations and Future Work

We presented a pipeline for detecting and tracking bacteria in hydrogels such as agar and collagen. Our results demonstrate that the proposed pipeline accurately predicts and tracks bacteria, including hard-to-detect bacteria. However, our pipeline also has limitations. For example, it can generate duplicate detections from the combination of multiple models, which can lead to duplicate tracks, affecting the accuracy of the system. To address this limitation, we plan to explore different post-processing techniques that can group similar detections together and produce a single track for each group. Deep learning-based tracking methods can also be explored to improve tracking, such as [27]. Another direction for improvement could be calibrating the confidence scores of the different models before carrying out the filtering process to ensure that NMS does not get biased towards one model with a skewed distribution of confidence scores.

References

[1] Bo An, Yu-Shan Lin, and Barbara Brodsky. Collagen interactions: Drug design and delivery. *Advanced drug delivery reviews*, 97:69–84, 2016. 1

[2] Athanasios D Balomenos, Panagiotis Tsakanikas, Zafiro Aspridou, Anastasia P Tampakaki, Konstantinos P Koutsoumanis, and Elias S Manolakos. Image analysis driven single-cell analytics for systems microbiology. *BMC systems biology*, 11:1–21, 2017. 2

[3] Alex Bewley, ZongYuan Ge, Lionel Ott, Fabio Ramos, and Ben Upcroft. Simple online and realtime tracking. *CoRR*, abs/1602.00763, 2016. 2, 3

[4] Tapomoy Bhattacharjee and Sujit S. Datta. Confinement and activity regulate bacterial motion in porous media. *Soft Matter*, 15:9920–9930, 2019. 2

[5] David S. Bolme, J. Ross Beveridge, Bruce A. Draper, and Yui Man Lui. Visual object tracking using adaptive correlation filters. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 2544–2550, 2010. 3

[6] Maryline Bonnet, Jean Christophe Lagier, Didier Raoult, and Saber Khelaifia. Bacterial culture through selective and non-selective conditions: the evolution of culture media in clinical microbiology. *New microbes and new infections*, 34:100622, 2020. 1

[7] J.A. Cornwell, J. Li, S. Mahadevan, J.S. Draper, G.L. Joun, H. Zoellner, N.S. Asli, R.P. Harvey, and R.E. Nordon. Trackpad: Software for semi-automated single-cell tracking and lineage annotation. *SoftwareX*, 11:100440, 2020. 2

[8] Dmitry Ershov, Minh-Son Phan, Joanna W Pylvänäinen, Stéphane U Rigaud, Laure Le Blanc, Arthur Charles-Orszag, James RW Conway, Romain F Laine, Nathan H Roy, Daria Bonazzi, et al. Bringing trackmate into the era of machine-learning and deep-learning. *BioRxiv*, pages 2021–09, 2021. 2

[9] Yohsuke T Fukai and Kyogo Kawaguchi. LapTrack: linear assignment particle tracking with tunable metrics. *Bioinformatics*, 39(1), 12 2022. btac799. 2

[10] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross B. Girshick. Mask R-CNN. *CoRR*, abs/1703.06870, 2017. 2

[11] Berthold Horn and Brian Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 08 1981. 3

[12] Rudolf E. Kálmán. A new approach to linear filtering and prediction problems" transaction of the asme journal of basic. 1960. 3

[13] Johannes Klein, Stefan Leupold, Ilona Biegler, Rebekka Biedendieck, Richard Münch, and Dieter Jahn. Tlm-tracker: software for cell segmentation, tracking and lineage analysis in time-lapse microscopy movies. *Bioinformatics*, 28(17):2276–2277, 2012. 2

[14] Harold W. Kuhn. The hungarian method for the assignment problem. *Naval Research Logistics (NRL)*, 52, 1955. 3

[15] Tsung-Yi Lin, Priya Goyal, Ross B. Girshick, Kaiming He, and Piotr Dollár. Focal loss for dense object detection. *CoRR*, abs/1708.02002, 2017. 2, 3

[16] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott E. Reed, Cheng-Yang Fu, and Alexander C. Berg. SSD: single shot multibox detector. *CoRR*, abs/1512.02325, 2015. 2

[17] Claire McQuin, Allen Goodman, Vasiliy Chernyshev, Lee Kamentsky, Beth A Cimini, Kyle W Karhohs, Minh Doan, Liya Ding, Susanne M Rafelski, Derek Thirstrup, et al. Cellprofiler 3.0: Next-generation image processing for biology. *PLoS biology*, 16(7):e2005970, 2018. 2

[18] Erik Meijering, Oleh Dzyubachyk, and Ihor Smal. Chapter nine - methods for cell and particle tracking. In P. Michael conn, editor, *Imaging and Spectroscopic Analysis of Living Cells*, volume 504 of *Methods in Enzymology*, pages 183–200. Academic Press, 2012. 2

[19] Victoria Palma, María Soledad Gutiérrez, Orlando Vargas, Raghuveer Parthasarathy, and Paola Navarrete. Methods to evaluate bacterial motility and its role in bacterialndash;host interactions. *Microorganisms*, 10(3), 2022. 1

[20] Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement. *CoRR*, abs/1804.02767, 2018. 2

[21] Shaoqing Ren, Kaiming He, Ross B. Girshick, and Jian Sun. Faster R-CNN: towards real-time object detection with region proposal networks. *CoRR*, abs/1506.01497, 2015. 2

[22] Caroline A Schneider, Wayne S Rasband, and Kevin W Eliceiri. NIH image to ImageJ: 25 years of image analysis. *Nat. Methods*, 9(7):671–675, July 2012. 2

[23] David R Stirling, Madison J Swain-Bowden, Alice M Lucas, Anne E Carpenter, Beth A Cimini, and Allen Goodman. CellProfiler 4: improvements in speed, utility and usability. *BMC Bioinformatics*, 22(1):433, Sept. 2021. 2

[24] Jean-Yves Tinevez, Nick Perry, Johannes Schindelin, Genevieve M. Hoopes, Gregory D. Reynolds, Emmanuel Laplantine, Sebastian Y. Bednarek, Spencer L. Shorte, and Kevin W. Eliceiri. Trackmate: An open and extensible platform for single-particle tracking. *Methods*, 115:80–90, 2017. Image Processing for Biologists. 2

[25] Jack Valmadre, Luca Bertinetto, João F. Henriques, Andrea Vedaldi, and Philip H. S. Torr. End-to-end representation learning for correlation filter based tracking. *CoRR*, abs/1704.06036, 2017. 3

[26] Quanli Wang, Jarad Niemi, Chee-Meng Tan, Lingchong You, and Mike West. Image segmentation and dynamic lineage analysis in single-cell fluorescence microscopy. *Cytometry Part A: The Journal of the International Society for Advancement of Cytometry*, 77(1):101–110, 2010. 2

[27] Nicolai Wojke, Alex Bewley, and Dietrich Paulus. Simple online and realtime tracking with a deep association metric. *CoRR*, abs/1703.07402, 2017. 3, 4