

Developing Image Analysis Pipelines for Quantitative Microscopy of Nematode Embryos, Cytoskeleton-Motor Mechanics, Cell Shape Analysis and Adhesion

विद्या वाचस्पति की
उपाधि की अपेक्षाओं की आंशिक पूर्ति में प्रस्तुत शोध
प्रबंध

A thesis submitted in partial fulfillment of the requirements of the
degree of Doctor of Philosophy

द्वारा / By

छात्र का नाम / Dhruv Khatri / ध्रुव खत्री

पंजीकरण सं. / Registration No.: 20193627

शोध प्रबंध पर्यवेक्षक / Thesis Supervisor: Prof. Chaitanya A. Athale



भारतीय विज्ञान शिक्षा एवं अनुसंधान संस्थान पुणे
INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH PUNE

2024

Certificate

I certify that the thesis entitled "Developing Image Analysis Pipelines for Quantitative Microscopy of Nematode Embryos, Cytoskeleton-Motor Mechanics, Cell Shape Analysis and Adhesion" by Mr. Dhruv Khatri represents his original work which was carried out by him at IISER, Pune under my guidance and supervision. The work presented here or any part of it has not been included in any other University or institutions. I further certify that the above statements made by her in regard to her thesis are correct to the best of my knowledge.



Date : 20-May-2024
IISER Pune

Dr. Chaitanya A.
Athale
SUPERVISOR

डॉ. चैतन्य आठले / Dr. Chaitanya Athale
प्रध्यापक / Professor
भारतीय विज्ञान शिक्षा एवं अनुसंधान संस्थान
Indian Institute of Science Education & Research
पुणे / Pune-411 008, भारत / India

Declaration by Student

Name of Student: Dhruv Khatri

Reg. No.: 20193627

Thesis Supervisor(s): Prof. Chaitanya A. Athale

Department: Biology

Date of joining program: 02/01/2019

Date of Pre-Synopsis Seminar : 29/ 11 / 2023

Title of Thesis : Developing Image Analysis Pipelines for Quantitative Microscopy of Nematode Embryos, Cytoskeleton-Motor Mechanics, Cell Shape Analysis and Adhesion

I declare that this written submission represents my idea in my own words and where others' ideas have been included; I have adequately cited and referenced the original sources. I declare that I have acknowledged collaborative work and discussions wherever such work has been included. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

The work reported in this thesis is the original work done by me under the guidance of

Dr./Prof. Prof. Chaitanya A. Athale

Date: 16/05/2024

Signature of the student



Acknowledgements

My thesis presents the end of my journey at IISER, Pune. My research journey would not have been possible without the guidance and patience of my research supervisor, the support and camaraderie of my friends, and the everlasting love from my family.

Perhaps the most important goal of anyone's life is to find their passions - things that one enjoys learning and, even more so, doing. Image analysis has provided me with excitement and awe, and I have thoroughly enjoyed the things that I could learn and implement during my research. For this, I have no one other than my supervisor to thank. His patience and understanding during my initial years were pivotal in allowing me the time and space required to learn new skills. He was equally thrilled at solutions to problems as I was at testing new methods. Without the guidance of my supervisor, the published work that we managed to output would have remained as nascent ideas. In this regard, I would also like to express my gratitude towards the many collaborators who shared their expertise and interesting problems on which I could work. This has allowed me exposure to multiple fields that otherwise would not have been possible. I would like to especially thank Dr. Marie Delatter from ENS Lyon, whose passion for nematode systems led to a fruitful research publication. I am also grateful to Shivani Yadav, Tanvi Kale, Bhagyashree Deshmukh, Soumya Bhattacharyya, Dr. Sunish, and Dr. Wen Lu for sharing interesting yet challenging image datasets.

During my academic journey, I have had the privilege of receiving constructive criticisms and feedback from senior researchers which have improved my communication, presentation, and academic writing skills. The constant reviews from my supervisor, well-placed criticisms, not only showed me important deficiencies in my work but also instilled in me the qualities of attention to detail and thoroughness - the cornerstones of a solid academic journey. I am thankful to my supervisor for inculcating these qualities in my work. In this regard, I also thank my seniors, Dr. Neha, Dr. Anushree, Yash Jawale, and Dr. Kunalika, for being the role model for me to look up to. Their passion for their work and timely advice on how to tackle research were extremely useful.

Among the very skilled and esteemed professors of IISER Pune, I would like to thank Dr. Deepak Barua and Dr. Sutirth Dey for their initial review of my colloquium. I also want to thank Dr. Pranay, Dr. Madhusudhan, Dr. Krishnapal, Dr. Collins, Dr. Sudha, Dr. Suhita, and Dr. Sanjeev for their interactions and allowing me to work with them, however brief the period may have been. I have learned a lot about managing

students, taking lectures, data science, protein modeling, and malaria biology in my interactions with them. My very brief interaction with Dr. Vladimir Gelfand at the M2T2 meeting, when I had just joined the institute, gave me a firsthand look at how established and exceptional academics tackle work and foster collaboration.

Among those that I have had the privilege of knowing, I express my deepest gratitude towards Iktesh, Suraj, Aman, Digvijay, Aasim, Akashdeep, Narayan, Deepak S., Sushant and other residents of the 9th floor of Hostel-2. Among those that I look up to, Saiket Da and Akashdeep have been like an elder brother to me. The common place outside my room proved to be the meeting grounds of the most caring and supportive people I have come across in my life. Their constant chatter, light-hearted banter, and sometimes extreme altercations provided enough impetus to forget about my personal worries. The innumerable trips, lunch, dinner and outings with them, I thank you for each and every occasion. These are the diya's that have light up my life outside of work. In my room, I enjoyed online gaming in my free time, where I met fantastic competitive people. Moox3, SuperMamaReddy, Balan, Niko, Strang3, Egohurt69, Sukuna, and Sabreclaws are among the many "online" (*sometimes offline*) comrades in my FPS adventure. My best moments were spent with my most affable batchmates - Akshay, Keerti, Mamatha, Shilpa, Netra, Nikita, and Soumyajit. I thank all of you for your time, love, and care that you have showered in the past years. Among the dearest, I express my deepest gratitude, love, and affection towards *ladybird* and *bandari* who have been the solace of my mind and my heart.

I would like to extend my thanks to IISER and the various departments, mainly academic, IT, and Biology, that have made this journey possible. My sincere thanks to Tushar, Sayali, and Swapnil, who have always been there to help and solve the issues pertaining to academics. My thanks and apologies to Mrinalini and Mahesh for their help and for disturbing them on many departmental occasions. My due respects to the housekeeping staff at my residence and workplace who have worked tirelessly to maintain our institute.

Lastly, I would like to thank my family, the cornerstone of my life. Their love and care have supported my journey through its ups and downs. My dearest brothers, Ballu and Shubham, their love is always with me. The pillar of our family, my mother, I cannot thank you enough for your hard work and sweat that you perform for our family. My father has taught me my life's greatest lesson, patience and forgiveness. I thank each and every member of my extended family for their love and affection. I am thankful to the supreme being who looks after me in my times of need. My knowledge and skills are gained from articles and books. I hope that my thesis is able to provide

some inspiration and insights to others, just as I gained from the work of other authors. For their work, I am deeply grateful.

Synopsis

Introduction

Microscopy plays an indispensable role in modern biological research by enabling researchers to visualize complex biological environments. This visualization capacity allows for the study of intracellular processes, their regulation, and their organization in a spatio-temporal context. While direct visualization of cellular environments is feasible, the inherent complexity of these systems necessitates robust quantification methods [201]. A significant challenge in this field is the analysis of large datasets. Modern microscope technologies, such as high-throughput imaging platforms and super-resolution 3D platforms like FIB-SEM [290], generate terabytes of data that are impractical for manual analysis. Consequently, the evolution of modern microscopy drives concurrent advancements in methods for quantifying image datasets [67, 70, 258, 300]. Image quantification involves the use of algorithms to characterize entire images or specific components within them. Spatial description includes attributes like size, shape, and position relative to other regions within an image. This requires object segmentation, a process where a segmented object is identified as a complete binary unit within a specific image region [105]. Notably, segmentation of cells and their intracellular structures has been done for various cell types [94, 156, 161, 198, 269, 279]. Temporal description defines changes in spatial features over time, such as tracking a moving object in a time series. Algorithmic tracking surpasses manual analysis in robustness, reproducibility, and the ability to track objects over long time series. Single particle tracking (SPT) is now integral to studies related to intracellular environments [114], membrane or protein dynamics, cell migration [268], and development [225]. These algorithms for segmentation and tracking are part of digital image processing (DIP).

In this thesis, I have developed image analysis pipelines for the analysis of cytoplasmic viscosity and divisional timescale in nematode embryos, quantification of microtubule geometry, bacterial morphology, and RBC detection in microfluidic chambers. The primary objective of each task is detailed in the following section where a breakdown of the chapters is also provided. The work collectively enables researchers to extract meaningful insights from complex microscopy datasets.

Estimating cytoplasmic viscosity of nematode embryos through particle tracking microrheology 1.¹

During the first mitotic division of nematode embryos, diverse spindle behaviors are observed, potentially influenced by cytoplasmic viscosity changes. We hypothesized that gradual viscosity alterations contribute to these variations. Using DIC imaging, we studied spindle dynamics in *C. elegans* embryos and quantified viscosity via microrheology and endogenous granule tracking. Our MATLAB tool, DICOT, integrates image filtering based detection methods, achieving an F1-score of 0.8 for granule segmentation [39]. DICOT combines image preprocessing, modified Gaussian filters based detection and automated threshold detection for segmentation, and distance-based single-particle tracking algorithm for generating trajectories. These improvements to our previous gradient based detection pipeline offers a more reliable and robust solution for the precise segmentation of cytoplasmic granules which is used for cytoplasm microrheology. The tool was used to estimate cytoplasmic viscosity in six evolutionary-related nematode species. However, our quantification of cellular viscosity proved insufficient to explain the differences in spindle mobility observed in the studied embryos. Consequently, we proceeded to quantify the nature of forces acting on the spindle.

Evolutionary divergence of spindle mechanics in nematode embryos is constrained by antagonistic pulling and viscous forces.²

Cellular functions like cell division are highly conserved across phyla, yet the evolutionary principles guiding their organization are less explored. We compared six nematode species with asymmetrically positioned mitotic spindles during embryonic division. While *C. elegans* spindles oscillate during displacement, others elongate without oscillations. Laser microsurgery revealed varied cortical pulling forces across species. Using a viscoelastic model and quantified cytoplasmic viscosity, we determined cytoplasmic

¹Primary data analysis and tool development: Dr. Anushree Chaphalkar and Yash Jawale. Parts of this chapter have been published as Chaphalkar, Anushree R., et al. Quantifying intracellular particle flows by DIC object tracking. Biophysical Journal 120.3 (2021): 393-401.

²This chapter has been published as Khatri et al. Evolutionary divergence of anaphase spindle mechanics in nematode embryos constrained by antagonistic pulling and viscous forces. Molecular Biology of the Cell 33.6 (2022): ar61.

drag, cortical pulling forces, and spindle elasticity. Viscosity variations correlated with oscillation speeds, while absence of oscillations in low viscosity species indicated smaller pulling forces. This suggests a trade off between viscosity and pulling forces normalized by embryo size, shaping spindle mobility. Additionally, species with higher viscosity exhibited longer cell division times. To investigate this further we developed an automated method to mark visible checkpoints during first mitotic division for diverse nematode species.

Automated cell cycle time quantification during first mitotic division in diverse nematode embryos

Our analysis of nematode cytoplasmic viscosity uncovered a correlation between cell cycle times and viscosity in nematode embryos. To explore this further, we developed CNNs for accurate identification of cell stages during the first mitotic division in embryos of 23 *Caenorhabditis* species. Time series recordings from a public database were annotated manually for four stages and used for robust network training [272]. Four CNN models, including a custom-designed EVOCELLNET, demonstrated high accuracy in stage categorization, with gradient visualization techniques providing insights into their decision-making processes. The classification pipeline facilitated automated identification of cellular events with reasonable temporal accuracy, showcasing the potential of deep learning in studying evolutionary diversification of cellular dynamics.

In vitro microtubule oscillations quantified using graph based segmentation approach.³

in vivo spindle oscillations in *C. elegans* are a result of complex interactions involving axonemal geometry, motor mechanics, microtubules (MTs), elastic linkers, and biochemical regulation. However, the essential components and constraints driving MT oscillations remain elusive. While *in vitro* methods are commonly used to study motor-driven MT dynamics, our novel beating assay reveals rapid changes in MT filament geometry, particularly high curvature. To address challenges in tracking self-intersecting MTs in such assays, we designed a segmentation pipeline. This approach,

³All data acquisition experiments were performed by Shivani Yadav. Parts of this work are published in Yadav et. al (2024)

employing graph-based methodologies and conventional image processing techniques, facilitates tracking of single filaments undergoing bending and buckling. By minimizing distances within directed graphs across consecutive frames, we reconstruct filament traces, crucial for understanding motor-microtubule interactions, especially in scenarios with highly curved contours. Our method's utility is demonstrated in a modified gliding assay, accurately quantified through comparison with simulated data. This advancement holds promise for multi-filament tracking at high MT densities, contributing to understanding motor-microtubule-induced buckling properties, particularly relevant in *in vivo* scenarios such as spindle oscillations in *C. elegans*.

Bacterial cell morphology analysis.⁴

The cell surface area (SA) increase with volume (V) is determined by growth and regulation of size and shape. Most studies of the rod-shaped model bacterium *Escherichia coli* have focused on the phenomenology or molecular mechanisms governing such scaling. Here, we developed image processing pipeline to examine the role of population statistics and cell division dynamics in such scaling. Segmentation pipelines are developed for fluorescence, DIC and phase contrast microscopy images. An improvement in separating adjacent cells based on intensity profile is proposed. All segmentation pipelines utilize common steps of identification of initial contour using intensity threshold followed by edge optimization of these contours using active contours. Automated quantification allowed the examination of the role of population statistics and cell division dynamics on surface area and volume scaling for *E. coli*.

RBC cytoadhesion detection in microfluidics chamber.⁵

Malaria presents a significant health challenge. An important aspect of the malaria life cycle is the adherence of infected red blood cells (RBCs) to vital organs. Antibodies that target these interactions hold tremendous potential as vaccines. However, traditional assays lack the capability to simulate complex of *in vivo* conditions. To bridge this gap, an automated system utilizing microfluidic chambers to investigate the dynamics of RBC adhesion was proposed by our collaborators [60]. To accurately track

⁴All data acquisition experiments were performed by Tanvi Kale. Parts of this chapter have been published in Kale, Tanvi, Dhruv Khatri, and Chaitanya A. Athale. "Allometry of *Escherichia coli* surface area with volume: effect of size variability, filamentation and division dynamics." Physical Biology (2023).

⁵All data acquisition experiments were performed by Bhagyashree Deshmukh.

the dynamics of RBC adherence, we developed an image analysis pipeline to quantify cytoadhesion. The pipeline involves two main steps: image filtering with a Laplacian of Gaussian (LoG) filter and post-processing classification to distinguish in-focus RBCs from background noise. The filtered RBCs are then tracked and checked if stationary, with the detection process guided by intensity peaks corresponding to RBC locations.

Contributions

C.A.A Chaitanya A. Athale

T.B Thibault Brugiere

M.D Marie Delattre

B.D Bhagyashree Deshmukh

D.K Dhruv Khatri

A.C Anushree Chaphalkar

S.Y Shivani Yadav

Y.J Yash Jawale

T.K Tanvi Kale

K.K Krishnapal Karmodiya

Chapter 1: Conceptualization: M.D & C.A.A. Data resources: M.D & T.B.

Chapter 2: Conceptualization: M.D & C.A.A. Data resource and experimental methodology (ATP depletion data): M.D & T.B. Software: A.C, Y.J & D.K. Formal Analysis and Visualization: D.K

Chapter 3: Data resource: here. Conceptualization, methodology, software, formal analysis, validation, data annotation and writing: D.K.

Chapter 4: Experimental investigation & Conceptualization: S.Y & C.A.A. Filament assays: S.Y. Data analysis methodology, software, validation, data curation, writing and visualization: D.K.

Appendix E: Experimental investigation & Conceptualization: T.K & C.A.A. All data time series: T.K. Data analysis methodology, software, validation, data curation, visualization and method writing: D.K.

Appendix F: Experimental investigation & Conceptualization: B.D & K.K. Microfluidic assays: B.D. Data analysis methodology, software, validation, data curation, visualization and method writing: D.K.

Abstract

Automated quantification of microscopy datasets plays an important role in understanding biological phenomena. This requires robust computational methods for accurate object detection, tracking, and segmentation. In this thesis I present the development of image processing pipelines to quantify nematode embryo viscosity, spindle dynamics and cell cycle progression in evolutionary related nematode species, dynamics of single microtubule filaments driven by molecular motors, bacterial shape allometry and RBC adhesion from microscopy datasets.

Spindle dynamics, observed during the first mitotic division, show large variations across nematode species. To investigate these differences, we first interrogated the rheological properties of the embryo cytoplasm using a DIC object tracking software (DICOT). By combining granule microrheology with force estimates acting on the mitotic spindle, quantified through spindle laser ablation, we reveal a trade-off between cytoplasmic viscosity and spindle forces on the appearance of spindle oscillations during the first anaphase stage. Additionally, our quantification revealed a strong correlation between cellular viscosity and cell division time. Motivated by this, we developed a Deep Learning based model to automatically reconstruct the progression of the first mitotic division in diverse *Caenorhabditis* embryos from label-free DIC time series. The classifier predictions rely on stage-dependent morphological patterns, and in the future, the tool can be used to study cell stage-dependent changes in intracellular rheological properties.

The mechanics of spindle oscillations are driven by complex interactions between MT and molecular motors. These interactions have been studied using *in vitro* reconstitution systems. To understand the mechanics of MT-driven spindle oscillations, our lab previously devised an *in vitro* system to reproduce single filament oscillations. However, quantification of filament geometry using existing tools proved to be limiting. To address these limitations, we developed a segmentation pipeline, named as *KnotResolver*, to accurately trace intersecting or branched filaments that enables the quantification of oscillation frequencies.

Morphology of cells shows high heterogeneity in allometric scaling of surface area and volume. To understand this scaling relationship, we developed a cell morphology analysis pipeline that can be applied to multiple imaging modalities. In preliminary work, we have also developed a pipeline to quantify RBC cytoadhesion in *in vitro*

microfluidic chambers. Overall, the computational image-analysis pipelines provide valuable insights into various cellular biological processes, facilitating further advancements in the field.

Abbreviations

DIP Digital Image Processing

CV Computer Vision

DoG Difference of Gaussian

LoG Laplacian of Gaussian

SoG Scaling of Gaussian

CNN Convolutional Neural Networks

SVM Support Vector Machine

NN Nearest Neighbor

NCC Nucleocentrosomal complex

MT Microtubules

DIC Differential Inteference Microscopy

DICOT DIC object tracking

ROI Region of interest

SPT Single particle tracking

KV Kelvin Voigt

FG Force generators

UMAP Uniform Manifold approximation and projection

SNR Signal to noise ratio

Contents

1	Introduction	2
1.1	Cellular biology and image quantification	3
1.2	Computational analysis of microscopy images	5
1.2.1	Image filtering	6
1.2.2	Image Classification	7
1.2.3	Object segmentation in images	9
1.2.4	Object tracking in time series data	12
1.3	Evolutionary diversity in anaphase spindle mechanics of nematode embryos	13
1.3.1	Mechanics of spindle oscillations in <i>C. elegans</i>	15
1.3.2	Cell size scales with spindle morphology in <i>C.elegans</i>	17
1.3.3	Species dependent variation in spindle motion of related nema- tode embryos during the first asymmetric cell division	19
1.4	Thesis overview	21
2	Estimating cytoplasmic viscosity of nematode embryos through par- ticle tracking microrheology	23
2.1	Introduction	24
2.2	Materials and Methods	27
2.2.1	RNAi experiments	27
2.2.2	Passive microrheology	27
2.2.3	Image acquisition of nematode embryos and strain maintenance	29
2.3	Results	30
2.3.1	Annotation of cytoplasmic granules as ground truth to quantify detection accuracy	30
2.3.2	Filter parameter optimization and comparison	33
2.3.3	Fast nearest neighbor based particle tracking	35

2.3.4	Cytoplasmic viscosity changes six-fold between closely related species	37
2.4	Conclusions	42
2.5	Discussion	44
3	Evolutionary divergence of spindle mechanics in nematode embryos is constrained by antagonistic pulling and viscous forces	47
3.1	Introduction	48
3.2	Material and Methods	49
3.2.1	Laser ablation	49
3.2.2	Oscillation speed	50
3.2.3	Data analysis and fitting	51
3.2.4	Statistical tests and correlation	53
3.3	Results	53
3.3.1	Dynamics of anaphase spindle pulling varies independently of oscillations across species	53
3.3.2	The net pulling forces and elasticity forces acting on centrosomes also vary between closely related nematode species	55
3.3.3	Towards the definition of a parameter space for spindle positioning and oscillations	59
3.4	Discussion	64
4	Automated cell cycle time quantification of first mitotic division in diverse nematode embryos	69
4.1	Introduction	70
4.2	Materials and Methods	72
4.2.1	Manual annotation of cell stages	72
4.2.2	EvoCellNet architecture	73
4.2.3	Training parameters	74
4.2.4	Evaluation metrics	74
4.2.5	UMAP representation of network activations	75
4.2.6	Transition boundary annotation using SVM	76
4.2.7	Code availability	76
4.3	Results	77
4.3.1	Annotating Images of nematode embryo development for CNN classification	77

4.3.2	Accurate classification of distinct stages of first mitotic division across neamtode genera	79
4.3.3	Uniform manifold approximation and projection (UMAP) of cell stage progression	83
4.3.4	Network predictions are based on robust morphological features	86
4.3.5	Automated cell stage checkpoint recognition across diverse <i>Caenorhabditis</i> embryos	88
4.4	Discussion	91
4.5	Conclusion	92
5	<i>In vitro</i> microtubule oscillations quantified using graph based segmentation approach	94
5.1	Introduction	95
5.2	KnotResolver overview	97
5.2.1	Filament segmentation	97
5.2.2	Graph representation of self-intersecting filaments	98
5.3	Materials and Methods	102
5.3.1	Generating simulated image of oscillatory filaments	102
5.3.2	Segmentation accuracy metrics	105
5.3.3	Analysis of filament contours	106
5.3.4	Manual correction of contours	106
5.3.5	Code Availability	107
5.4	Results and Discussion	108
5.4.1	Performance validation of microtubule segmentation on simulated data	108
5.4.2	Performance evaluation of branch resolution on simulated data .	109
5.4.3	Qualitative comparison with existing methods	113
5.4.4	Semi-automated tracing of filaments in a pinned gliding assay .	113
5.5	Conclusions	115
6	Conclusion and Outlook	118
A	Supporting material to Chapter 2	127
A.1	Supporting Figures	127
B	Supporting material to Chapter 3	132
B.1	Supporting Figures	132

C	Supporting material to Chapter 4	135
C.1	Supporting Tables	135
C.2	Supporting Figures	137
D	Supporting material to Chapter 5	143
D.1	Figures	143
D.1.1	Supporting videos	144
E	Bacterial cell morphology analysis	148
E.1	Abstract	148
E.2	Introduction	148
E.3	Methods	149
F	RBC cytoadhesion detection in microfluidics chamber	152
F.1	Abstract	152
F.2	Introduction	153
F.3	Methods	154
G	Microtubule tip detection and tracking in gliding assays using deep learning	158
G.1	Abstract	158
G.2	Introduction	159
G.3	Methods	161
G.3.1	A gliding assay database	161
G.3.2	Filament tip detection network training	162
	Publications	165
	Bibliography	167

List of Figures

1.1	Schematic of image convolution.	6
1.2	Comparison of image classification framework between DL and ML. . .	9
1.3	Image segmentation using Otsu’s threshold.	10
1.4	First mitotic division in <i>C. elegans</i>	15
1.5	Linear scaling of spindle with cell size.	18
1.6	Diversity of spindle dynamics across nematodes.	20
2.1	Precision of yolk-granule detection in DIC	32
2.2	Comparing filters for DIC object detection	34
2.3	Precision of yolk-granule detection in DIC	36
2.4	Measuring granule diffusion coefficient for multiple species.	39
2.5	Cytoplasmic viscosity estimates from yolk granule mobility.	40
2.6	Effect of ATP depletion on granule mobility.	42
3.1	Spindle elongation trajectories from the analyzed nematode species. . .	52
3.2	Centrosome trajectories post laser ablation characterized by recoil velocity and decay constant	56
3.3	Pulling forces estimated from viscoelastic model fit to recoil data	60
3.4	Effect on spindle mobility of divergence in cell mechanics.	63
4.1	Stereotypical patterns during the first embryonic division across diverse nematode species.	78
4.2	CNN architecture and training metrics.	80
4.3	Confusion matrix for the validation set	83
4.4	UMAP projection of the validation dataset and gradient visualization using Grad-CAM	85
4.5	Network performance drops in the absence of stage dependant cytoplasmic features.	87

4.6	Cell stage boundaries combined with radial density trace in VGGNet classified UMAP space.	89
5.1	Filament graph representation	99
5.2	Filament crossover resolution by graph representation and minimization of curve dissimilarities	100
5.3	Identifying the optimal contour by minimizing deviation from an un- branched contour	103
5.4	Interactive resolution of complex knots from experimental data	107
5.5	Quantifying segmentation performace of KnotResolver and alternative segmentation methods with changing ‘noise’	110
5.6	Identifying the limits of detection accuracy from simulated data and manual path-input	112
5.7	Qualitative comparison with of KnotResolver with threshold-fitting and active contour based approaches	114
5.8	Contour analysis in time of representative experimental filaments	116
A1	Detection accuracy estimated by F1-score to compare filters.	128
A2	Filters compared for sensitivity, precision and F1-score	129
A3	MSD from the embryo mid-plane.	130
A4	Comparison between drift diffusion model (DDM) and anomalous diffu- sion model (ADM).	131
B1	Effect of spindle cutting on centrosome recoil.	133
B2	Spindle oscillation speed.	134
C1	Training dataset with image augmentation	137
C2	Training and validation metrics for loss and accuracy	138
C3	Performance comparison of the trained models in terms of precision, recall and F1 score	139
C4	Confusion matrix for the test dataset	140
C5	Cell stage boundaries combined with radial density trace for EvoCellNet, ResNet and EfficientNEt in UMAP space.	141
	SVCIGrad-CAM comparison between four models	142
D1	Representative time-series of filament with tracked contours	143
D2	Simulated filament images with increasing SNR used to measure position detection error comparing segmentation methods	144

SVD	Representative time-series of filament knot resolution with tracked contour	145
SVD	Representative time-series of filament knot resolution with Frenet–Serret vectors overlaid.	146
SVD	Representative time-series of filament knot resolution with multiple in- tersecting filaments.	147
D1	<i>E.coli</i> cell analysis pipeline for DIC, phase and fluorescence images. . .	151
F1	RBC detection using LoG filter	155
F2	Classification of LoG detections to detect "in-focus" RBCs using CNN	157
G1	Simulating experimental time series using cytosim.	162
G2	Tip detection inference on simulated dataset.	164

List of Tables

2.1	The viscosity of nematode embryonic cytoplasm from six species was estimated based on tracking the diffusion	43
3.1	Calculated cell length , cell cycle duration and centrosome size for the studied species.	54
3.2	Fit parameters from recoil data fitting.	58
4.1	Configuration of trained models	79
4.2	Network performance for each stage for validation dataset	81
4.3	Absolute mean error in checkpoint prediction	90
C.1	Training metrics	135
C.2	Validation metrics	135
C.3	Network performace of each network on each stage for test dataset . . .	136
C.4	Average inference time of trained on a single image.	136
G.1	Gliding assay simulation parameters	161

Chapter 1

Introduction

1.1 Cellular biology and image quantification

Microscopy is our primary means of observing cellular life, providing us with spatial information on cellular morphology, sub-cellular structure, and associated dynamics. Rapid advances in microscopic imaging aim to improve spatial resolution, i.e., the ability to distinguish between two adjacent objects, while also allowing for longer duration of observation [91]. A plethora of imaging modalities exist, which cover a wide range of techniques that cater to specific research needs. For example, fluorescence cell imaging can be used for protein specific labelling and is a powerful tool for studying biological processes such as G-protein activation and regulation [125], protein-protein interactions [212], trafficking [263], and development [178]. Modern imaging modalities like multi-photon microscopy, with higher depth penetration [143, 250] and light-sheet imaging, with minimal phototoxicity [122, 278] offer improved spatial resolution and precision for visualizing live cell and tissue structures and their dynamics over longer durations. Fluorescence can be combined with super-resolution techniques such as STED (Stimulated Emission Depletion) [275] and SIM (Structured Illumination Microscopy) [118] enables higher spatial resolution beyond the diffraction limit of light i.e, Abbe’s limit [1], aiding to the understanding of organelle structure and function [192, 287]. Label-free imaging modalities such as phase contrast [29] and DIC microscopy [220], along with other scattering-based methods [89], allow visualization of unperturbed cellular dynamics and are widely utilized for analyzing cell morphology and embryo development [45, 133, 282]. On the other end of the spectrum is electron microscopy which, which requires fixed samples and provides an ultrastructural overview of cellular architecture. Recent advancements such as FIB-SEM (Focussed Ion Beam Scanning Electron Microscopy) offer high-throughput imaging of cellular architecture at nanoscale (~ 20 nm) resolution [165, 261].

Despite the richer context to cellular life offered by contemporary microscopy tech-

niques, drawing quantitative knowledge from such datasets often proves to be limiting [54]. This is primarily due to the complexity of cellular environments, wherein, it is not possible to manually track or detect a signal of interest surrounded by varying noise levels and other cellular components. Cellular events also occur at scales that may not be caught by visual inspection alone and will be missed without a thorough data analysis [56]. Another key challenge is the sheer size of microscopy datasets. It is feasible to manually analyze a few images, but extracting knowledge from large datasets is extremely difficult. For example, electron microscopy and live cell datasets easily exceed gigabytes of memory [50].

In response to these challenges, there has been a parallel development of computational methods that allow a more reproducible, automated, and complete analysis of microscopy data. The development of these techniques comprises the nascent field of bio-image informatics [217]. The key aim of this field is to adapt or develop novel image processing algorithms that can be utilized for microscopy datasets [195]. Other aspects of this field include improving logistics for image storage, pushing open access and developing easy visualization of complex datasets [259].

The development of these algorithms is rooted in the fields of digital image processing (DIP) [105] and computer vision (CV) [56]. Together, they facilitate quantitative extraction of information from imaging datasets which, today, plays a pivotal role in understanding biological phenomena [183]. This involves development of robust computational methods for accurate object classification [196], segmentation [286], and tracking [43]. Many tools and software have been developed that implement these algorithms in a user-friendly manner, catering to the analysis of microscopy datasets. Notable examples include CellProfiler [34], ilastik [15], and Fiji [236].

Image processing algorithms have wide-ranging applications, however, their applicability to biological datasets is often limited by the low signal-to-noise ratio [189]. These algorithms often require manual tuning of parameters before application, even

for similar datasets, which requires an understanding of the underlying algorithms as well as some programming fluency [99]. The recent explosion of Artificial Intelligence (AI) promises to overcome these limitations. AI has revolutionized the field of image processing and, consequently, aided in the development of bio-image informatics. It offers automated solutions to diverse tasks of recognition, detection, and tracking without the overhead of tuning or knowledge of the underlying algorithms. Many methods that offer automated segmentation of cells [251], organelles [98], and tracking [176] have emerged in recent years. Although the significance of DL has been in its ease of applicability, DL algorithms have their own limitations which is failing to generalize on data that falls outside of their training data distribution [160, 210]. There is also a requirement to educate the users on the expected failures of the proposed algorithm. Nonetheless, until a robust solution to every problem is devised, the field of bio-image informatics will continue to ingeniously develop new algorithms for specific study cases. In addition to a sound knowledge of these algorithms, quantitative extraction of information from microscopy images also requires a thorough understanding of the biological process in question and the imaging modality that generates the data.

In the following section, I will first discuss the various categories of image processing algorithms relevant to this thesis, followed by an introduction to the biological motivation underlying my research. Since most of the work is a result of collaboration, the discussion on the imaging modality used is limited.

1.2 Computational analysis of microscopy images

Traditional computational analysis, to separate tractable algorithms from black box methods of AI, of images includes key steps of image pre-processing, segmentation, feature extraction followed by the final analysis [209]. In comparison, AI methods are more end-to-end applicable and provide an output given an input image. Nonetheless,

AI methods can be combined with algorithms for pre- and post-processing to improve results [283]. Generally, image pre-processing includes algorithms for noise reduction and contrast enhancement, which improve the separation between the object of interest and the background [149]. Key to many preprocessing and detection techniques is the convolution operation, which is performed during image filtering.

1.2.1 Image filtering

Image filtering is a core operation in image processing where a filter kernel, a small symmetric matrix, is applied to every pixel in an image. This convolution process is performed by sliding a filter kernel over the image, computing the weighted sum of the pixel values in the neighborhood defined by the kernel, and replacing the central pixel with the computed value (Figure.1.1). This operation is repeated for every pixel in the image, resulting in a filtered output [105]. This process allows specific features in the image to be highlighted or suppressed based on the filter design. Additionally, the filtering effect could be noise removal [151], blurring and edge detection [188] for image enhancement [38, 151]. There are two main types of image filters: linear and non-linear [38, 214].

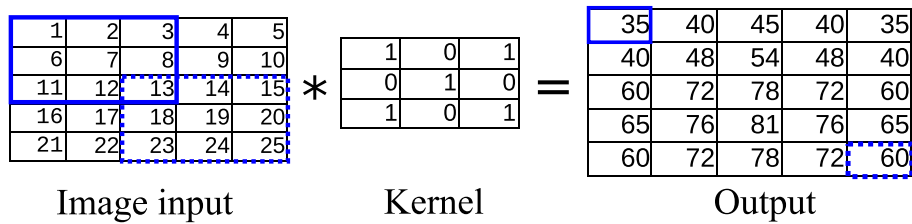


Figure 1.1: **Schematic of image convolution.** Schematic of image filtering through a convolution process with image input shown on the left, filter in the middle, and the output matrix on the right. The kernel filter size is 3×3

In linear filters the output pixel value is a linear combination of the input pixel values within the filter kernel's neighborhood. Examples of linear filters include mean

filters, Gaussian filters (smoothing), and sobel filters (edge detection) [105]. In non-linear filters the output pixel value is a non-linear function of the input pixel values, such as ranking and sorting, within the filter kernel's neighborhood. Examples of non-linear filters include median and bilateral filters [105].

Two specific types of linear filters, the Difference of Gaussians (DoG) [19] and Laplacian of Gaussian (LoG) kernels [110], are widely used for detecting edges in images [88, 265, 291]. The equation for the LoG filter is given by:

$$\text{LoG}(x, y) = -\frac{1}{\pi\sigma^4} \left(1 - \frac{x^2 + y^2}{2\sigma^2} \right) e^{-\frac{x^2 + y^2}{2\sigma^2}} \quad (1.1)$$

where: x & y : pixel coordinates and σ : standard deviation of the Gaussian. The equation for the Difference of Gaussians (DoG) filter is given by:

$$\text{DoG}(x, y) = \frac{1}{\sqrt{2\pi}} \left(\frac{1}{\sigma_1} e^{-\frac{x^2 + y^2}{2\sigma_1^2}} - \frac{1}{\sigma_2} e^{-\frac{x^2 + y^2}{2\sigma_2^2}} \right) \quad (1.2)$$

where: σ_1, σ_2 : standard deviations of the two Gaussians. These are used for detection or segmentation of structures with blob-like appearance, such as cells, nuclei, or other organelles in microscopy images [84, 280].

In this thesis, image filtering is extensively used for preprocessing of images in Chapter 5, Chapter 2 and Appendix E. A detailed comparison of filter performance for object detection tasks is performed for granule detection for nematode embryos in Chapter 2.

1.2.2 Image Classification

Image classification is a prediction task where the goal is to assign a label or a category to an input image based on its visual content [163]. Image classification is crucial for tasks like cell classification, disease diagnosis and organism identification [185, 240]. The effectiveness of the classification largely depends on the quality of the extracted

features and the choice of the appropriate classification algorithm. For AI methods, performance is dependent on the quality of pretrained embeddings [106], dataset size [87], and optimal training [164].

Machine learning (ML) based image classification typically involves key steps of image preprocessing, feature extraction, and model development. Image preprocessing involves preparing the input image for further processing, which may include operations like resizing [10], normalization [144], or applying filters to enhance or suppress certain features [164, 248]. In the feature extraction step, relevant features are extracted from the input image using various techniques known as feature engineering. These features can also be extracted from previously trained deep learning networks, which are known as pre-trained network embeddings [264].

In DL the time-consuming step of feature engineering is not required, and the model is therefore end-to-end trainable (Figure.1.2(A)). DL methods include convolutional neural networks (CNNs), which are particularly effective for image classification tasks [150]. ML methods use techniques like support vector machines (SVMs) [255], random forests [159], or logistic regression for the prediction model [162]. These algorithms require a feature extraction step, as performed during training (Figure.1.2(B)). The choice of classification algorithm depends on factors such as the complexity of the problem, the available training resources, and the trade-off between accuracy and computational efficiency [281]. After training the classification model, the algorithm is evaluated using appropriate metrics, such as accuracy, precision, recall, or F1-score, depending on the benchmark requirement [164].

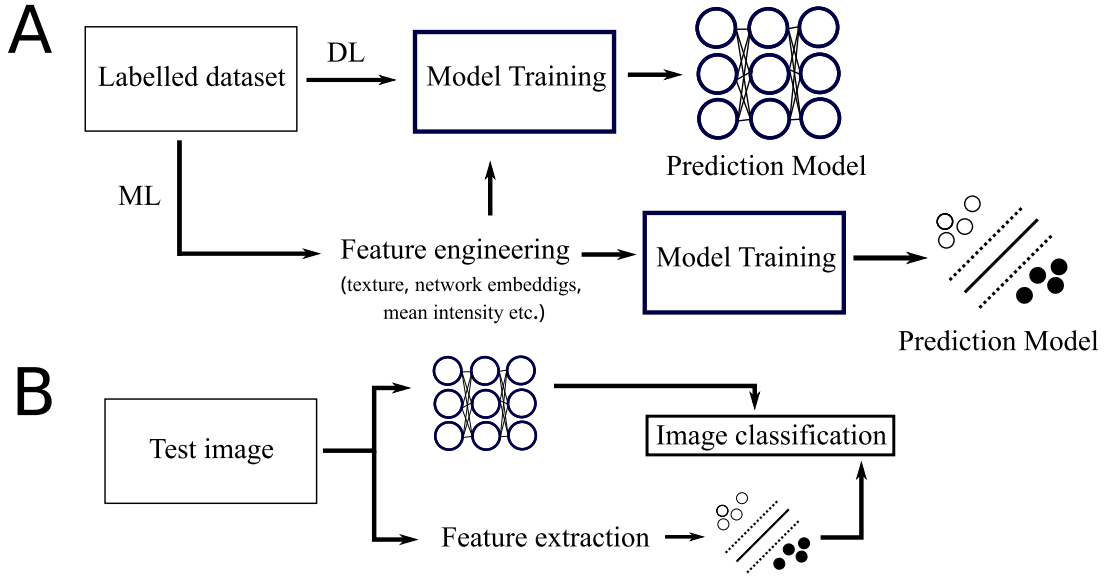


Figure 1.2: **Comparison of image classification framework between DL and ML.** (A) The difference between the training steps in DL and ML is shown. ML methods require a time-intensive process of feature engineering. (B) Prediction from an ML model requires an additional feature extraction step.

In this thesis, we have developed an automated cell stage classification pipeline based on convolutional neural networks (CNNs) in Chapter 4. A simple RBC detection pipeline is developed in Appendix F where we combine image filtering based detection with image classification based on CNNs.

1.2.3 Object segmentation in images

Although image classification can assign labels to an input image, interrogation on the object size, shape, and spatial characteristics requires complete and accurate separation of the object of interest from the rest of the image, i.e., the background. This is achieved by object segmentation, which involves partitioning an image or video frame into multiple segments or regions. Object segmentation methods can be broadly categorized into intensity-based methods [207], region growing [36, 134] and DL-based segmentation algorithms [231].

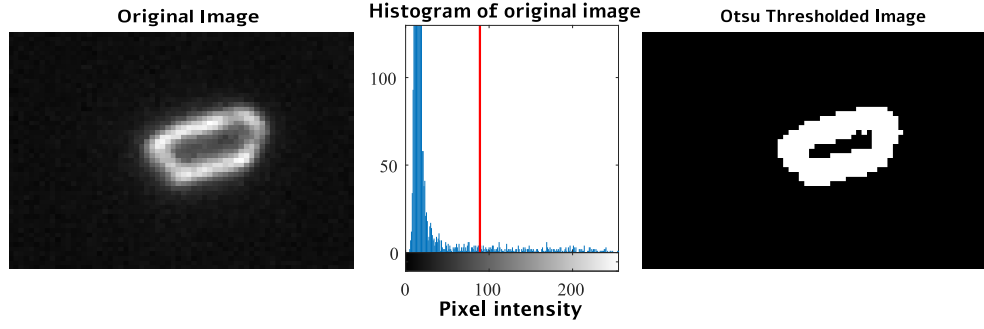


Figure 1.3: **Image segmentation using Otsu’s threshold.** From left to right: an input image of an *E. coli* cell labeled with FM 4-64 dye [86]. Intensity histogram of the original image with Otsu’s threshold shown as the red line. The binary output image.

Intensity-based methods, as the name suggests, rely on the intensity values (grayscale or color) of the image pixels to segment objects. Some popular intensity-based methods include intensity thresholding, wherein objects are separated from the background based on their intensity values relative to a defined threshold. A widely used method is Otsu’s threshold, which maximizes inter-class variance while selecting a suitable threshold for segmentation [207] (Figure 1.3). In this approach, all values above the threshold are assigned a value of 1, and below the threshold are assigned as 0, resulting in a binary output. Edge detection is another pixel intensity-based method for segmentation wherein sudden changes in pixel intensity are detected to help identify object boundaries [78, 303]. Since edges are also image features they can be generated by methods discussed in the image filtering section, e.g., Sobel filter [131]. Unsupervised clustering methods are based on the assumption that regions within an image exhibit similarities in terms of their pixel intensities, colors, textures or other derived features [48]. In this method pixels are grouped into clusters based on their intensity values or other derived features, with each cluster representing a distinct type of an object or region. A popular clustering-based approach is k-means clustering, which assigns pixels to different classes based on their proximity to the cluster centroids in the feature space [63].

Region-growing methods involve iterative grouping of pixels into regions or de-

lineating object boundaries based on some predefined criteria [221]. Some examples include region growing from seed points, where pixels are added to regions based on an intensity or texture similarity criteria. "Snakes" or active contours are another class of region growing methods wherein an initialized region is deformed to fit the object boundaries by minimizing an energy function that considers image features and contour smoothness [134]. The energy function is a combination of internal ($E_{internal}$) and external ($E_{external}$) energy components which can be expressed as:

$$E_{total} = \gamma E_{internal} + \delta E_{external} \quad (1.3)$$

where γ & δ are weight coefficients that balance out smoothness and contraction of the contour. Level set methods represent contours as level sets of a higher-dimensional function, which are more rubber-band-like instead of iterative "snakes" and can handle complex topology and shape deformations [206].

Deep learning-based methods have gained significant popularity for object segmentation due to their ability to learn complex features directly from annotated data [239]. Most popular deep learning architectures for object segmentation are based on Convolutional Neural Networks (CNNs) [244]. These are specialized neural networks designed to automatically learn hierarchical features from image data, making them highly effective for object segmentation tasks. U-Net is a specific CNN architecture that has been widely used for semantic segmentation tasks, where each pixel is classified into a specific object class [231].

The discussed object segmentation approaches are widely used for various cell [182], tissue [66], organism [179], and anatomical structure segmentation [154]. In cellular biology, identifying and separating individual cells or cell nuclei in microscopy images is essential for cell counting, tracking, and analysis. Similar to image classification, the selection of the appropriate segmentation method should consider the imaging modal-

ity, the characteristics of the objects of interest, the available computational resources, and the specific requirements of the application. Deep learning-based approaches, such as U-Net and its variants, have demonstrated remarkable performance in biomedical image segmentation tasks, potentially outperforming traditional methods. However, traditional methods are relatively easy to implement, and model-based designs provide accurate segmentation accuracy around cell boundaries, which is essential for morphometric analysis [65].

In this thesis, algorithms for image segmentation based on intensity threshold and region-growing method such as active contours [134] are extensively used in Chapter 2, Chapter 5 and Appendix E.

1.2.4 Object tracking in time series data

Object tracking is essential for understanding processes of cell migration, division, and cell-cell interactions [183]. These algorithms monitor and trace individual cells or cellular structures over time in microscopy images or videos. This allows researchers to analyze cell behavior, study dynamic cellular processes, and extract other valuable quantitative data [302]. At the molecular level, it reveals the dynamics of proteins, vesicles, and other particles inside cells [35, 40]. Tracking is also essential for studying embryogenesis by following every cell division and movement within a developing embryo [3, 113].

Object tracking algorithms typically involve two main steps: (i) *object segmentation*, discussed in the previous section, (ii) *object linking*, where the association between consecutive frames is generated to construct trajectories. The simplest approach for (ii) is to connect objects based on distance minimization in consecutive frames; the algorithm is known as the nearest neighbor (NN) tracker [186]. This could be extended to include similarities in size, shape, or dynamics. However, this approach is not suitable for crowded environments where it could lead to spurious linkages. More

advanced methods of linking include complex cost functions that optimize tracking to methods that account for future positions of the particle based on previous history [18, 285]. Since object tracking is a field of intense research, a more comprehensive survey of different methods can be found elsewhere [44, 176, 298]. In our work, we have only implemented the NN algorithm for tracking in Chapter 2, and detailed description is provided therein.

1.3 Evolutionary diversity in anaphase spindle mechanics of nematode embryos

One of the most remarkable aspects of life is its evolutionary and ecological diversity, which is also seen at the cellular scale [170]. The comparison of cellular architecture such as the cell wall, mitochondrion and chloroplasts, across species has provided key insights into the delineation of the key domains of life [200]. Among the many cellular organelles, comparative studies of the nucleus and mitochondrion across species have aided in the understanding of how cellular function evolves [62, 175]. However, our understanding of the processes that shape the evolution of cellular biology remains limited [170]. Quantifying cellular traits and systematically comparing them across species offers a powerful approach to understanding the interplay between genetics, selection, mutation, and cellular adaptation [170]. This requires including non-model species in a thorough comparison of known cellular processes, a field referred to as Evolutionary Cell Biology [170].

Nematode genera are key candidates for such studies [101, 112, 247]. The ease of culturing nematodes, their relatively short life cycle, and the prominent representation of *C. elegans* as a model system for studying embryonic development, aging, and behavior-related studies make them suitable for comparative analysis [25]. These worms are widely adapted to diverse environmental conditions and exhibit variations

in their morphology across species [59, 273]. Due to their transparent eggshell, it is feasible to compare the developmental processes of the early embryo across species through quantitative image analysis [20, 81]. The first mitotic division, in particular, has been the avenue of intense research, and while being a checkpoint for further embryo development, large variations within the first mitotic division are observed in related nematode species [24, 59, 273].

The division process is well described in the model system *C. elegans* (schematic is shown in Figure.1.4). The process starts with fertilization, with is the entry of the male gamete into the embryo. The newly formed pro-nuclei migrate toward each other and meet near the posterior end. The fusion of these two pro-nuclei forms the nuclear centrosome complex (NCC). The NCC complex migrates towards the center and aligns itself along the anterior-posterior axis. The nuclear envelope dissolves, and the mitotic spindle is formed. The spindle elongates asymmetrically, followed by spindle oscillations that end with the formation of the cytokinetic furrow, completing the first mitotic division. The formed daughter nuclei are asymmetric, and this asymmetry is critical for further fate determination, which produces the founder cells.

While the asymmetric division is under strong selection, quantitative analyses of first mitotic division have revealed striking differences in how the preceding stages progress [24, 273, 273]. Differences in how the spindle rotates, where it meets, its size, elongation dynamics, and aspects of cell division time are also variable across the nematode species [273]. In my thesis work, I have focused on addressing spindle mobility differences and quantification of cell division timings across diverse nematode species. We complement our analysis with a tool for quantification of *in vitro* MT oscillations.

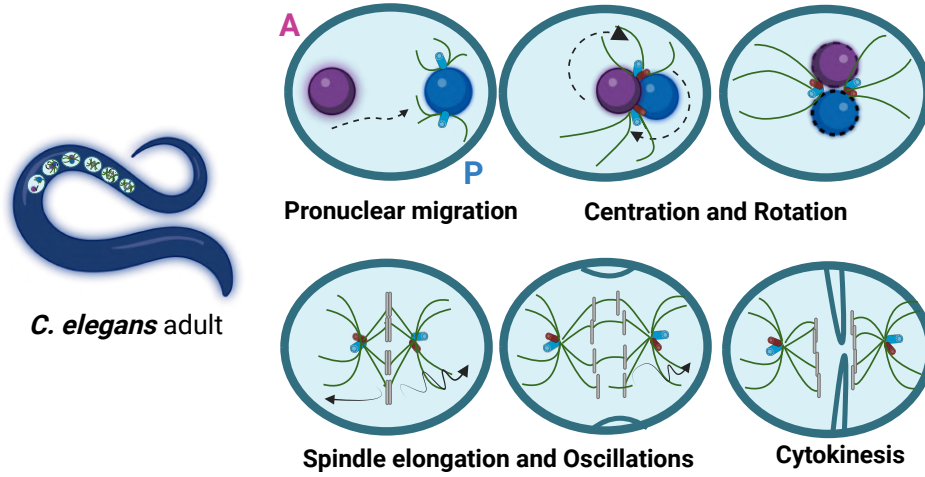


Figure 1.4: **First mitotic division in *C. elegans*.** Schematic of key events that occur during the mitotic division process. The two pro-nuclei male (blue) and female (magenta) are shown.

1.3.1 Mechanics of spindle oscillations in *C. elegans*

In *C. elegans* single-celled embryos, the spindle is initially centrally located. Asymmetric cortical pulling forces are responsible for its posterior displacement during anaphase. The entire spindle is first slightly shifted to the posterior in metaphase and early anaphase due to an imbalance of forces between the two sides of the spindle [107, 108, 152, 202]. Next, the posterior centrosome displaces faster and to a greater extent than the anterior centrosome during anaphase, generating asymmetric elongation [107]. Laser ablation of the central spindle at the onset of anaphase resulted in both centrosomes accelerating toward the cell poles, with the posterior centrosome velocity greater than the anterior, demonstrating that opposing and asymmetric pulling forces act on astral MTs to displace the centrosomes [107]. These movements result from the activity of a conserved dynein-containing protein complex anchored at the cortex [145]. During spindle displacement, the anterior and posterior centrosomes also oscillate back and forth along the transverse axis, in a manner that mimics anti-phase oscillations (offset by half a wavelength). These stereo-typical movements, with re-

producible build-up and die-down phases and specific frequency and amplitude, are referred to as spindle oscillations. Upon mild inactivation of the dynein-containing complex, spindle oscillations are abolished, yet the cell divides asymmetrically [215]. Hence, although not essential, spindle oscillations have been the subject of considerable research because they serve as a readout of the collective mechanics, as summarized in a theoretical model of spindle oscillations [109]. A more recent model of spindle oscillation mechanics suggests an additional role for cell size in modulating spindle oscillations [127]. Such collective mechanics resulting in spatial oscillations or instabilities is seen in multiple models of MT motor systems such as a multiaster system with cortical dynein and dynamic instability [138].

C. elegans spindle oscillations were predicted by a model to be driven by cortical motors, and astral MT attachments predict that oscillations emerge above a threshold of active forces [215]. At the same time, Kozłowski et al. (2007) predicted that a low cytoplasmic viscosity was also required for oscillations with speeds greater than $0.2 \mu\text{m/s}$, since higher viscosity resulted in a strong damping force opposing the motor forces [146]. Indeed *C. elegans* spindle oscillations have also been quantitatively reproduced by models of the mechanics of MT motors [108, 146]. Cortical force generators (dynein-containing complexes) pull from each side of the cortex (upper and lower), which should leave the centrosome in a stable central position [171]. In metaphase spindles of *C. elegans* embryos, these forces maintaining the central position have recently been estimated by fitting a viscoelastic model to perturbation-relaxation experiments on intact spindles [92], and these are likely to be comparable to forces acting at anaphase. The transverse displacement of the centrosomes is explained in the model by a positive feedback mechanism as a result of which a slight displacement of the centrosome toward the upper cortex, for instance, is amplified because pulling forces increase as the centrosome comes closer to the cortex, a scenario that results from the load per motor decreasing with decreasing distance to the centrosome [216]. The restoring

force acting to recenter the centrosome could be generated by astral MTs pushing on the cortex as they polymerize [146], or by the buckling of these MTs, extending laterally to the oscillation axis [216]. The tug of war between these pulling and restoring forces generates the oscillations. Pulling forces must also counter-balance the damping force generated by the viscous cytoplasm to launch the oscillations. Spindle motion is thus caused by the complex interplay between the driving mechanical forces and the material properties of the cell and of the spindle.

1.3.2 Cell size scales with spindle morphology in *C.elegans*

The mechanics of cytoskeleton and motors play a prominent role in cell division, transport and regulation. This would also suggest that these properties governed by physical constants must scale to compensate for the diversity across evolution, for example, with cell size and shape differences. Potentially they may also constrain the evolutionary diversification of a widely conserved process such as cell division. Size scaling or allometry of intracellular components with cell size has been described [37]. It lacks a simple explanation due to the diversity in mechanisms and missing principles regulating the size of intracellular organelles and structures [229]. The ubiquitous nature of the mitotic spindle has attracted similar approaches at size scaling through development and across evolution, since it involves the interplay of mechanical components of microtubules (MTs), motors, chromosomes, and regulators across species [80]. Recent work has proposed MT nucleation as a mechanism that determines spindle size scaling in development [230] that could provide a mechanism for the linear scaling of spindle and cell sizes during embryogenesis across species (Figure 1.5) [53, 80]. However, understanding the evolutionary scaling of dynamical properties of spindles their mobility, asymmetry, and motion is expected to be more complex, due to the time dependent nature of the process.

Large changes in cellular properties may also be possible as long as the phenotype

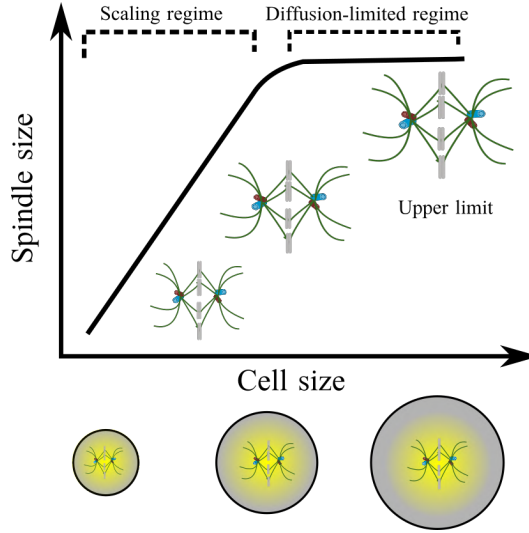


Figure 1.5: **Linear scaling of spindle with cell size.** The schematic illustrates the scaling of spindle size with cell size. In smaller cells, the scaling behavior is governed by the number of microtubules. When the cell size surpasses a certain threshold, microtubule dynamics become limiting, leading to the spindle reaching its maximum possible size. Yellow region indicates diffusion limited zone. Figure is adapted from [230].

under selection is not modified. These non-adaptive changes, or cryptic changes, help delineate the flexibility of biological mechanisms and their constraints [169]. Therefore to what extent the physical properties can evolve without affecting the underlying variation remains to be explored. In nematode species, similar to what has been described for *C. elegans*, an anterior/posterior (A/P) polarity axis is already established at the time of the first cell division [59]. One manifestation of this polarity is the asymmetric displacement of the mitotic spindle during the first anaphase, from a central to a posterior position. Consequently, the division generates two daughter cells of unequal size and of unequal fate. Understanding how the movements of spindle displacement differs between species could help in understanding how cryptic changes in the cellular parameters govern the physical constraints in the evolution of spindle motion.

1.3.3 Species dependent variation in spindle motion of related nematode embryos during the first asymmetric cell division

In this thesis we explore which cellular parameters can explain the absence of spindle transverse oscillations in the studied nematode species. Among those previously described by [272], we selected 4 species without spindle oscillations belonging to 4 distinct genera: *Pristionchus pacificus*, *Oscheius tipulae*, *Diploscapter* species 1 JU359 (D. sp. 1) and *Caenorhabditis monodelphis*. We also chose *C. remanei* and *C. elegans* as control species displaying anaphase spindle oscillations. The differences in spindle dynamics are illustrated with one representative embryo per species Fig.1.6. These species are also characterized by differences in cell size. Specifically, *P. pacificus* embryos are 10% longer and *D. sp. 1* are 20% shorter than *C. elegans* embryos. As previously shown on a larger set of species, the lack of spindle oscillation is not restricted to small or large embryos, thus cell size change alone is unlikely to be responsible for evolutionary changes in spindle oscillations [272].

Measuring biophysical properties of spindle dynamics

Our quantification of spindle motion across nematode species shows that spindle transverse oscillations are restricted to *Caenorhabditis* species [273]. In *C. monodelphis*, which is the most basal *Caenorhabditis* species, and in all species outside of this genus, the anaphase spindle is asymmetrically displaced without any transverse oscillations. We ask, what changes in cellular parameter can account for the absence of oscillations. A simple hypothesis is that the viscosity of the cytoplasm could give rise to these differences, based on predictions from simulations of *C. elegans* spindle-oscillatory mechanics [146]. However, these predictions remain to be tested experimentally, since altering cellular viscosity without affecting cell physiology is technically challenging.

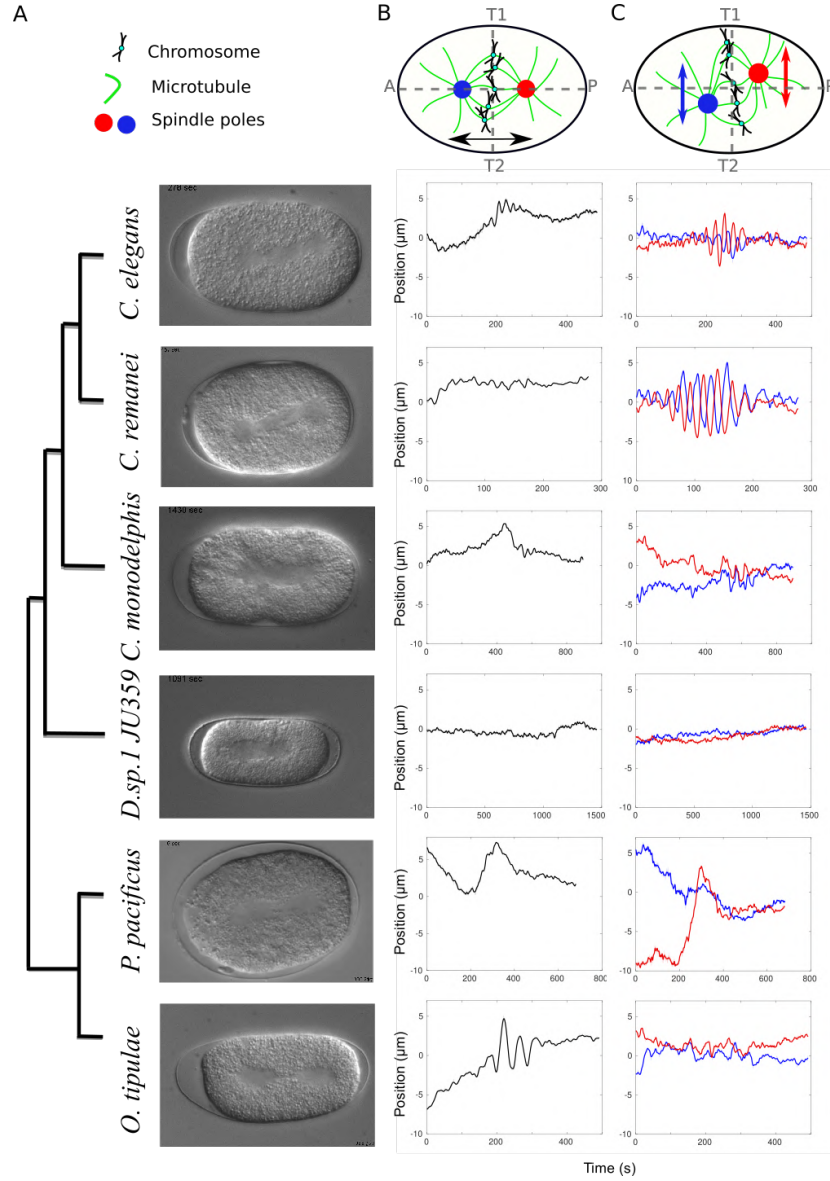


Figure 1.6: **Diversity of spindle dynamics across nematodes.** (A) The diversity in spindle movement in single-celled embryos of evolutionarily related nematodes is illustrated with one representative image per species taken from the study by Valfort et al. (2018). (Left) The phylogenetic relationship between the chosen nematode species and (Right) representative still images of single-celled embryos acquired in DIC are shown. Spindle and centrosome positions were tracked over time from DIC time series movies. (B) The change in position of the spindle midplane along the AP axis with time is plotted relative to the cell center (black lines). The position $0\mu\text{m}$ is the intersection between AP and T1-T2. Figure is adapted from Valfort et. al. 2018 [273]

Alternatively, if the net pulling forces are reduced, we also expect the loss of oscillations. Across species, spindle oscillation buildup may be hindered by an increase in cytoplasmic viscosity or by a reduction in cortical forces arising from gene expression changes, reduced cortical localization, or other parameter change such as cell size and MT mechanical properties.

In the next chapter, we address the question of which biophysical features explain the diversity of spindle motion observed between species. In doing so, we specifically asked which combination of parameters, specially viscosity and cortical forces, have been retained over the course of evolution as good solutions to sustain asymmetric cell division.

1.4 Thesis overview

We first present a MATLAB-based single-particle tracking tool for mobility analysis of objects in in vitro and in vivo DIC time-series microscopy. The tool integrates contrast enhancement, multiple modified Gaussian filters, automated threshold detection for segmentation, and two-dimensional single-particle tracking. We showcase its utility in quantifying subcellular dynamics, such as detecting birefringent yolk granules in *Caenorhabditis elegans* embryos and characterizing cytoplasmic flows driven by spindle oscillations. My major contributions were optimizing the granule detection kernel and comparison of existing blob detection filters with novel SoG kernel. We used this optimized detection framework to investigate evolutionary principles of cellular organization, in the context of spindle motion during the first embryonic division across nematode species. Through laser microsurgery and a viscoelastic model, we investigate the biophysical parameters influencing spindle mobility. We identify a trade-off between cytoplasmic viscosity and pulling forces, shedding light on the mechanical constraints underlying evolutionary diversification of spindle mobility.

Furthermore, to advance rapid quantification of cell cycle times we employ deep learning approach, specifically convolutional neural networks (CNNs), to classify distinct cellular stages during the first mitotic division of nematode embryos based on DIC image time-series. The successful classification enables quantification of the time spent in each stage, providing a proof of principle for the utility of deep learning in studying evolutionary diversification of cellular dynamics.

Microtubule properties are associated with spindle oscillations, cell size, and cell division regulation. Our lab previously proposed an *in vitro* system to recapitulate sustained MT oscillations under controlled motor density and MT length parameters [295]. The analysis of these oscillations could shed light on the oscillatory dynamics of MT during spindle oscillation. However, analyzing buckled MT geometry proved to be challenging. We address the challenge of filament tracking in high-density and curved contour conditions. By employing a graph-based approach, we successfully solve the problem of intersecting filaments and reconstruct filament traces. This approach proves useful in analyzing reconstituted pinned filaments exhibiting beating motion. Using this tool we explore the spontaneous oscillations driven by collective molecular motor transport of microtubules. Through *in vitro* experiments and an *in-silico* model, we uncover the role of motor binding, MT bending, and buckling in the emergence of periodic oscillations. These findings shed light on the minimal essential components and constraints contributing to the oscillations. Two additional studies are shared in the appendix. In the first work we developed a segmentation pipeline for *E. coli* in fluorescence, DIC, and phase contrast image datasets. In the last study we build a RBC detection pipeline in the microfluidic chamber. Overall, these collective research contributions underscore the important role of computational methods in advancing quantitative analysis of cellular biology.

Chapter 2

Estimating cytoplasmic viscosity of nematode embryos through particle tracking microrheology¹

¹Microscopy data was acquired by Thibault Brugière and Dr. Marie Delattre, ENS Lyon, France

2.1 Introduction

Cell viscosity quantification is important for understanding the complex nature of the cytoplasm within cells [166]. The crowded cell environment, filled with membranous organelles, RNA, ribosomes, and cytoskeletal polymers, makes the cytoplasm viscosity higher ($\sim 1\text{-}2$ cP, *Saccharomyces cerevisiae*) than that of water [167]. The high viscosity, combined with the presence of semi-flexible polymers and active processes driven by molecular motors, collectively shapes the viscoelastic nature of the cytoplasm [173, 187]. This viscoelastic nature is crucial for maintenance of cell structural integrity while allowing shape adaptation to external forces. Internally, the cytoplasm can generate size-dependent viscoelastic reactive forces which are capable of moving spindles and other passive objects back to their original position [197, 288]. These forces, can reach hundreds of piconewtons in magnitude and highlight the crucial role of bulk cytoplasmic properties in controlling organelle positioning and cellular organization [6]. Additionally, a change in cell viscoelasticity is also observed in metastatic cells [51]. These findings collectively emphasize the importance of cell viscoelasticity in cellular growth and intracellular mechanics.

In low Reynolds number regime, viscous forces dominate over inertial forces and it is expected that the mobility of intracellular organelles and structures, such as the spindle, would experience a viscous drag opposite to their motion [73, 116]. A change in cellular viscosity should therefore affect the mechanics of spindle-driven division processes. Indeed, in a simulation framework, Kozłowski et. al. have shown that increasing the cytoplasmic viscosity in their system dampens spindle oscillations [146]. However, changing the cell viscosity of cellular systems is difficult. Studies have demonstrated that changing the temperature of cells can result in alterations in cell viscoelasticity. Rapid changes in temperature makes the cell viscosity lower, making them more fluid, while long exposure affects cell elasticity [140]. The effect may be variable in different

cell types and contrasting results are observed when using different techniques [254]. Brigasse et al.'s work on *C. elegans* and *C. briggsae* demonstrated that altering the embryo culture temperature reduces the overall cell cycle time within the characteristic thermal limit of the species [13]. While qualitative changes in the mechanics of pro-nuclear migration and spindle elongation were observable, these differences were not quantified. A reduction in cell viscosity would decrease the dampening of critical diffusion processes in the cell such as reaction rates, mobility and transport thereby fastening the cell division process. However, cell systems are expected to adapt and maintain cell homeostasis with changing temperatures, as observed in plasma membranes [28, 243]. The emerging concept of visco-adaptation in cell cytosol complicates the understanding of the contributing factors to this apparent change [218]. A simplistic hypothesis would suggest that variations in the stable viscosity of the cytoplasm have resulted in apparent changes in the spindle mechanics of divergent nematode embryos. Therefore, quantification of the cell viscosity across these species could potentially establish a direct link between cytoskeletal mechanics and cellular properties.

Cell viscoelastic properties can be quantified using microrheology [172]. These techniques can measure local viscoelastic properties of the cell at a microscopic length scale. Most commonly used method for cell rheology use inert probes that are incorporated into the cytosol or the material of interest. For purely viscous materials, the Brownian diffusion of these particles can be related to cell viscosity using the Stokes-Einstein relation [71, 256]. This is known as passive microrheology. In active methods, the inert particles are caught using optical traps or are magnetic and exposed to calibrated forces, and the response of the cell reflects its viscoelastic properties. Recent work by Yamada et al. has shown that endogenous lipid granules are ideal candidates for passive microrheology, alleviating the need for exogenous probes [296]. This approach eliminates the need for cell micro injections and resulting deformations, allowing for completely noninvasive measurements of cell mechanics. Endogenous granules were

shown to have similar viscoelastic spectra to artificial beads, which was checked for gelatin, a viscoelastic medium [296]. The development of tracking methods to monitor endogenous granules in the cytoplasm of nematode embryos can be instrumental in revealing comparative differences among studied nematode species. The primary motivation for these tracking methods is to gain a deeper understanding of intracellular mechanics, particularly in living cells, without resorting to invasive or disruptive techniques.

Since the embryo of *Caenorhabditis elegans* and related nematode species are optically transparent and lack any pigments, they are ideal samples for particle microrheology. The cytoplasm of these embryos is abundant in endogenous lipid granules. Using differential interference contrast (DIC) imaging, these granules can be clearly visualized and tracked. Previously, DIC microscopy-based mutational screens for *C. elegans* have been used to understand embryonic cell divisions [102] and spindle mechanics [107]. Algorithms used to identify features in *C. elegans* embryos include a shortest path method to automatically identify embryo stages and quantify boundary dynamics during divisions [21, 97]. Similarly, active contours have been used to track spindle dynamics [47, 79]. The analysis of intracellular dynamics by segmentation-based approaches has been limited to quantifying centrosome and spindle positional dynamics [20, 81].

Here, we describe a DIC-object-tracking (DICOT) method that combines a novel filtering approach that we refer to as scaling of Gaussian (SoG). Combined with a distance-based, single-particle tracking approach, we employ it to analyze *in vivo* cytoplasmic mobility of endogenous granules in a one-cell nematode embryos to estimate cytoplasmic viscosity. Parts of this tool were already developed by members of the lab. My major contributions to this tool which were related to improving and comparing the detection accuracy of SoG with other existing filters and implementing the tracking module of the analysis pipeline.

2.2 Materials and Methods

2.2.1 RNAi experiments¹

The RNAi experiments were performed by Thibault Dr. Marie Delattre, ENS Lyons and are summarized here for completeness. RNAi experiments on *C. elegans* were performed by feeding. Wild type L4 larvae were fed for 24 hours, with HT115 bacteria producing the double-strand RNA of *atp-2* and *cyc-1* genes. The RNA interference was considered successful when the one-cell embryos did not show signs of spindle movements during mitosis.

2.2.2 Passive microrheology

Passive microrheology is a biophysical technique that allows quantification of the viscoelastic behavior of biological materials, operating on the principle that the thermal motion of microscopic particles suspended within a material reflects its viscoelastic properties [301]. Briefly, the mean squared displacement profiles are utilized to calculate the diffusivity of particles exhibiting Brownian dynamics. This setup requires non-interacting probes immersed in the biological sample, a microscopy system for recording the motion of these probes, and detection and tracking software to monitor these objects over time. Analysis of these trajectories allows for the computation of the particle's Mean Squared Displacement (MSD) profile using either time or ensemble averages, from which the diffusivity of the particles can be determined by the following relation:

$$\langle MSD \rangle = 2nDt, \quad (2.1)$$

Here, $\langle MSD \rangle$ represents the mean squared displacement, D denotes the diffusivity, n represents the tracking dimension (which can be either 2D or 3D), and t

¹Sample preparation and microscopy experiments by Thibault Brugière and Dr. Marie Delattre, ENS Lyon, France

stands for time. This equation illustrates how the mean squared displacement of particles increases linearly with time, and the slope of this relationship is related to the diffusivity. This assumes absence of any active transport processes with zero drift. The ensemble average $\langle MSD \rangle$ is calculated by the following relation [90]:

$$\langle MSD \rangle (\tau) = \sum_t (x_i(t + \tau) - x_i(t))^2, \quad (2.2)$$

Here, N is the total number of particles in the ensemble, τ is the time lag, $x_i(t)$ represents the position of the i_{th} particle at time t . For purely viscous solutions the diffusivity is related to the viscosity by the Stokes-Einstein relation [71, 256]:

$$D = k_b T / 6\eta\pi a, \quad (2.3)$$

where K_b is the Boltzmann constant, T is the temperature, a is the granule radius and η is the fluid phase viscosity.

Granule density correction

Granule radii (r) are very similar $\sim 0.2\text{--}0.3\mu m$ across species (Figure.2.1(D)) and are used in the estimate of viscosity for each species. To account for the crowded nature of the cytoplasm due to granule packing, the effective viscosity (η_{eff}) was estimated based on the approximation for soft spheres [223], known as Quemada correction, by accounting for the packing fraction of the embryo due to the high 2D granule-packing fraction (ϕ). This approach is better than the alternative of Dougherty kreiger method which is suited for dilute suspensions [148]. The granule-packing fraction for each species was calculated from eight representative ROIs each from the anterior and posterior regions, with ~ 400 granules per species. The granule density per unit area,

$$\rho_g = N_g / A_{cell}, \quad (2.4)$$

where A_{cell} is the embryo size and N_g is the number of granules. The area of each granule (g) was measured from the granule radius to arrive at the granule packing fraction:

$$\phi_{2D} = \rho_g \times A_g, \quad (2.5)$$

was used to calculate the effective viscosity:

$$\eta_{eff} = \eta_f \times \left(1 - \frac{\phi_{2D}}{\phi_{max}}\right)^{-2}, \quad (2.6)$$

where ϕ_{max} is the maximal packing fraction, taken to be 0.64 for random packing [30], and η_f is the fluid phase viscosity.

2.2.3 Image acquisition of nematode embryos and strain maintenance²

All strains were maintained at 20°C on nematode growth medium seeded with *E. coli* OP50, as described in [272]: *Caenorhabditis elegans* (N2), *Caenorhabditis remanei* (PB219), *Caenorhabditis monodelphis* (SB341), *Diploscapter* sp. 1 (JU359), *Oscheius tipulae* (CEW1), *Pristionchus pacificus* (PS312). For embryo recording, females were dissected in M9 and one cell-stage embryos were placed between slide and cover-slip on a 2% agar pad. Embryos were observed with a Zeiss Axio-imager A1 or A2 with a 100× DIC Plan Achromat NA 1.4 lens. For video recording of the cell division, we took two images per second with a digital Kappa camera (DX4-285FW)

²Sample preparation and microscopy experiments by Thibault Brugi re and Dr. Marie Delattre, ENS Lyon, France

2.3 Results

DICOT is a GUI based tracking tool that implements a Scaling of Gaussian (SoG) filter to detect lipid granules present in the cytoplasm of nematode embryos [39]. The SoG filter scales a standard Gaussian filter (h_G) with the mean (2.7). The formulation is given as:

$$h_{SoG} = h_G - \langle h_G \rangle (\lambda - \phi), \quad (2.7)$$

where $\langle h_G \rangle$ is the mean of the gaussian kernel h_G . When the image contains bright objects against a dark background, typical for both DIC and fluorescence, image λ is set to 1, whereas it is set to 0 for dark objects against a bright background. The sensitivity factor (ϕ) is used to enhance the degree of contrast between the foreground and background. To compare the performance of SoG with other detection filters such as Difference of Gaussian (DoG) and Laplacian of Gaussian (LoG) we first created a ground truth dataset of manually annotated lipid granules.

2.3.1 Annotation of cytoplasmic granules as ground truth to quantify detection accuracy

A blob detection pipeline can be used to identify the birefringent yolk granules in a DIC image time series of *C. elegans* embryos as identifiable by eye [142]. The enhanced contrast achieved by Scaling of Gaussian (SoG) SoG filtering allows object segmentation by automated thresholding. However, there are existing filters such Laplacian of Gaussian (LoG), Difference of Gaussian (DoG) that can be used similarly to SOG. To compare and contrast the extent of correctly detected yolk granules, we annotated six randomly selected frames in a *C. elegans* time series from two representative regions of interest (ROIs), one anterior and one posterior, to account for any spatial differences Figure. 2.1(A), and marked yolk granules manually to serve as the GT to measure

detection accuracy Figure.2.1(B). Total 450 annotations were made , ~ 75 per roi, and saved for further analysis. In addition to this manual annotation we also quantified the granule radii across the 6 species of nematodes in our dataset Figure.2.1(D).

The detection output from DICOT analysis is classified based on the euclidean distance between the granule coordinates detected by the algorithm and the annotated coordinates from the ground truth into the following classes: True positives (TP); distance of the centroid identified by the algorithm lies within a threshold (t_{dist}) distance of the interactively annotated point Figure.2.1(C). A t_{dist} of two pixels was chosen to account for human error while selecting centroid of the granule. False positives (FP); detection obtained from the algorithm at a distance to the GT greater than t_{dist} are classified as false positives. False negatives (FN); all manual annotations that were not detected by the algorithm are classified as false negatives. To quantify how well DICOT works at object detection, we estimate three variables: sensitivity (S), precision (P), and the $F1$ score. S is defined as the ratio of correct detection by the algorithm (TP) and all detection both true and false, as described by the following expression:

$$S = TP / (TP + FN) \quad (2.8)$$

An alternative measure, P is the fraction of correctly classified detections out of total detections from the algorithm as given by the following:

$$P = TP / (TP + FP) \quad (2.9)$$

Finally, the F1-score (F1) combines both measures of P and S as follows:

$$F = 2 \times P \times S / (P + S) \quad (2.10)$$

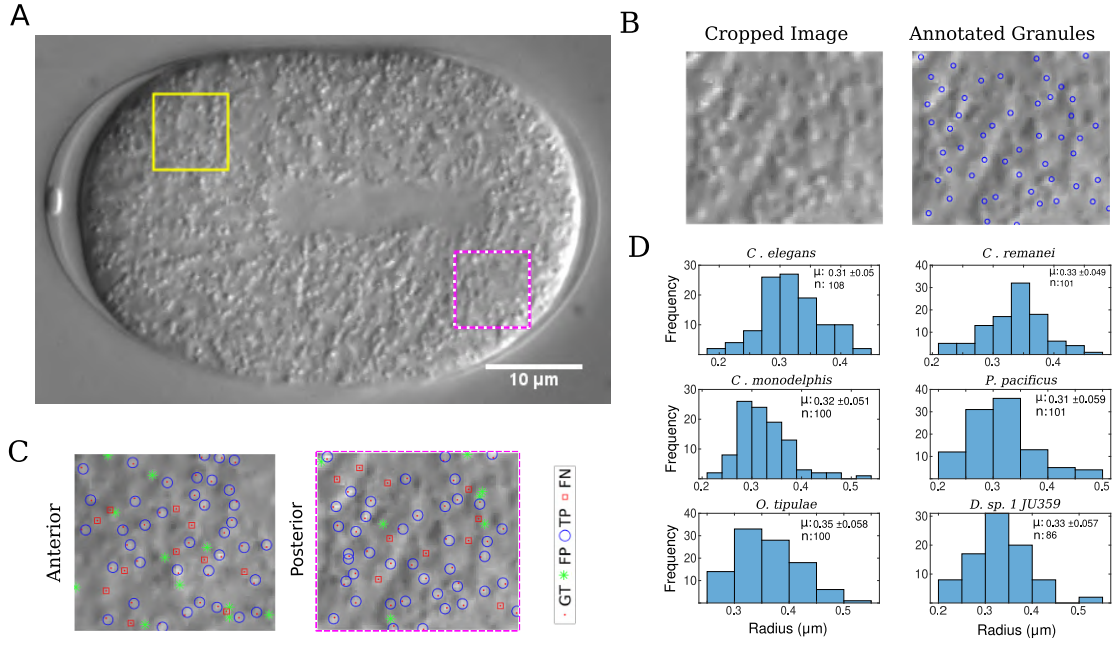


Figure 2.1: **Precision of yolk-granule detection in DIC.** (A) ROIs from the anterior (line) and posterior (dashed line) regions of a DIC image time series of a *C. elegans* embryo were used to estimate. (B) Representative ROI with the annotations shown in blue dots. (C) the ground truth (GT; point) through manual annotation of granules and classify a detection based on the algorithm as either true positives (TP) (open circle), detected by the algorithm, as well as manually false positives (FP) (asterisk), those identified by the algorithm but not manually, and false negatives (FN) (open square), those that were identified manually but the algorithm failed to detect. (D) Granule radii estimate for the analyzed species

2.3.2 Filter parameter optimization and comparison

The quality of single-particle tracking (SPT) depends on the accuracy of the object detection. To achieve high accuracy of detection, we compared multiple filtering approaches; Gaussian, inverse Laplacian of Gaussian (iLoG), difference of Gaussian (DoG), and the scaling of Gaussian (SoG). These filters were then combined with segmentation by thresholding based on Otsu’s method to detect filtered objects in DIC images [207]. The SoG filter is a simpler version of the DoG filter $h_{DoG} = h_{G1}(\sigma_1) - h_{G2}(\sigma_2)$, where h_{G1} and h_{G2} are two Gaussians, and σ_1 and σ_2 define the respective spread. For a DoG filter when $\sigma_2 \gg \sigma_1$, the difference approximates to a SoG kernel. At the same time, the LoG filter, based on the second derivative of a Gaussian, can also be mapped to DoG at specific values of Gaussian filters [7]. However, the simplicity in using SoG scaling on a Gaussian by a scalar value, by the use of a single ϕ parameter and estimation of filter size from object size, as compared with multiple or derivatives of Gaussian, suggests a simplicity that could be more user-friendly for optimization in the analysis of diverse experimental data sets. A visual comparison of these filters along with the segmentation and detection results is shown in Figure.2.2 (A-C). For a fair comparison all parameters of different compared kernels were optimized using an exhaustive search of optimal parameters within predefined limits shown in Figure.A1 (A-D). The optimized parameters for each filter are as follows: SoG : ($k= 7$, $\sigma = 0.77$, $\phi = -0.009$), Gaussian : ($k= 3$, $\sigma 0.1$), LoG : ($k= 6$, $\sigma 0.40$), DoG : ($k_1, k_2= 11$, $\sigma_1 = 1.63$, $\sigma_2 = 0.94$)

In our analysis of positional detection errors, as shown in Figure.2.3, we find that the mean positional detection error is lowest for the SoG filter when compared to other tested filtering methods, as indicated by both mean and median values, and also with regard to outliers (Figure.2.3(A)). To avoid any biases, we scanned multiple parameters for each filter systematically across a range of comparable values and used those parameters that maximized the F1 score. Although the error resulting from

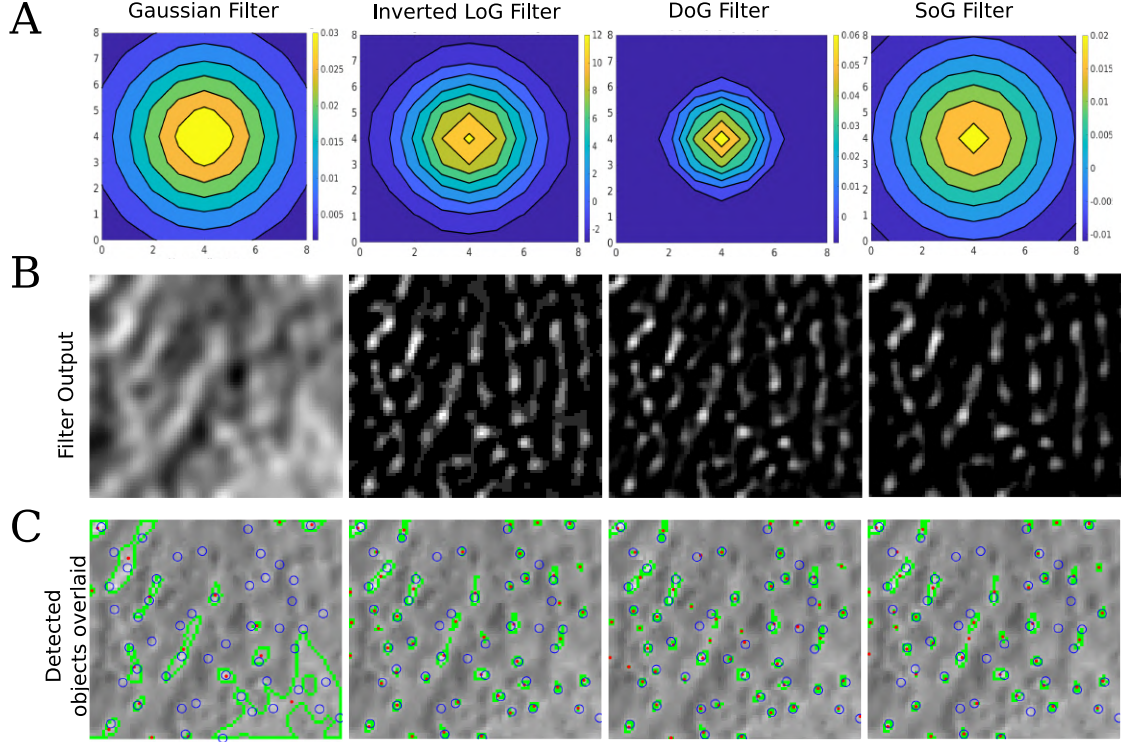


Figure 2.2: **Comparing filters for DIC object detection.** (A) An ROI from a mid-plane DIC image of a *C. elegans* embryo was manually annotated to mark lipid granules (blue circles). (B) Four different image filters were tested: Gaussian (blue), Inverted LoG (red), DoG (ochre) and SoG (purple). (C) Contour maps of multiple filters used to convolve the data are compared: Gaussian $\sigma = 2.25$, inverted Laplacian of Gaussian (iLoG) $\sigma = 2.25$, difference of Gaussian (DoG) taken between two functions with $\sigma = 1.25$ and 2.25 and Scaling of Gaussian (SoG) $\sigma = 2.25$. All filters have the same kernel size, i.e. 9 pixels. (C) The filtered output images (D) are then segmented by an automated threshold and the contours of detected objects (green) are used to estimate centroids (red dots) and compared to manual annotations (blue circles).

filtering by SoG, iLoG, and DoG is comparable, we find increasing noise in the image changes the positional detection error with SoG and iLoG, showing the smallest error and thus, greater robustness compared with the other filters Figure.2.3 (B-C). Speckle noise with variance range from 0 to 0.08 was added using *imnoise* function. Visual overlay of the filter design and the filtered output is shown Figure.2.2. This suggests that our method is optimal for the analysis of *in vivo* time series of *C. elegans* embryos. We additionally implemented a regional max search over the filtered objects to further improve the sensitivity of our detection method as shown in Figure.A2(a-c). This improvement is attributed to the identification of distinct peaks in cases where two closely situated granules are not separated during segmentation.

2.3.3 Fast nearest neighbor based particle tracking

To track detected particles across frames we implemented a nearest neighbour tracking algorithm [203]. The algorithm first identifies the closest neighbor in successive frames to track the movement of detected objects. The input parameters to the tracking algorithm include search radius in microns ($4\ \mu m$), scaling factor ($0.065\ \mu m$), minimum track length (2 s), time interval between frames (s), distance unit (μm), and time unit (s). The micron search radius defines the radius of a circle within which the algorithm searches for the next potential object detection. Objects falling within this defined circle are considered candidates for association in the tracking process.

For each frame, the algorithm associates objects between consecutive frames based on their distances. The association is performed using a custom sorting function, which considers the distances and a specified pixel search radius. Objects are labeled and tracked over time, forming trajectories. Trajectories with less than three data points are excluded from further analysis, ensuring that only well-connected object paths are considered. This can be tuned using the minimum track length parameters. Finally, the algorithm writes the reassembled trajectories, including object IDs, frame

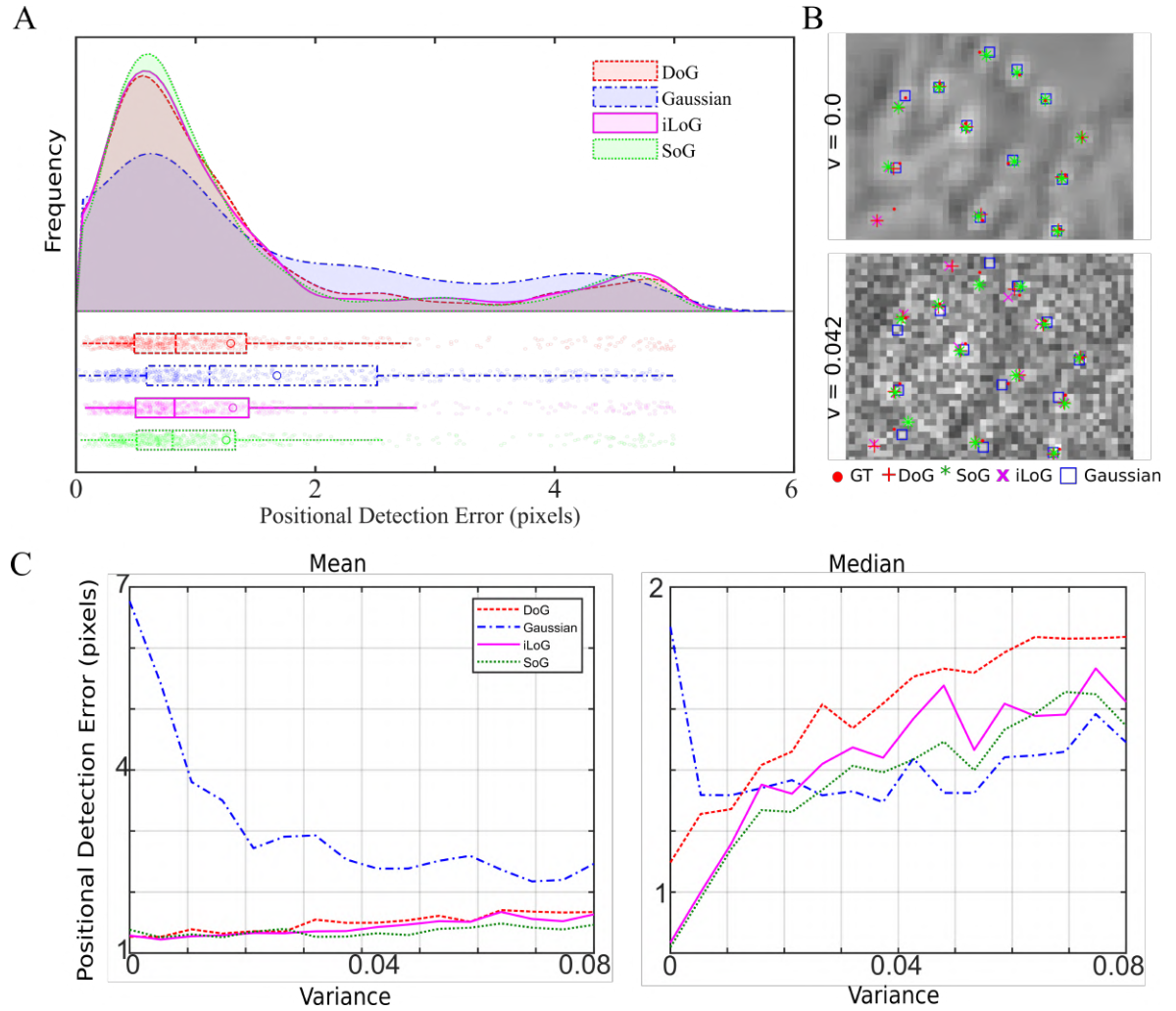


Figure 2.3: **Precision of yolk-granule detection in DIC.** (A) Shown is the frequency distribution of the positional detection error in pixels based on filtering by either Gaussian (dash-and-dot line), inverted iLoG (solid line), DoG (dashed line), or SoG (dotted line) is plotted with (bottom) the box-plot representing the data scatter. Here, the following symbols have the described meaning: open circle, mean; midline, median; and whiskers, minimum and maximum. (B) The effect of multiplicative speckle noise on positional detection error was examined for a range of variance values of the noise function. Two representative images of variance 0 and 0.042 were filtered based on the optimal parameters for detection by filtering and segmentation based on either Gaussian (open circle), iLoG (cross), DoG (plus sign), or SoG (asterisk) filters. The error is the distance from the GT (solid circle). (C) The mean (left) and median (right) positional detection error for increasing variance of noise ranging is plotted. To see this figure in color, go online.

numbers, X and Y coordinates, object area (area), and time to an output text file named 'trajectories.txt' in the specified output folder. Based on the input parameters distance and time metrics are converted to the suitable scale. This output file provides the raw record of the tracked objects movement and is saved for further analysis.

2.3.4 Cytoplasmic viscosity changes six-fold between closely related species

In DIC images, nematode embryo cytoplasm is prominently packed with clearly visible yolk granules [46, 119]. We proceeded to estimate the cytoplasmic viscosity in the different species using granule microrheology, similar to a previous study on *C. elegans* embryo, where exogenous particles were used as probes [107]. Since granules undergo streaming due to spindle movements and are at times even actively transported, we chose to follow granule mobility during interphase and in the top plane of the embryos far from the spindle plane to minimize the effects of active transport on granule mobility measurement. Based on the distinct active motion in mid-plane contrasted to the diffusive motion in the top plane, all further analysis of granule diffusive motility was performed only in the top plane. A representative MSD plot from the embryo mid-plane showing super-diffusive MSD profile is shown in Figure.A3. In contrast the MSD profiles from the embryo mid-plane were linear as shown in Figure.2.5 (C). MSD profiles for the anterior and posterior region for one sample from every species is shown in Figure.2.4. Each embryo in the time series was partitioned into anterior, middle and posterior regions and over 2,000 granules in the anterior posterior portion of each embryo were tracked. Multiple embryos of each species were analyzed and the mean squared displacement (MSD) of granules was calculated as described in the methods section. Granule MSD plots were averaged over time and across multiple granules and fitted to the diffusion model equation (2.1).

Individual granules display heterogeneity in individual MSD profiles, with some granules appearing to be superdiffusive, while others appear to follow sub-diffusive behavior. On the average, they appear diffusive, due to the linear nature of the MSD plot at short time scales Figure.2.5(C). We proceeded to fit diffusion equation (2.1) (Materials and Methods) to the data in order to estimate a single free parameter, the diffusion coefficient D . This result is shown in Figure.2.5(D). We find that the analyzed species have strong differences in the diffusivity of lipid granules with *C. monodelphis* and *D.sp.1 JU359* showing slowest mobility of endogenous granules. On the other hand *P. pacificus* had the highest diffusivity in the cytoplasm. This could be due to transient flows that occur predominantly in the posterior region. To estimate the effect of these flows on our D measurement we compared the calculated D with the diffusion coefficient estimate from anomalous diffusion model D_{adm} and found them to be non-significantly different Figure.A4.

To estimate the fluid phase viscosity (η_F), which implicitly combines cytoplasmic viscosity and the effect of cytoskeletal elements and organelles, we measured the radius of lipid granules from each species averaged across multiple individuals and found them to be comparable Figure.4.1(D). Using the Stokes–Einstein (eq.(2.3)) we calculate the value of viscosity [14, 72]. Due to the visually apparent crowding of granules in the cytoplasm, we need to include the effect of crowding by the highly packed yolk granules where packing density was corrected for their density, based on the theory of the effect of soft-sphere packing in a crowded colloidal suspension on viscosity [223]. An additional motivation for applying this correction is that the granules have a high packing fraction (ϕ_{2D}) and a diameter of $\sim 0.3\mu m$, comparable to the size of the centrosomes (diameter $\sim 2 - 4\mu m$). Indeed, granules occupy between 29 and 42% of the cell area comparable between species Figure.2.5(E). The packing fraction is then used to estimate the effective cytoplasmic viscosity η_{eff} , Figure.2.5(F). The packing fraction is similar for the top and the mid-plane Figure.A3. We find it to be 6 to 10-fold

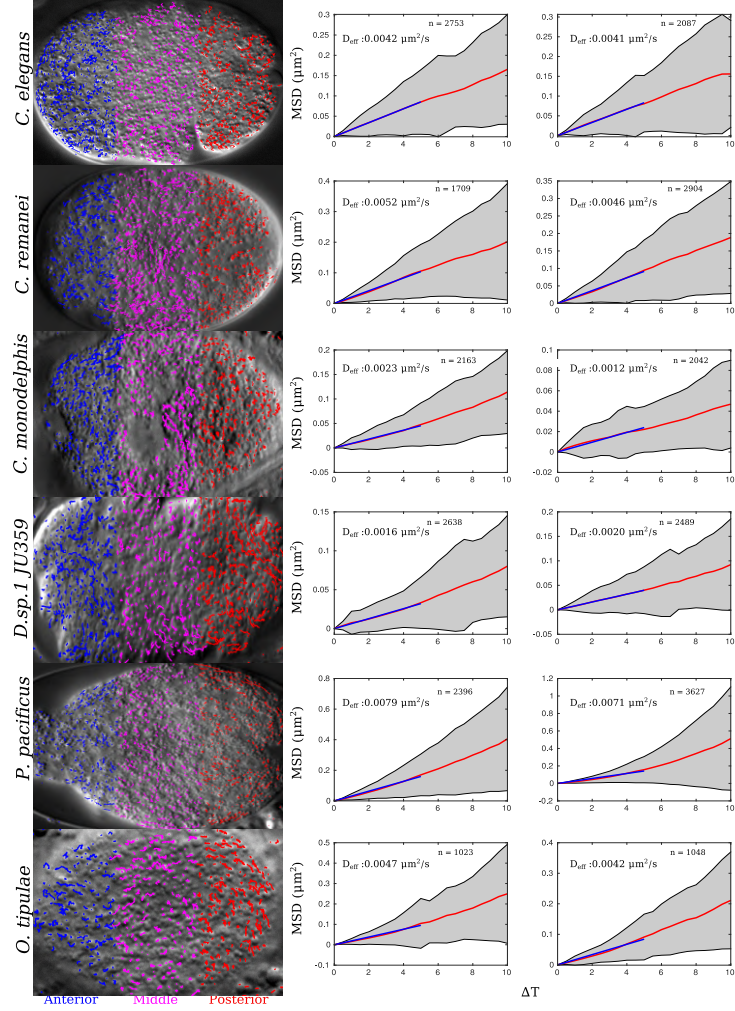


Figure 2.4: **Measuring granule diffusion coefficient for multiple species.** (left) Single images in DIC of embryos of the 6 species analyzed are overlaid with the tracks of granules from individual species. Colors indicate blue: anterior, magenta: middle and red: posterior regions based on 33% of the major axis length. (right) The MSD values of granule tracks from the anterior (left) and posterior (right) regions are plotted and fit to the diffusion model (till 5s) to the mean MSD curve. The values obtained by fitting the ADM model (blue line, Equation 8, Materials and Methods) to the mean MSD curve (red line) are D_{eff} and α . The number of trajectories analyzed are indicated by n.

higher than the fluid phase viscosity η_F , accounting for both fluid-phase and crowding effects. The mean *C. elegans* effective cytoplasmic viscosity η_{eff} of 0.67 Pa-s is comparable to previously reported values, which ranged from 0.1 Pa s [55, 93] to 1 Pa s [55].

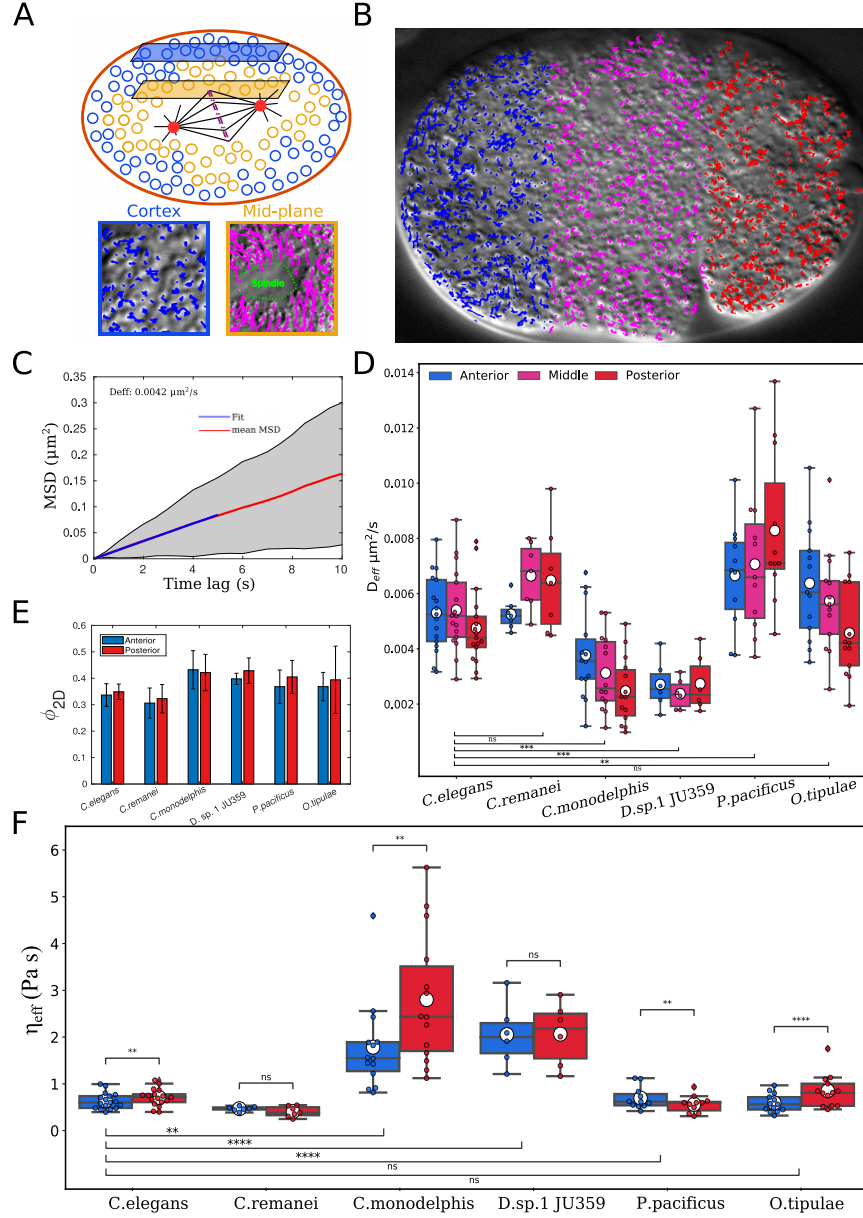


Figure 2.5: **Cytoplasmic viscosity estimates from yolk granule mobility.** (A) Yolk granule mobility in the one-celled stage of the nematode embryo from midplane (yellow circles) and cortical (blue circles) is depicted. Lines: tracks of granules, midplane: yellow box, and cortex: blue box. (B) Granule movement was automatically tracked from DIC microscopy time series of a *C. elegans* embryo imaged in the cortical plane. Lines: tracks, blue: anterior, magenta: middle, and red: posterior. (C) The mean squared displacement (MSD) averaged over granules from the whole embryo is plotted as a function of time interval (red line). Data are fitted to a diffusion model (blue line, Equation 4, Materials and Methods). Gray region: SD across individual trajectories ($n \sim 4,000$). Fit parameter $D_{eff} = 0.0042 \mu m^2/s$. (D) The D_{eff} of anterior, middle, and posterior cortical granules from all six species are represented as box plots. (E) The area packing fraction of granules (ϕ 2D) from the anterior (blue) and posterior regions (red) is plotted for the six nematode species analyzed (mean $\pm$ SD). (F) The effective viscosity η_{eff} of each nematode species from the anterior (blue) and posterior (red) regions is plotted as a box plot. White circle: mean, horizontal line: median. Points indicate outliers. The asterisks indicate * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, and **** $p < 0.0001$, while ns = not significant for t tests (paired within species, independent between species.)

This estimate of viscosity is based on the hypothesis that diffusion is the primary process driving granule motion, that is, passive, thermal Brownian motion. As a further test of the influence of active transport on diffusion, we examined the effect of ATP depletion on granule mobility, which is expected to block active intracellular transport. If D_{eff} is truly representative of the passive Brownian motion of granules, ATP depletion should not change the diffusion coefficient. Therefore, *C. elegans* embryos were treated with RNAi against *atp-2* (ATP synthase subunit) or *cyc-1* (Cytochrome-c1) as seen in Supplemental Figure.2.6(A), which was shown to inhibit mitochondrial ATP production [64, 270]. While it is expected that depleting ATP will affect all physiological processes in a nonspecific manner, the yolk granule mobility of treated compared with untreated *C. elegans* embryos was found to be quantitatively unchanged in terms of MSD profiles Fig.2.6(B), the diffusion coefficient Figure.2.6(C) and viscosity Figure.2.6(D). Taken together, the evidence strongly supports the hypothesis that granule mobility we describe is indeed diffusive.

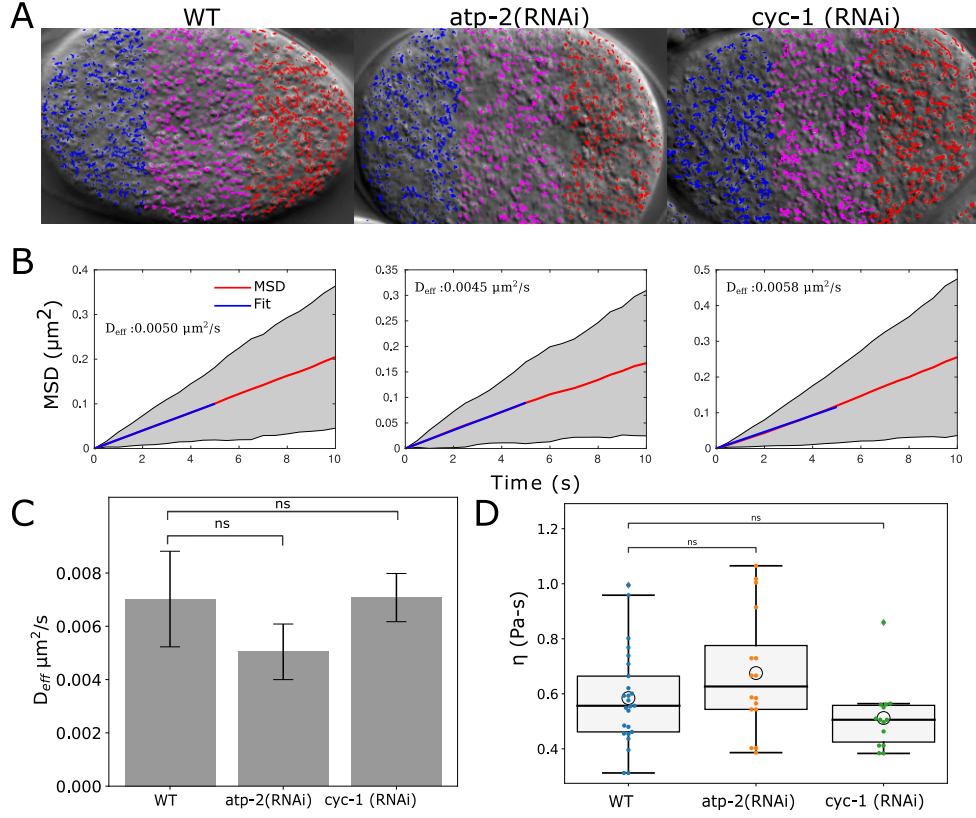


Figure 2.6: **Effect of ATP depletion on granule mobility.** (A) Representative tracked granules from WT are compared to embryos treated with RNAi targeting ATP-2 and CYC-1 are overlaid on DIC images of the embryo in the anterior (blue), mid-cell (magenta) and posterior (red) regions. (B) The MSD with time is plotted for whole embryos corresponding to (A). The grey area is the s.d. The data was fit using a diffusion model (blue) to estimate the effective diffusion coefficient (D_{eff}). (C) The mean D_{eff} (error bar: 2 SEM) and (D) effective viscosity are compared between untreated and treated embryos. The mean values were compared using a KS test; ns: not significant.

2.4 Conclusions

We find that effective cytoplasmic viscosity (η_{eff}) differs between species. *C. monodelphis* shows the highest viscosity of 2.8 Pa s (Figure.2.5(F); Table 2.1). We also find that viscosity differs greatly between anterior and posterior regions in *O. tipulae*, and to a lesser extent in *C. monodelphis* and *P. pacificus*, and only minimally (factor of 1.1) in *C. elegans*. The viscosity in the anterior and posterior halves of the embryos

was, however, similar in *C. remanei* and *D. species 1* (Figure.2.5(F), Table 2.1). These results reveal that even closely related species show a diversity in cytoplasmic viscosity, with an ~ 7 -fold difference between the most extreme values measured.

Species	Fluid phase viscosity of cytoplasm, η_F (Pa s)		Granule radius (μm)	Packing Fraction, ϕ_{2D}		Effective cytoplasmic viscosity corrected for packing fraction, η_{eff} (Pa s)	
	Anterior	Posterior		Anterior	Posterior	Anterior	Posterior
<i>C. elegans</i> (n=19)	0.14 ± 0.04	0.16 ± 0.04	0.32	0.33	0.35	0.63 ± 0.17	0.71 ± 0.18
<i>C. remanei</i> (n=8)	0.13 ± 0.01	0.11 ± 0.03	0.34	0.29	0.33	0.47 ± 0.05	0.41 ± 0.11
<i>C. monodelphis</i> (n=15)	0.22 ± 0.12	0.35 ± 0.17	0.32	0.42	0.41	1.78 ± 0.97	2.80 ± 1.40
<i>D. sp. 1. JU359</i> (n=5)	0.27 ± 0.10	0.27 ± 0.10	0.33	0.4	0.42	2.08 ± 0.75	2.07 ± 0.76
<i>P. pacificus</i> (n=13)	0.12 ± 0.04	0.09 ± 0.03	0.31	0.4	0.4	0.70 ± 0.24	0.56 ± 0.18
<i>O. tipulae</i> (n=14)	0.11 ± 0.04	0.15 ± 0.06	0.36	0.4	0.4	0.59 ± 0.20	0.85 ± 0.35

Table 2.1: **The viscosity of nematode embryonic cytoplasm from six species.** The estimation is based on tracking the diffusion of ~ 1000 lipid granules (per species) to calculate MSD profiles. The fluid phase viscosity (η_F) for all six species was calculated from the effective diffusion coefficient obtained by fitting the MSD to the MSD profiles. To account for self-crowding by a high density of lipid granules, we measured their radii (μm) and area packing fraction (ϕ_{2D}) in order to correct the viscosity to arrive at the effective viscosity η_{eff} (Pa s). Between 8 and 19 embryos were analyzed per species, indicated by n. All values of viscosity are mean $\pm$ SD.

While the increase in viscosity across species (see Figure.2.5(F)) correlates qualitatively with slower spindle dynamics in *C. monodelphis* and *D.sp.1 JU359*, it is unclear if viscosity alone can explain, for example, the absence of bonafide spindle oscillations in some species such as *P. pacificus*, whose cytoplasmic viscosity is very close to that of *C. elegans*. Nonetheless, the quantification of cytoplasmic viscosity differences allows us to measure the viscous drag experienced by large objects moving inside the cytoplasm, such as the mitotic spindle. Using laser ablation, we can estimate cortical forces acting on the spindle during its recoil. We will discuss these results in the next chapter.

2.5 Discussion

The intracellular aqueous environment of macromolecules and presence of cytoskeletal proteins has made the study of viscous and elastic properties of cells vital for our understanding of cellular mechanobiology. However, how much these properties can change without affecting the functions they sustain is still unclear. This is in part because experimental perturbations of these properties are challenging. An alternative approach is to measure the physical properties of cells from different species, because they reveal the range of solutions retained over the course of evolution, and hence the range of possible changes.

Early work measuring cytoplasmic viscosity in vertebrate cell lines reported a viscosity of 0.282 Pa.s by following the Brownian motion of cytoplasmic inclusions [4]. A more recent study using micrometer-sized beads and magnetic tweezers to measure the viscoelastic properties of the cytoplasm in *C. elegans* during spindle positioning also reported viscosity of 0.159 Pa.s [93]. However, tracking the microrheology of the mobility of injected nanospheres in one-celled embryos of *C. elegans* revealed a spatially uniform cytoplasmic viscosity of 1 Pa.s [55]. Accumulating evidence of the probe size dependence of such measurements due to macromolecular crowding [74, 190] could explain this order-of-magnitude difference reported by different workers in the same species. Our measurements retrieve a viscosity value that is in between these two extremes for *C. elegans*. We report the effective cytoplasmic viscosity to be ~ 600 times higher than the viscosity of water. Using the same approach to multiple species, we have uncovered large variations, up to an order of magnitude, between species, with *C. monodelphis* showing the highest viscosity. The differences we measure could arise from multiple factors such as differences in protein concentrations, presence of different densities of cytoskeletal meshwork, or higher organelle densities.

There is a slight but non-significant increase in viscosity observed in ATP-2 RNAi-

treated embryos (see Figure.2.6(C)), which may suggest a gradual change in cytoplasmic viscosity. Notably, increased cell viscosity is associated with cell death in yeast cells [135]. In this study, the efficacy of ATP knockdown was qualitatively assessed by the cessation of spindle motion and embryo mortality. However, the time elapsed between treatment across the embryos sample was not measured. To better understand why granule mobility and measured viscosity remain unchanged, it would be crucial to track these endogenous granules over extended time in RNAi treated embryos and measure if there are gradual changes to cell viscosity over time. Additionally, this investigation could be complemented by techniques to measure cytoplasmic ATP levels, such as luciferase-based assays [77, 168], providing a more concrete understanding of the interrelationship between ATP levels and cytoplasmic viscosity. Another important consideration is the nature of transport of these endogenous lipid granules. Existing evidence suggest that yolk granules in egg laying animals function as heterogeneous organelles that take part in embryo development [224]. However, the molecular machinery and mechanisms for yolk mobilization throughout development remain poorly explored [224]. Therefore, a more detailed investigation into the transport of lipid granules during embryo development would help in understanding the mobile nature of these yolk granules.

Multiple studies have demonstrated that changing the temperature of cells results in alteration of cell viscoelasticity [140, 254], however a similar investigation is lacking for the nematode system. Briggase’s work on *C. elegans* and *C. briggsae* has demonstrated that altering the embryo culture temperature reduces the overall cell cycle time within the characteristic thermal limit of the species [13]. While qualitative changes in the mechanics of pro-nuclear migration and spindle elongation were observable, these differences were not quantified. In the future, a careful comparison of changes in granule mobility, spindle dynamics, and cell cycle time between species with changing temperature could potentially allow us to address the question of how these differences in

spindle dynamics could arise. Nonetheless, to address our initial question of potential differences that could explain the variations in spindle dynamics of the evolutionary related species, we proceeded to quantify the forces acting on the spindle during the first mitotic division.

Chapter 3

Evolutionary divergence of spindle mechanics in nematode embryos is constrained by antagonistic pulling and viscous forces

3.1 Introduction

Our quantification of spindle motion across nematode species revealed that spindle transverse oscillations are restricted to *Caenorhabditis* species [272]. Only in *C. monodelphis* which is the most basal *Caenorhabditis* species, and in all species outside of this genus, the anaphase spindle is asymmetrically displaced without any transverse oscillations. Our initial attempt was to understand if viscosity change accounts for this variability. However, this simple hypothesis proved to be insufficient explanation of absence or presence of oscillations.

The work of Grill et. al, 2005 presents a theoretical framework to explain spindle oscillations during first cell division for *C. elegans* embryo [109]. Key to the model is the concept of load-dependent detachment rates, where force generators detach from microtubules at a rate depending on the load force exerted on them. This load force is determined by the extension of elastic linkers connecting force generators to the cell cortex. By considering the interplay between attachment and detachment rates, the model predicts the occurrence of spindle oscillations beyond a critical number of force generators. Furthermore, the model is supported by experimental evidence showing that depletion of G proteins (encoded by *gpr-1* and *gpr-2*), that regulate cortical force generators, abolishes or attenuates oscillations, suggesting a threshold of motor activity necessary for oscillations to occur [216]. The oscillatory behavior is therefore described by key parameters of a threshold number of motors and microtubule contacting the cortex. Around the same time, Kozłowski et al. (2006) predicted that a low cytoplasmic viscosity was also required for oscillations with speeds greater than $0.2 \mu\text{m}/\text{s}$, since higher viscosity resulted in a strong damping force opposing the motor-forces [146].

While our understanding of spindle mechanics in *C. elegans* is relatively comprehensive, elucidating similar mechanisms across species presents challenges. Our initial effort to quantify cytoplasmic viscosity changes is incomplete without the quantifica-

tion of net pulling forces across various species, because spindle oscillation buildup may be hindered by an increase in cytoplasmic viscosity or by a reduction in cortical forces arising from gene expression changes or reduced cortical localization. Therefore, our next step is to quantify the net pulling forces across different species.

Mechanical forces acting on the spindle can be revealed by laser ablation of the central spindle at the onset of mitosis. Following spindle cutting, the centrosomes recoil towards the cell pole if they are initially pulled [107]. This is because the central spindle connects the poles and holds the balance, much like a stretched rubber band. In species where the mitotic spindle elongates by inside-out pushing forces, severing the spindle leads to the collapse of the centrosomes at the center of the spindle [139, 267]. Subsequently, parameters of spindle velocity, force and elastic components can be quantitatively measured. In this chapter, we use a visco-elastic model to measure the cortical pulling forces responsible for spindle motion in all species. We then ask which changes in the force, viscosity and recoil kinetics across species correlate with spindle motion kinetics. This comparative approach could enable us to gain a deeper understanding of the evolutionary conserved mechanisms underlying spindle oscillations and their role in asymmetric cell division.

3.2 Material and Methods

3.2.1 Laser ablation¹

One-cell embryos in prophase or prometaphase were mounted between slide and coverslip as described above. Embryos were then recorded on an inverted spinning disk confocal microscope (Leica DMI4000B- CSU 22 Yokogawa) with a 100X immersion objective (HCX PL APO 1.4 oil) controlled by Metamorph. Images were acquired with an iXon3 897 Andor camera every 0.5 seconds. Spindle severing was performed using

¹See 1

a UV laser module ($\lambda=355$ nm) iLas2 Roper, as described in [107]. For each species, the laser power was adjusted so that the cut, performed at the onset of spindle elongation, generated a rapid movement of the centrosomes (due to spindle severing), but did not arrest the cells (due to excess laser power).

The number of samples analyzed were: *C. elegans*: 14, *C. remanei*: 14, *C. monodelphis*: 14, *P. pacificus*: 17, *O. tipulae*: 13, and *D. sp. 1 JU359*: 9. The data had to be processed, and stationary trajectories that showed no recoil were removed. Before analysis, we removed all missing values from the recoil data. All recoil trajectories were manually aligned with the zero point, which marks the laser ablation treatment. The anterior trajectories were flipped by taking their absolute values to overlay with posterior trajectories for a better visual comparison. Trajectories with noise that appeared frequently near the end of every trajectory were manually removed. The raw recoil trajectory along with time stamps of before and after laser ablation for representative species is shown in Figure B1(A and B).

3.2.2 Oscillation speed

Spindle oscillations were analyzed in representative trajectories of each species by measuring distance of the centrosomes from centerline, defined by the A-P axis, with time and the data was smoothed using the discrete cosine transform to reduce low-frequency noise with the threshold greater than 1.5 for *C. monodelphis* and *O. tipulae*, and 0.56 for the remaining species, as shown in Figure.3.1(A). The speed of oscillation was calculated as the change of position over successive windows in 5 second intervals, throughout the smoothed trajectory and the distribution was used to estimate mean oscillation speed for anterior and posterior centrosomes. Data smoothing was performed using inverse discrete cosine transform of frequency components with values higher than 0.56. The smoothed trajectory and speed over Successive frames for *C. elegans* is shown in Figure.B2(A and B). The oscillation speed was calculated for each

species using the Fast Fourier Transform (FFT) analysis, the power spectrum of frequency is shown for *C. elegans* and *C. remanei* is shown in Figure.3.1(B). Spindle oscillation speed comparison is shown in Figure.B2(C).

3.2.3 Data analysis and fitting

The initial recoil velocity and rate of decay from the recoil trajectories was obtained by fitting the anterior and posterior centrosome recoil trajectories to the function:

$$f(t) = A * (1 - e^{-t/\tau}) + c \quad (3.1)$$

where A is amplitude of recoil and τ is decay constant of the exponential and V is the recoil velocity given by $V = A/\tau$ based on previous work [252]. Oscillation speed predictions from simulations as a function of increasing viscosity were taken from previous work by Kozlowski et al., (2007) by digitizing the plot (webplotDigitizer) and fitting it to a 4-parameter sigmoid function described previously [137, 146]:

$$v = d + a/(1 + \exp^{-\frac{(\eta-c)}{s}}) \quad (3.2)$$

Where v is the speed, η is the viscosity, d and a are the minimal and maximal speeds, c is the half-maximal viscosity and s is the steepness of the profile. Individual data from centrosome recoil of all species were ignored if the goodness of fit measure (R²) was less than 0.8, except for *D.sp.1* where this cutoff was 0.5. As a result of this we do not report v and τ values for this species (Figure.3.2(C, D)).

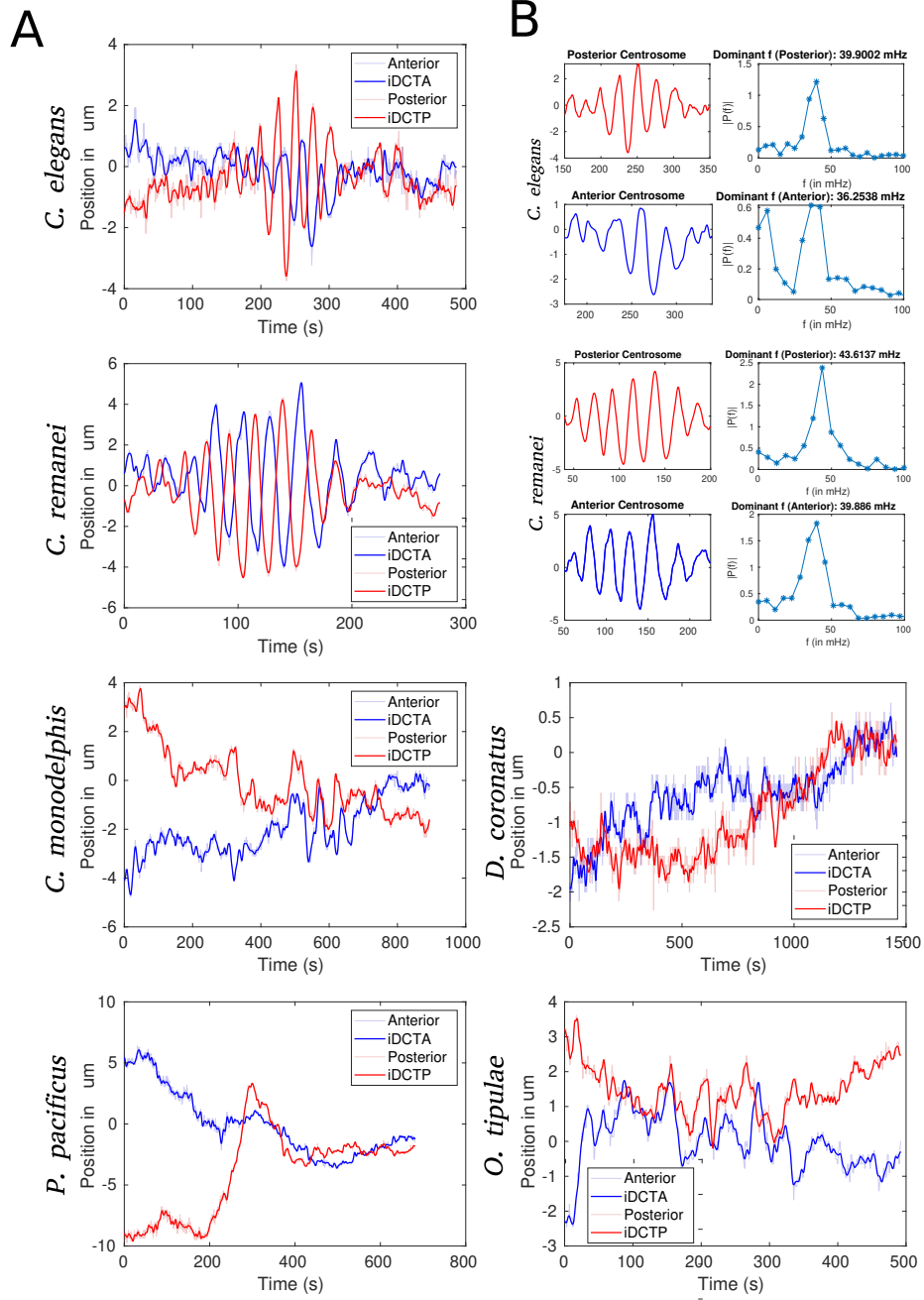


Figure 3.1: **Spindle elongation trajectories from the analyzed nematode species.** (A) Raw trajectories (solid) lines for anterior (blue) and posterior (red) are shown for every species. The inverse discrete cosine transform (iDCT) is shown in the background. (B) Oscillation frequency analysis using FFT analysis is shown for *C. elegans* and *C. remanei*. On the left the input trajectory is shown and on the right the power spectrum with the dominant peak as the highest point.

3.2.4 Statistical tests and correlation

All correlations are reported as Pearson correlation coefficients [237]. Regression fits for all correlations was done using *regplot* from by Seaborn 0.11.0 [284] , Python3. The confidence interval was set to 0.68 which falls within 1 standard deviation. All statistical tests were performed in SciPY 1.5.2 [277] . Viscosity between species and within a species across A/P regions were compared using the t-test.

3.3 Results

3.3.1 Dynamics of anaphase spindle pulling varies independently of oscillations across species

Previous work has shown spindle movement, including oscillations, is mainly driven by pulling forces acting on the spindle in *C. elegans* embryos [107, 152]. However, in some distant species, for instance in the yeast *S. pombe* [267] and during meiotic divisions in some cell types, for instance some oocyte's meiotic divisions, spindle positioning is independent of microtubule-based pulling forces from the cortex [5]. We first asked whether the absence of spindle oscillations in some nematode species reflects a mechanism that is independent of pulling forces. Mechanical forces acting on the spindle can be revealed by laser ablation of the central spindle at the onset of mitosis. Following spindle cutting, the centrosomes recoil towards the cell pole if they are initially pulled [107]. This is because the central spindle connects the poles and holds the balance, much like a stretched rubber-band. In species in which the mitotic spindle elongates by inside-out pushing forces, spindle severing leads to the collapse of the centrosomes at the center of the spindle [139, 266].

We used a pulsed UV laser to sever the spindle at the onset of anaphase in all the 6 species Figure.3.2(A) and analyzed the recoil trajectories of the anterior and the

posterior centrosomes after the cut. The centrosomes of most of the species appear to recoil towards the cell pole Figure.3.2 (B). The movement of *D. sp. 1* centrosomes was very limited Figure.3.2 (B) with the characteristic recoil time being almost as long as the time of acquisition (27 s) with $\tau = 35.75$ and 25 seconds, for the anterior and posterior pole, respectively. We found a velocity of ~ 0.02 and $0.11 \mu\text{m/s}$ for the anterior and posterior pole, respectively. This is suggestive of a near absence of recoil in *D. sp. 1*. Spindle poles of *C. elegans* and *C. remanei* on the other hand, recoil faster with half times of ~ 9 s and ~ 5 s for anterior and posterior poles respectively. Compared to *C. elegans*, the A/P average recoil time of *C. monodelphis* is longer but *P. pacificus* and *O. tipulae* are statistically comparable Figure.3.2 (C). The recoil velocity of *C. elegans* centrosomes after cutting is $\sim 1.17 \mu\text{m/s}$ (anterior) and $\sim 1.82 \mu\text{m/s}$ (posterior), a difference of $\sim 36\%$ (Figure. 3.2(D)). This asymmetry in anterior and posterior pole recoil velocity is comparable to previous reports [107], and the ratio is identical in *C. remanei* centrosomes. Interestingly, we found a lack of significant A/P asymmetry in recoil half-times and velocities for *C. monodelphis*, *P. pacificus* and *O. tipulae*, despite their asymmetric cell division. These values are summarized in table 3.1 below.

Species	Recoil velocity (mm/s)		Recoil time, t (s)	
	Anterior	Posterior	Anterior	Posterior
<i>C. elegans</i> ($n=13$)	1.17 ± 0.50	1.82 ± 0.44	9.43 ± 4.60	5.18 ± 1.62
<i>C. remanei</i> ($n=14$)	1.48 ± 0.32	2.34 ± 0.43	8.84 ± 4.61	4.88 ± 1.72
<i>C. monodelphis</i> ($n=7$)	0.44 ± 0.14	0.34 ± 0.13	24.00 ± 10.98	26.47 ± 12.63
<i>D.sp.1</i> ($n=1$)	0.02 ± 0.00	0.11 ± 0.00	35.76 ± 0.00	24.99 ± 0.00
<i>P. pacificus</i> ($n=12$)	0.80 ± 0.32	1.03 ± 0.54	11.63 ± 8.25	7.89 ± 3.89
<i>O.tipulae</i> ($n=12$)	0.84 ± 0.35	0.78 ± 0.41	8.38 ± 7.33	15.43 ± 16.03

Table 3.1: Calculated cell length , cell cycle duration and centrosome size for the studied species.

Thus we find recoil trajectories of centrosomes after laser ablation cut vary between species in terms of the recoil velocity, half-time of recoil, final position of the centrosome

and anterior and posterior asymmetry Figure.3.2 (B-C). For all species that do not display oscillations, the half time of recoil was higher than those of *C. elegans*, whereas the initial velocity was systematically lower. While these results suggest the forces acting on the spindles could explain the differences between anterior and posterior centrosomes for some species, we need to estimate the forces. To do this from the recoil experiment requires our previous estimate of effective cytoplasmic viscosity that opposes the motion.

3.3.2 The net pulling forces and elasticity forces acting on centrosomes also vary between closely related nematode species

While spindle laser ablation experiments have been used in the past to estimate relative rates of pulling by forces acting on the astral MTs, the quantification of absolute forces requires a mechanical model of the spindles, motors and the cytoplasm. In recent work the absolute forces centering *C. elegans* metaphase spindles of 1-cell stage embryos has been estimated by measuring relaxation after mechanical perturbation and fitting to a model with both elastic and viscous components, i.e. a viscoelastic model [93]. Indeed, multiple studies have concluded that the mechanical properties of the cytoplasm are best explained by viscoelasticity [17, 76]. The Kelvin-Voigt (KV) model is used to quantify the spring force (F) that acts on laser ablated centrosomes pulling them backwards and also the viscous drag (γ) due to the presence of cytoplasm, crowding by granules and the elasticity (k) of the half-spindle and actin meshwork Figure.3.3(A). All species tested show a recoil of centrosomes on ablation, except for *D. sp. 1 JU359*, which had a negligible recoil post laser ablation. Nonetheless we can use the same physical model to characterize this recoil across species. Based on the KV-model the

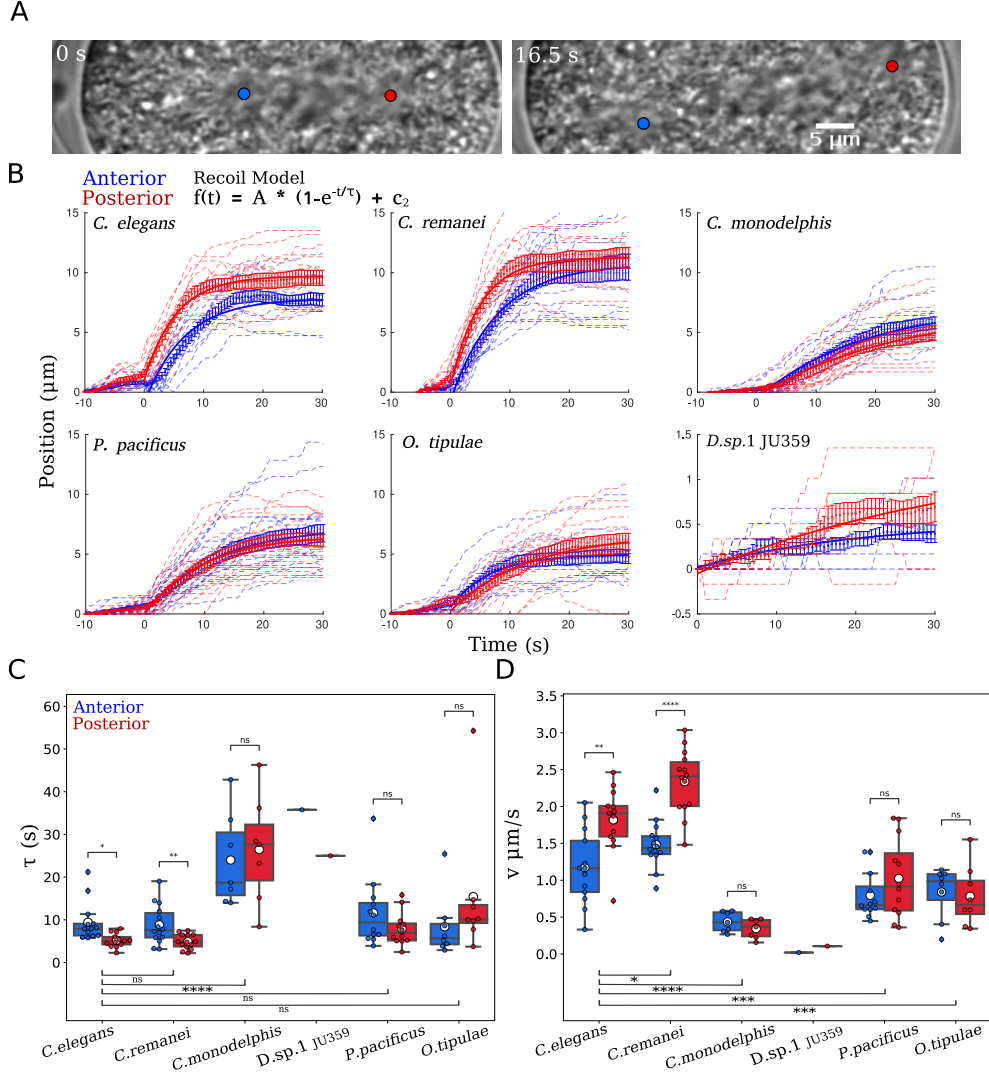


Figure 3.2: Centrosome trajectories post laser ablation characterized by recoil velocity and decay constant. (A) Representative images of a *C. elegans* embryo before (0 s) and 16.5 s after laser ablation of the spindle, with the anterior (blue circle) and posterior (red circle) centrosomes marked. (B) Recoil trajectories of the anterior (blue) and posterior (red) centrosomes after laser ablation ($t = 0$ s) for different species were averaged (filled circles) with error bars indicating SEM. The data were fitted to the recoil model (bold line) given by Equation 8. The individual profiles are also plotted (thin lines). (C, D) The fit parameter distributions for each species from the anterior (blue) and posterior (red) trajectories are plotted as box plots (τ : left, v : right) with mean (white circle), median (horizontal line), and first and second quartile (box) indicated. The differences between anterior and posterior parameters within a species were compared using a paired t test, while interspecies comparisons between means for each species were made using an independent t test. The asterisks indicate $\star p < 0.05$, $\star\star p < 0.01$, $\star\star\star p < 0.001$, and $\star\star\star\star p < 0.0001$ while ns = not significant.

position of the centrosome as a function of time, $p(t)$ is fit to the following equation:

$$p(t) = \frac{F}{k}(1 - \exp^{-\frac{kt}{g}}) + c1 \quad (3.3)$$

where F is the effective force driving the movement of the centrosome as it recoils after cutting, k is the elasticity of the medium, g is the drag coefficient and $c1$ is the position at time $t = 0$. Assuming the centrosome and MT aster can be treated as a spherical object, the Stokes drag coefficient can be estimated from the expression:

$$g = 6\pi\eta R \quad (3.4)$$

Here, η is the viscosity of the medium and R represents the centrosome aster radius. Centrosome sizes are estimated based on their smooth disk-like appearance that excludes the cytoplasmic granules due to the high density of astral microtubules around the pericentriolar material, as seen in DIC microscopy images. The radius of this disk serves as a measure for centrosome size for each species. This simplification is necessary since a centrosome is a complex structure with many microtubules extending from it, preventing a precise measurement of a radius. Also implicit in our approach is the assumption that centrosome size remains constant during anaphase. The effective viscosity η_{eff} for each species that combines the effect of granule crowding, by using quemada correction [223], as discussed in the previous chapter is used to estimate the viscous drag acting on the centrosome.

In the Kelvin-Voigt model, the centrosome position is governed by a pulling force F , here exerted on astral microtubules by cortical force generators proportional to the deformation from resting length and the elasticity k that results from a combined effect of cytoplasmic components, actin meshwork and astral MTs and the viscous drag g , that measures the opposition to the motion of the aster, Figure.3.3 (A). Given the free parameters, a good fit to the recoil trajectories requires constraints. We use the

effective cytoplasmic viscosity η_{eff} , as discussed above, to constrain the drag coefficient g . Substituting these values into Equation 3.3 we reduced the free parameters of the fit to only two: the force F and rigidity k , which were estimated from fits to the data from multiple recoil experiments. The results of this fitting are shown in Figure.3.3 (C) and summarized in table 3.2.

Species	Force (pN)		K (pN/ μ m)	
	Anterior	Posterior	Anterior	Posterior
<i>C. elegans</i> ($n=14$)	42.21 $\pm$ 15.43	101.67 $\pm$ 35.24	4.50 $\pm$ 1.45	11.38 $\pm$ 5.58
<i>C. remanei</i> ($n=14$)	35.61 $\pm$ 13.90	56.15 $\pm$ 15.31	3.37 $\pm$ 2.15	5.12 $\pm$ 1.97
<i>C. monodelphis</i> ($n=14$)	46.62 $\pm$ 23.62	42.55 $\pm$ 27.96	5.36 $\pm$ 4.77	4.81 $\pm$ 5.81
<i>D. sp. 1. JU359</i> ($n=2$)	2.70 $\pm$ 1.32	4.35 $\pm$ 0.88	2.87 $\pm$ 3.70	2.85 $\pm$ 3.03
<i>P. pacificus</i> ($n=16$)	26.68 $\pm$ 13.60	24.97 $\pm$ 17.54	3.75 $\pm$ 2.52	3.68 $\pm$ 2.13
<i>O. tipulae</i> ($n=10$)	27.82 $\pm$ 17.28	27.74 $\pm$ 17.05	5.54 $\pm$ 3.82	3.82 $\pm$ 3.26

Table 3.2: **Fit parameters from recoil data fitting.** Individual trajectories of centrosome recoil data from spindle cutting were fitted to the Kelvin–Voigt model and fit parameters averaged for each species. Centro-some sizes were measured from DIC images of spindles based on the maximal diameter of the clear region of the spindle poles. Brackets indicate the number of trajectories analyzed (n). Only those fits with $R^2 \geq 0.80$ (0.50 for *D. sp.1 JU359*) were used for analysis.

Our results revealed the net pulling force acting on the anterior and posterior centrosome of *C. elegans* spindles to be 47 pN and 71 pN, respectively Figure.3.3(C). The A/P asymmetry in pulling forces acting on half spindles is found only in *C. elegans* and *C. remanei*, with A/P force asymmetry absent in all the remaining species. This result is in apparent contradiction with the fact that all species divide asymmetrically. However, in these species, as in *C. elegans*, the entire spindle is slightly shifted posteriorly at the onset of anaphase. We hypothesize that this initial asymmetry is sufficient to maintain the spindle in an asymmetric position, despite symmetric pulling forces during late anaphase. We find *C. elegans*, *C. monodelphis* and *P. pacificus* spindles experience comparable magnitudes of forces. However, compared to *C. elegans*, those in *C. remanei* and *O. tipulae* are lower, while *D. sp.1* appears to have the lowest values measured amongst all 6 species. Differences in pulling forces are very limited (~ 3 fold)

for most species, except for *D. sp. 1* which is ~ 10 fold smaller than *C. elegans*.

The rigidity (k) values range between ~ 1 to ~ 9 pN/ μm , consistent with values reported for *C. elegans* embryonic spindle elasticity associated with centering [93]. Although variations of the rigidity appear more constrained than pulling forces, they show a similar trend between species. Here, only *C. remanei* shows clear differences with *C. elegans*, and to a lesser extent *O. tipulae* Figure.3.3(C).

The differences in spindle mechanics between species in terms of pulling forces, elasticity and viscous drag suggests a global trend that might explain differences in unperturbed spindle behavior. To investigate this we proceed to correlate these variables in order to find possible correlation patterns.

3.3.3 Towards the definition of a parameter space for spindle positioning and oscillations

Overall, our results show substantial and independent variation from one another, even between closely related species. Interestingly we find the spindle dynamics measured by recoil velocity, v and time-constant τ after cutting correlate strongly with effective cytoplasmic viscosity (mod $r > 0.7$) Figure.3.4. We find little correlation between viscosity and the spindle pulling forces Figure.3.4(C) and elasticity Figure.3.4(D). To our surprise, cell cycle time shows a strong correlation with viscosity with $r > 0.8$, although more species need to be examined to confirm this trend. Cell length, which we would have expected to affect the recoil properties only weakly correlates with the recoil velocity v and time of recoil η .

In the crowded environment of the cell, it is not surprising that viscosity plays an important role in intracellular mobility. When we measure the average unperturbed spindle oscillation speed of intact spindles and correlate it with viscosity, we find species with low viscosity have faster oscillating spindles, while more viscous cytoplasm result

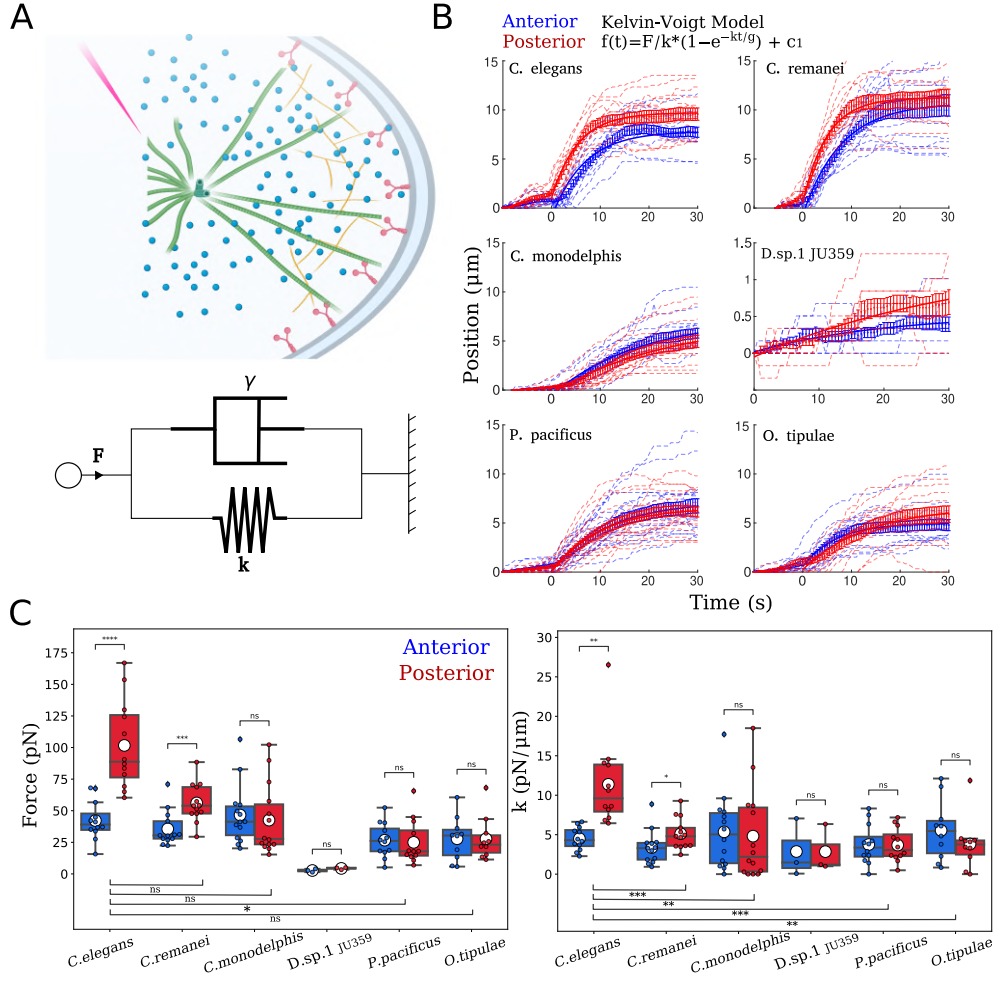


Figure 3.3: Pulling forces estimated from viscoelastic model fit to recoil data. (A) Schematic of laser ablation of the spindle midplane that results in recoil motion of the centrosome modeled by the Kelvin–Voigt viscoelastic model represented here by a spring with elasticity k and a dashpot with viscous drag coefficient γ , which acts to damp movement due to the pulling force F . (B) Recoil trajectories of the anterior (blue) and posterior (red) centrosomes after laser ablation for different species were averaged (filled circles). Error bars: SEM. The mean data were fitted to the Kelvin–Voigt model (bold line), Equation 1 (Materials and Methods). The faint dashed lines represent individual profiles of anterior (blue) and posterior (red) centrosomes. (C) The distributions of the fit values of force, F (Left), and elasticity, k (Right), for each species from the anterior (blue) and posterior (red) trajectories are plotted as box plots with mean (white circle), median (horizontal line), and first and second quartile (box) indicated. The differences between anterior and posterior within species were compared using a paired t test, while interspecies comparisons were done using an independent t test, similar to Fig 2. The asterisks indicate $\star p < 0.05$, $\star\star p < 0.01$, $\star\star\star p < 0.001$, and $\star\star\star\star p < 0.0001$ while ns = not significant.

in slower oscillations. When we attempt to distinguish species that show bonafide oscillations (Ce, Cr) from the rest (Pp, Ot, Dsp and Cm) that do not, by finding the mid-point of viscosity separating the species at the edge (Ce and Pp), we find a viscosity of ~ 0.6 Pa s emerges as a predicted threshold Figure.3.4(F). When compared to the predictions of the effect of viscosity on *C. elegans* spindle oscillation speed from a previously described computational model, we find sharp decline in oscillation speeds at a qualitatively comparable threshold of viscosity predicted by Kozlowski et al., 2007. In order to quantify the simulation outputs, the data was fit to 4-parameter sigmoid function (Equation 9, Materials and Methods). The resulting minimal oscillation speed from fitting is $0.15 \mu\text{m/s}$ (y-axis of Figure.3.4(F)). We use this as a quantitative measure for the loss of oscillations, that separates the experimental data of spindle oscillation speeds into 'oscillating' and 'non-oscillating'. By comparing the predictions from previous work [147], with comparative measurements using evolutionary diversification of properties, we overcome the difficulty of experimentally testing the effect of changing viscosity, since it is expected to have non-specific effects on cell physiology.

However, the viscosity threshold alone appears to be too small a change to be the sole determinant of loss of oscillations, and other parameters need to also be considered. The spindle pulling forces are well known to be the primary determinant of spindle oscillations and we therefore expect a trade-off between viscosity and pulling forces could explain sustained oscillations. This is supported by the case of *C. remanei* in which spindles oscillate whereas both forces and viscosity are reduced compared to *C. elegans* as discussed previously. On the other hand, *P. pacificus* experiences lower forces compared to *C. elegans* but only small differences in viscosity, which could explain the absence of oscillations. In order to account for a potential trade-off between the pulling forces and viscous drag that opposes it, and inspired by previous work by Juang et al. (2017) on the evolutionary scaling of mechanical stiffness of avian eggs by a dimensionless parameter combining mass, elasticity and egg size, we estimate a

ratio ω , that combines spindle pulling forces (F), viscosity (η) and embryo size (L). While this measure has dimensions length/time we find it may predict the propensity of spindles to oscillate [129]. We define it as:

$$\omega = \frac{F}{\eta \times L} \quad (3.5)$$

where η is the effective viscosity and L is the long-axis of the embryo. We find spindle oscillation speed plotted as a function of ω (with units $\mu\text{m/s}$) shows a clearer separation between the species and appears to predict the onset of oscillations with the midpoint between Pp and Cr ($\omega = 1.3$) serving as this threshold Figure.3.4(G). A low value of the ratio is suggestive of either low force or high viscosity, while a high value indicates either low viscosity or high force, both normalized by embryo size. Thus the measure ω appears to predict the onset of anaphase spindle oscillations in the species examined.

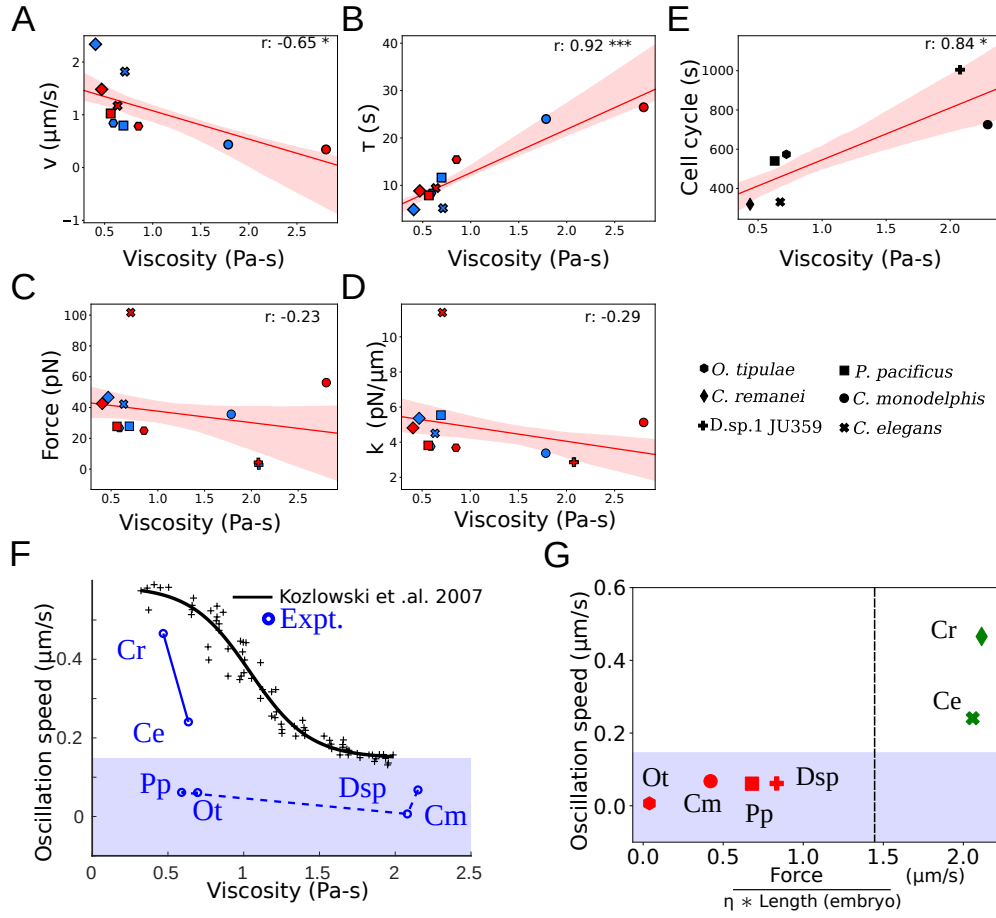


Figure 3.4: Effect on spindle mobility of divergence in cell mechanics. The covariation in all six species of (A) recoil velocity v , (B) decay constant τ , (C) spindle-pulling force (pN), (D) elasticity ($\text{pN}/\mu\text{m}$), and (E) cell cycle time with effective viscosity are plotted. Symbols: species, red line: regression fit, shaded area: one SD (confidence interval 0.68), r : Pearson's correlation coefficient, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, and **** $p < 0.0001$ while ns = not significant. (F, G) Spindle oscillation speeds (from Supplemental Figure S6) are plotted as a function of the effective viscosity, and ω , G , F , experimentally measured spindle oscillation speed: circles (o), simulations (Kozłowski et al., 2007): plus sign (+). Simulation data were fitted to a four-parameter sigmoid function. Fit parameters: maximal speed $a = 0.44 \mu\text{m/s}$, minimal speed $d = 0.15 \mu\text{m/s}$, viscosity at half-maximal oscillation speed $c = 1.04 \text{ Pa}\cdot\text{s}$, and steepness parameter $s = 0.2 \text{ Pa}\cdot\text{s}$. G, the oscillation speed of the spindle is plotted as a function of ω , the ratio of the pulling force to the product of viscosity and the length of the embryo ($\omega = F/(\eta L)$). The dashed line is the midpoint separating oscillating from non-oscillating species. F, G, shaded regions (purple): oscillation speeds less than the minimal speed $d = 0.15 \mu\text{m/s}$ from the fit to simulation data in F. Ce: *C. elegans*, Cr: *C. remanei*, Ot: *O. tipulae*, Pp: *P. pacificus*, Dsp: *Diploscapter sp. 1*, and Cm: *C. monodelphis*. Colors indicate non-oscillating (red) and oscillating (green) species.

3.4 Discussion

The intracellular aqueous environment of macromolecules and presence of cytoskeletal proteins has made the study of viscous and elastic properties of cells vital for our understanding of cellular mechanobiology. However, how much these properties can change without affecting the function they sustain is unclear. This is in part because experimental perturbations of these properties are challenging. An alternative approach is to measure the physical properties of cells from different species because they reveal the range of solutions retained over the course of evolution, and hence the range of possible changes.

In this part of the thesis we explored the reason why spindle motion varies between closely related species of nematodes. We had previously identified species for which spindle elongation and displacement at anaphase is accompanied by transverse oscillatory movements, similar to the situation found in *C. elegans*. In contrast, many nematode species do not undergo these oscillations, despite an identical off-center displacement of the spindle at anaphase. Comparing spindle movements is a difficult task because spindle elongation, displacement and oscillations occur simultaneously. Moreover, the spindle is composed of two centrosomes that are oscillating in an anti-phase manner and which are linked by a central spindle, whose mechanical properties are mainly unknown. Based on previous work demonstrating the role of motor-driven pulling forces (elastic) and the effect of cytoplasmic drag in damping this movement (viscous), we proceeded to examine the comparative viscoelastic properties of the spindle motion of multiple species. First we laser-ablated the central spindle and analyzed the recoil trajectories to estimate velocity and characteristic time constants. Using the drag to constraint a viscoelastic model of spindle recoil, we find spindle elasticity to vary by a factor of 10 and pulling forces to vary by a factor of maximum 3, except for *D. sp. 1* (~ 40 times lower than *P. pacificus*). Correlation of intact spindles as well as

recoil dynamics suggests a balance of forces acting on spindles determines the mobility. Cellular parameters such as centrosome size or cell length do not appear to correlate with spindle recoil parameters, while viscosity strongly correlates with both recoil parameters and intact spindle mobility. Inspired by studies to quantify egg stiffness by a single dimensionless number that revealed allometric invariance of avian egg mechanics we arrived at a measure of spindle stiffness ω , the ratio of force and viscosity scaled by embryo size [129]. This ratio predicts a threshold for the emergence of spindle oscillations above a critical value that could serve as a biophysical constraint on the evolution of spindle oscillations. Quantitative intracellular measurements of viscosity and spindle forces in multiple species will be required to test the validity of this prediction.

An important aspect of our study is the absolute quantification of drag, elasticity and forces acting on spindles. While measurement errors are inherent to any method, the error in size estimation of centrosomes is comparable across species, suggesting that while a more precise estimation of sizes will improve the accuracy of our estimates of spindle mechanics, it will not affect the relative trends. Indeed by considering a predictive ratio between the pulling force and opposing viscosity we possibly have a robust measure less sensitive to errors in measurement.

The results demonstrate that regardless of the presence of spindle transverse oscillations, all studied species are subjected to cortical pulling forces. Interestingly, although dynein, is a highly conserved protein, the protein responsible for its anchoring at the cortex, i.e. GPR- $\frac{1}{2}$ or LIN-5, are not found in the genomes of *P. pacificus*, *O. tipulae* or other members of the *Diploscapter* genus [59]. This raises the possibility of differences in the mechanical tuning of functional orthologs associated with cortical pulling across species. A thorough comparison and identification of key domains associated with cortical pulling are therefore essential for a molecular understanding of these differences. We found variations in the net pulling forces, by a factor of ~ 3 , for most species except *D. sp. 1*. Previous models have proposed that modulation of the force can be achieved

by changes in the number of motors, the individual force per motor, the attachment and detachment rate of motors, and microtubule dynamics, as seen in other systems [124, 257]. Establishing transgenic lines, in particular to follow live microtubules, in these different species is a necessary step towards more quantitative measurements of these parameters.

We also find the rigidity (spring elasticity) parameter differences between the species measured varies by a factor of ~ 10 , ranging between 0.92 and 8.6 $pN/\mu m$ (Table.3.2). In terms of the cellular components, this parameter can be understood to be a combined measure of cytoplasmic stiffness and microtubule components. The lack of a clear trend suggests that while the effective elasticity due to cytoplasmic inclusions and the density of microtubules and actin meshwork vary between closely related nematode species they appear to only weakly correlate with spindle mobility.

In order to dissect the molecular components that govern this effective elasticity parameter in future, a fluorescence based approach to compare these mesh-works could help understand why this mechanical parameter does not appear to play a role in spindle dynamics. This suggests that the forces of spindle pulling and drag are the primary determinants of the spindle behavior in these species. Embryonic division in all these species examined here is asymmetric. Yet our study finds only in *C. elegans* and *C. remanei* have measurably distinct A/P asymmetry in spindle pole retraction half times and velocity as well as pulling forces estimated. The asymmetry of the spindle pole displacement can arise either due to differences in forces driving of translocation or in the initial position, or both. We found that all species have a posteriorly shifted spindle at the onset of anaphase, similar to what had been described for *C. elegans* [107, 152]. In *C. elegans*, the initial shift of the spindle at metaphase and early anaphase is due to the asymmetry of pulling forces [152]. Our results thus suggest that an early asymmetry in the pulling forces is sufficient to maintain an asymmetric position and ultimately an asymmetric cell division in most species, regardless of the asymmetry of

pulling forces during the remainder of anaphase. Nevertheless, whether an imbalance of pulling forces is responsible for the initial shift in spindle positioning outside *C. elegans* remains to be demonstrated.

The evolutionary scaling of size, dynamics and elongation has been reported for inbred lines of *Caenorhabditis elegans* and was explained by a model of 1-1 stoichiometry of astral microtubules (MTs) and cortical force generator (FG) complexes [82]. Asymmetric spindle pulling in the model arises from a 20% greater number of FGs localized in the posterior of the cell. While this model addresses many of the limitations of previous physical models in explaining evolutionary diversification, the experimental data for reconciliation involves the use of inbred lines of the model. Therefore, by examining the nature of spindle dynamics and elongation in more than just *C. elegans*, we believe we see wider variation in measured properties, and identified potential constraints to the variability and the consequences of it in terms of cell size, viscosity and forces.

Overall, our results allow us to define which combination of parameters are now compatible with spindle transverse oscillations. Reduction of pulling forces can still lead to oscillations provided it is compensated by reduced cytoplasmic viscosity, as seen in *C. remanei*, per unit length of the embryo. However, changes in a single parameter, as seen in *P. pacificus* (reduced pulling forces) or in *C. monodelphis* (higher viscosity), critical for oscillations, do not suffice. We propose that a tradeoff between cortical pulling forces and cytoplasmic viscosity results in spindle oscillations, when the ratio of pulling forces to viscosity are high for a given embryo size. This provides a framework that could both be tested with more species of nematodes, as well as generalized to other cellular systems.

We also note that species that have high viscosity also have a longer cell cycle duration Figure.3.4(C) correlation coefficient ($R=0.87$). This result suggests that the only combinations that have been retained by natural selection are compensatory changes, where high viscosity is compensated by a slow down of the cell cycle or vice-

versa. Over the course of evolution, slow species could have afforded an increase in cytoplasmic viscosity because even though objects are slowed down by the viscous drag, they will have time to reach their final position. Conversely, a low viscous drag may have preconditioned the emergence of a fast cell cycle. Regardless of the orientation of changes, this interesting correlation raises the question of the selective pressure responsible for species-specific viscosity values.

Chapter 4

Automated cell cycle time

quantification of first mitotic division

in diverse nematode embryos

4.1 Introduction

The process of cell division is highly stereotypical both within and across nematode species. This provides a unique opportunity to investigate the divergence of cellular traits over evolutionary time scales. Comparative studies on spindle dynamics have revealed universal scaling relationships between embryo size and spindle dynamics [80], while variations in division asymmetry and cellular viscosity have been observed among different nematode species [136, 272]. However, there are no automated methods to quantitatively compare the dynamics of the spindle and other intracellular processes which can be applied to diverse nematode species. Such a method can be combined with existing tools for granule mobility, embryo and eggshell size quantification, and will aid in creating a database of intracellular parameters, which will be, invaluable in modeling how cellular traits diverge over evolutionary time scales.

Caenorhabditis elegans is a widely studied model system for embryogenesis due to its short generation time, eggshell transparency, and ease of genetic manipulation. The first mitotic division of the embryo, in particular, is extensively described. The division process starts with the entry of sperm into the oocyte, which establishes anterior-posterior polarity and triggers the completion of oocyte meiosis. Subsequently, the female pronuclei migrate towards the male pronuclei near the anterior of the embryo, followed by their fusion to form a nucleus-centrosome complex (NCC). The NCC moves to the center of the embryo and aligns itself along the anteroposterior axis, leading to the formation of a mitotic spindle that separates the chromosomes into two opposite poles. Finally, a cytokinetic furrow forms, dividing the cytoplasm and giving rise to two asymmetric daughter cells.

Using DIC microscopy, the sub-cellular morphology during this process can be clearly visualized. The two pronuclei, NCC and the mitotic spindle form a uniform intensity region surrounded by granular cytoplasm. This clear distinction between the

organelles and the cytoplasm is exploited for the identification of phenotype mutants in functional genomic screens [104, 219, 249, 304]. Additionally, automated analysis has also been developed to study embryo shape [20], pronuclear migration [141], spindle dynamics [81], and granule mobility [39] in DIC time series of developing embryos. However, these tools are limited in scope or are not directly applicable for diverse embryos that show morphological differences.

Recently, convolutional neural networks (CNNs) have been applied in various areas of biology, such as cancer classification [194], cell cycle analysis [196, 226], and cell phenotyping [69, 184]. CNNs excel at supervised learning, where they can automatically annotate input images into predefined labels used during network training. They are capable of recognizing robust features corresponding to a particular class [196] and can also learn the latent relationships between classes [75]. In particular, the work of Eulenberg et al. (2017) has shown that CNNs can reconstruct cell cycle and disease progression, which can be visualized using dimensionality reduction techniques such as t-SNE [75, 274]. Another example is *DeepCycle*, which demonstrates the ability of CNNs to quantify cell stage progression in fluorescent-tagged cells [227]. However, deep learning tools cannot be directly applied to novel datasets without considerable re-training and optimization.

The application of CNNs to nematodes have been limited to the study of *C. elegans* adult morphology [9, 11] and embryogenesis [33]. There is a lack of automated tools that can be applied to the first mitotic division across nematode embryos. We propose that a well-trained network capable of recognizing distinct stages of the first mitotic division will be capable of identifying subtle differences in how intracellular dynamics differ among related species. It can be used to accurately measure the duration of each phase, which is otherwise a time-consuming and laborious task. Additionally, an automated annotation tool can be not prone to human error which can occur when working with large datasets.

In this study, we trained several convolutional neural networks (CNNs) to accurately classify distinct stages of the first embryonic division across various *Caenorhabditis* species. Additionally, we optimized a shallow CNN, named EvoCellNet, which solely relies on morphological features for classification. The average F1-score for classification of all species is $\sim 95\%$. The CNN features, visualized using Grad-CAM [238], correspond to stage-dependent structures that were identified across the *Caenorhabditis* genera. The method can be applied to label-free DIC microscopy data with no tuning parameters.

4.2 Materials and Methods

4.2.1 Manual annotation of cell stages

All time-series data used in this study were obtained from a published database containing recordings of 42 nematode species undergoing first mitotic division [272]. Since the database focuses on comparing spindle mobility traits, most recordings start after the spindle has already formed. To avoid over-representation of any one stage during our annotation, we filtered those time series that start before pronuclear congression. Also, we limited our selection to only *Caenorhabditis* species, which form a major portion (84%) of the database. In total, we selected 213 recordings of 23 unique *Caenorhabditis* species, which initiate before pronuclear congression. We then manually annotated the selected time series based on visible sub-cellular morphology. For each time series, we marked three sub-cellular events as checkpoints: meeting of the two pronuclei, initiation of spindle elongation, and invagination of the embryo cortex. These checkpoints were used to segregate the time-series into four stages: pronuclear migration (PM), centration and rotation (CR), spindle elongation and displacement (SD), and cytokinesis (Cy). The nomenclature of these stages is motivated by similar use in previous studies on nematode embryo development [103, 104, 112]. Annotation of these stages

was done in ImageJ [49]. Out of the selected 213 recordings, 44 recordings representing 21 unique *Caenorhabditis* species were kept for testing. The remaining dataset was split into four folds for training and validation using stratified k -fold cross validation split based on species identity. The distribution of individual frames into four classes is shown in Figure.C1(B).

Data augmentation and image pre-processing

Data augmentation is an effective technique to reduce over fitting during network training which improves model generalization [241]. We applied random image transformations to the input training batch such as vertical and horizontal flips (*probability* = 0.6), affine and image rotation in range of (0,10) and (0,60) degrees respectively as shown in Figure.C1(A). All images were resized to (256, 256 px) in height and width and intensity normalized with mean and standard deviation of $\mu = 0.48$ and $\sigma = 0.15$, respectively. The mean and standard deviation were calculated from the entire dataset.

4.2.2 EvoCellNet architecture

The CNN architecture consists of 6 convolutional layers followed by a one-dimensional convolutional layer with four output filters. The layers have a 3x3 kernel size except for the first layer with a 5x5 kernel size. The network architecture was optimized by experimentation. The input to the layers is normalized using batch normalization and the output from each layer is passed through a ReLU activation[2]. The input to the first layer is a 2-D single channel tensor of 256 x 256 pixels (height x width). The network uses skip connections and layer concatenation. Max pooling is applied to obtain the same size of all feature maps before concatenation. Dropout ($p = 0.3$) was added after each layer except for the first 2 layers. The number of filters in each layer and the network layout is shown in Figure.4.2(A). The output from the last layer is

passed through global average pooling layer producing a ($N \times C \times 1 \times 1$) tensor, where N is batch size and C is the number of classes. The network has 787,636 parameters.

4.2.3 Training parameters

All networks were optimized using stochastic gradient descent algorithm [22] with a learning rate (α) of 1e-2 for 50 epochs and a mini-batch size of 64. The categorical cross entropy loss was minimized during optimization. The network training was done on a V100 NVIDIA (16GB) card. Python version 3.11.5 was used, with the *PyTorch* framework version 2.1.0+cu118 utilized for deep learning work [213]. A comparison of inference time of the trained networks is listed in TableC.4. Code implementation and network weights are shared on github (EvoCellNet).

4.2.4 Evaluation metrics

Descriptive metrics such as accuracy, precision and F1-score were used to evaluate network performance. Accuracy is calculated as:

$$Accuracy = (TP + TN) / (TP + FP + TN + FN), \quad (4.1)$$

where TP, TN, FP, FN are true positive, true negative, false positive and false negative count respectively. Precision quantifies the proportion of positives that are true from total positives informing us about mis-classification and is described by the expression:

$$Precision = \frac{TP}{TP + FP}, \quad (4.2)$$

The recall score in turn informs us about the numbers of images that were missed from being classified correctly and is given by:

$$Recall = \frac{TP}{TP + FN}, \quad (4.3)$$

A combined metric that is the harmonic mean of the recall and precision metrics as given by:

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall}, \quad (4.4)$$

4.2.5 UMAP representation of network activations

Uniform Manifold Approximation and Projection (UMAP) implementation in python was used for the cluster analysis [233]. The output feature vector from the trained CNNs was used as the input. Parameters used for UMAP were *n - neighbors*: 10, *seed*: 55, *min - dist*: 0.2, *n - components*: 2, *metric*: "euclidean". The input vector was scaled to range [0,1] using *StandardScaler* function provided by *sklearn*. The *StandardScaler* was fit to the features from the validation dataset for each fold, and was saved for future datasets. The UMAP features were rescaled again before plotting. The file used for this is *pseudoTime.py*

We partition the points in two-dimensional UMAP space into a specified number of sectors, determined by the *num_sectors* parameter. Utilizing the angular positions of each point relative to the center, computed with *np.arctan2*, we segregate the space into sectors of equal angular width. Next, we compute the mean coordinates of points within each sector and identify the predominant class label among them. The means are then connected in sequence starting with the pro-nuclear migration class and ending with the cytokinesis class. This overlay facilitates the analysis of spatial distribution and class composition within the UMAP embedding, akin to cell cycle trajectories obtainable with self-organization maps as used in DeepCycle [226].

4.2.6 Transition boundary annotation using SVM

Support vector machine (SVM) algorithm was utilized to determine the transition boundary between adjacent classes. The input to the SVM is a two-dimensional feature vector obtained from UMAP. The UMAP feature vector is derived from the prediction outputs of the neural network. The support vectors extracted from the SVM indicate potential cellular checkpoints. While support vectors typically lie on the maximum separating boundary between classes, non-separable clusters in linear SVM result in points falling within the SVM margins, leading to misclassification beyond the correct SVM boundary. The algorithm identifies these points as support vectors as well. To improve the accuracy of automated checkpoint annotation we apply a data filtering step where we select time frames that fall within the average time frame of the support vectors on both sides of the decision boundary. Then, we determine the final transition checkpoint by taking the average of the maximum time frame from the first class and the minimum time frame from the second class. The calculations and analyses were performed using *python ver 3.8.8*, specifically utilizing the SVM algorithm provided by the *sklearn* library version *0.24*.

4.2.7 Code availability

All scripts used for training, analysis and figure generation are shared here: [nematode-DL](#)

4.3 Results

4.3.1 Annotating Images of nematode embryo development for CNN classification

Nematode embryos are well suited for differential interference contrast (DIC) microscopy due to the transparent nature of the oocytes. The spindle dynamics can be clearly visualized that forms uniform intensity region surrounded by a granular cytoplasm. We obtained multiple time-series of species that are evolutionarily related to *C. elegans* that had previously been analyzed in terms of spindle dynamics [273]. The selected image time-series begin before fertilization where two distinct nuclei are visible- male and female pro-nuclei. All time series end with cytokinesis. The furrow ingression is clearly visible which leads to cell cytokinesis. Based on morphologically identifiable stages, each time series was manually annotated into four stages of first mitotic division. To be consistent, we followed the annotation schematic shown in Figure.4.1(A). A montage of 5 related *Caenorhabditis sp.* embryos namely *C. elegans*, *C. remanei*, *C. briggsae*, *C. angaria* and *C. portoensis* demonstrate qualitatively similar morphological changes as they progress through the stages of (i) pro-nuclear migration (PM), (ii) nuclear centration and rotation (CR), (iii) spindle displacement (SD) and (iv) cytokinesis (Cy) Figure.4.1(B).

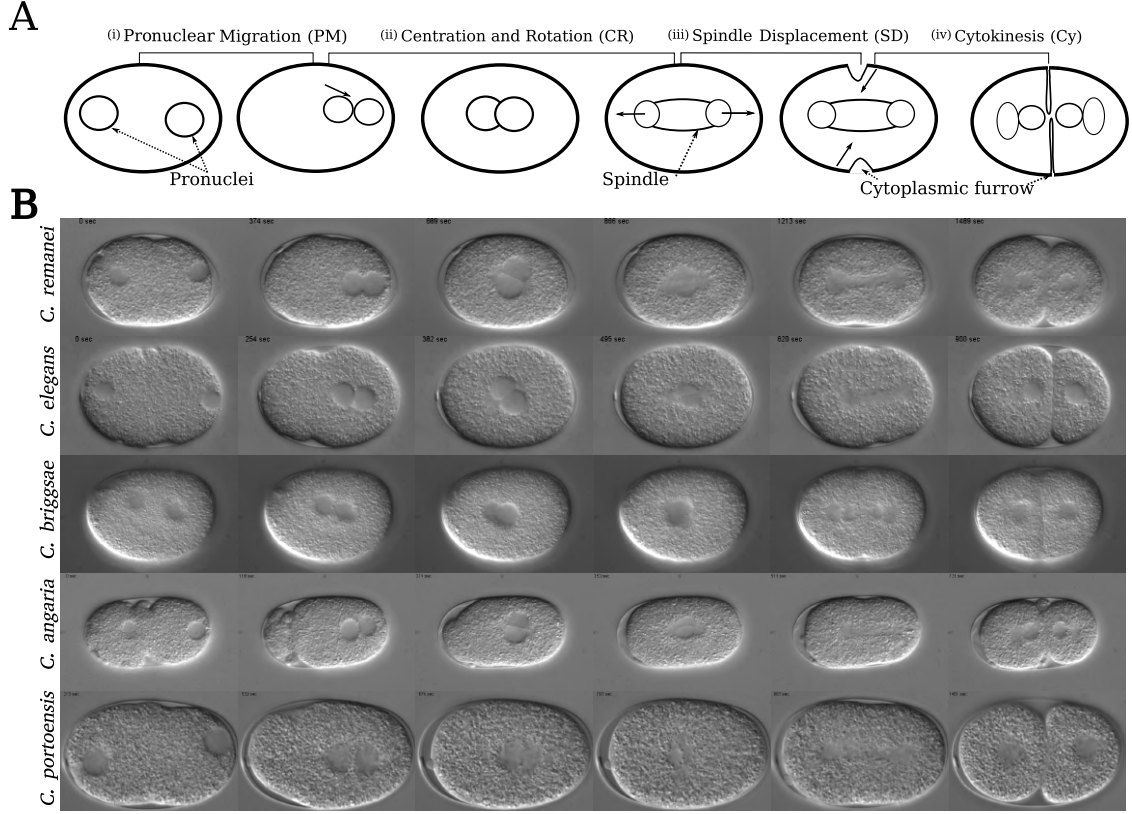


Figure 4.1: **Stereotypical patterns during the first embryonic division across diverse nematode species.** (A) A schematic of the first mitotic division in a typical nematode embryo is shown. We classify a given time-series of first embryonic division into four stages of (i) pronuclear migration (PM), (ii) centration and rotation (CR), (iii) spindle displacement (SD) and (iv) cytokinesis (Cy). Annotation of these stages is done by identifying sub-cellular events of pronuclear migration, spindle elongation and furrow indentation (indicated by bold arrows). (B) A montage of five *Caenorhabditis* species undergoing first embryonic division is shown. Species names is written at the start of every row. All DIC images taken from a published dataset [273]. The montage visualizes stereotypical morphology observed in a given stage across different species. These cytoplasmic features are formed by the two pronucleus and the mitotic spindle, which are visible as uniform intensity regions surrounded by a granular cytoplasm and a cell membrane.

Input data during training was augmented by using image transformations such as, affine transformations, image flipping in x- and y-axes and shearing Figure.C1(A). The training data of 169 time series was split by stratified K-fold cross-validation split strategy ($k=4$) based on species and cell stage identity, Figure.C1(C) Since the number

Network	Parameters ($\times 10^6$)	Convolutional layers	Linear layers
EvoCellNet	0.7	7	0
EfficientNet	11	237	1
ResNet	11	19	1
VggNet	134	13	3

Table 4.1: **Configuration of trained models.**

of frames in each stage for the original training dataset were different, with $PM = 43,821$, $CR = 97,715$, $SD = 53,197$ and $CY = 63,336$ and we oversampled the minority classes of the training split. All models were trained using the same hyperparameters for comparison, and model weights were saved based on a minimum validation loss criterion.

4.3.2 Accurate classification of distinct stages of first mitotic division across neamtode genera

Three CNN models from the *torchvision* library were trained for the image classification task in pytorch [213]. These are: ResNet [117], VggNet [242], and EfficientNet [262]. To these we added a shallow CNN that we named to as EvoCellNet, the architecture of this network is shown in Figure.4.2(A). We attempted to address whether network design based on previous developments could be optimized, given the trade-offs between complexity, accuracy and training time. Additionally we wanted to optimize the gradients to overlap with cytoplasmic features that are visible during first mitotic division. EvoCellNet network is the simplest in terms of number of parameters and numbers of convolutional layers as compared to the other models and is optimized to highlight cytoplasmic features observed during distinct stages. This is discussed in more details in the next section of results. A comparison of model configuration is shown in Table 4.1. All models were trained from scratch using the same hyperparameters, and model weights were saved based on the minimum validation loss criterion.

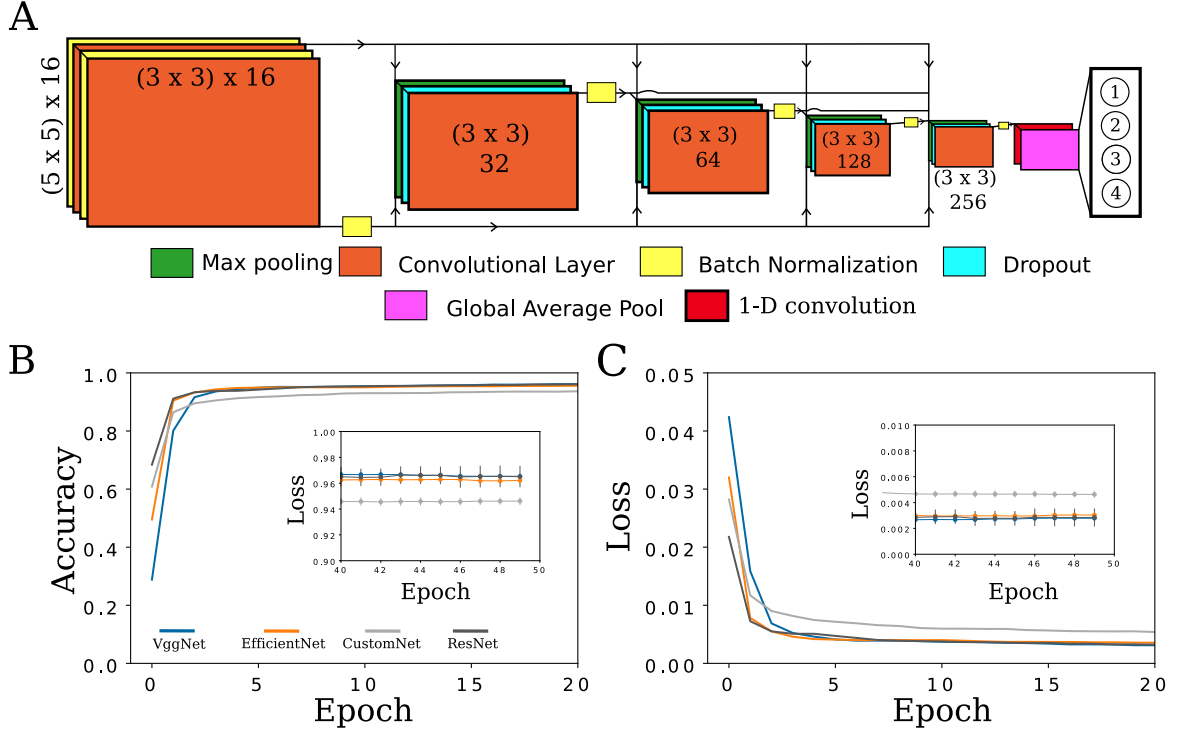


Figure 4.2: **CNN architecture and training metrics.** (A) The architecture of EvoCellNet is shown. The CNN consists of 6 convolutional blocks and each block consists of multiple color-coded functions: convolution (orange), max pooling (green), batch normalization (yellow), dropout (cyan), global average pool (magenta) and 1D convolution (red). The numeric values indicate filter size (height x width) and the number of filters in the respective layer. The outputs of the classifier (1-4) are fixed. (B) The accuracy (left) and loss (right) measure during network training are plotted with epoch number. These are averages of four folds of cross validation. Individual curves are shown in Figure.C2. The detailed description of these metrics is described in the *Methods* section. Multiple previously described CNNs- VggNet [242], EfficientNet [262] and ResNet [117] are compared to our network (EvoCellNet). Details are described in the *Methods* section.

	Pronuclear migration	Centration rotation	Spindle displacement	Cytokinesis
	Validation			
EvoCellNet				
Precision	0.97 ± 0.010	0.97 ± 0.004	0.88 ± 0.012	0.99 ± 0.003
Recall	0.99 ± 0.004	0.94 ± 0.013	0.96 ± 0.007	0.96 ± 0.009
ResNet				
Precision	0.99 ± 0.005	0.97 ± 0.008	0.94 ± 0.029	0.98 ± 0.007
Recall	0.98 ± 0.004	0.97 ± 0.010	0.93 ± 0.017	0.98 ± 0.015
VggNet				
Precision	0.98 ± 0.007	0.97 ± 0.013	0.94 ± 0.012	0.98 ± 0.006
Recall	0.99 ± 0.005	0.97 ± 0.008	0.93 ± 0.023	0.98 ± 0.010
EfficientNet				
Precision	0.98 ± 0.003	0.97 ± 0.004	0.93 ± 0.007	0.98 ± 0.004
Recall	0.98 ± 0.004	0.96 ± 0.007	0.93 ± 0.003	0.98 ± 0.005

Table 4.2: **Network performance for each stage for validation dataset.** The average precision and recall metrics with standard deviation for individual stages for all the four models used in this study. Then mean and standard deviation are calculated across the four folds of the k-fold split for each model.

Performance on validation dataset

The accuracy and loss curves during training, for all networks, converge by 5 epochs Figure.4.2(B-C). Each curve in the plot represents the average of the four cross-validation. The shallow EvoCellNet exhibits higher convergence loss compared to the deeper network architectures. The individual curves for each fold display differences in the smoothness of the validation accuracy and loss and are shown in Figure.C2. However, the final convergence values are similar. All the trained networks performed with high accuracy on the validation dataset, with EvoCellNet achieving an accuracy of 95%, compared to 97% for VggNet, 96% for EfficientNet, and 97% for ResNet. A comparison of precision, recall and F1 score for training, validation and test dataset is shown in Figure.C3(A). Despite the simpler architecture of EvoCellNet and significantly fewer parameters compared to ResNet ($\sim 6\%$ of the parameters), it achieved only 2% lower accuracy during validation. This indicates that fewer parameters are sufficient to capture the differences between the distinct classes.

The confusion matrix for the validation dataset from the first fold of cross-validation for all four models is shown in Figure.4.3. The results demonstrate that all networks ex-

hibited lower performance in identifying the Spindle displacement (SD) stage compared to the other stages. Specifically, EvoCellNet displayed a 9% lower precision, while the larger networks demonstrated approximately 3% lower precision for the "Spindle displacement" stage. This drop in accuracy is also reflected in the confusion matrix, where the misclassification values (off-diagonal) are higher for the CR-SD stage as shown in Figure.4.3. These results were expected since the transition from the CR stage to the SD stage is marked by the beginning of spindle elongation, making it challenging to precisely identify the exact point of elongation in a frame-wise classification scheme. In contrast, the other two checkpoints, namely the meeting of the two pronuclei and invagination of the cortex, are much easier to identify and can be distinguished at a single time point.

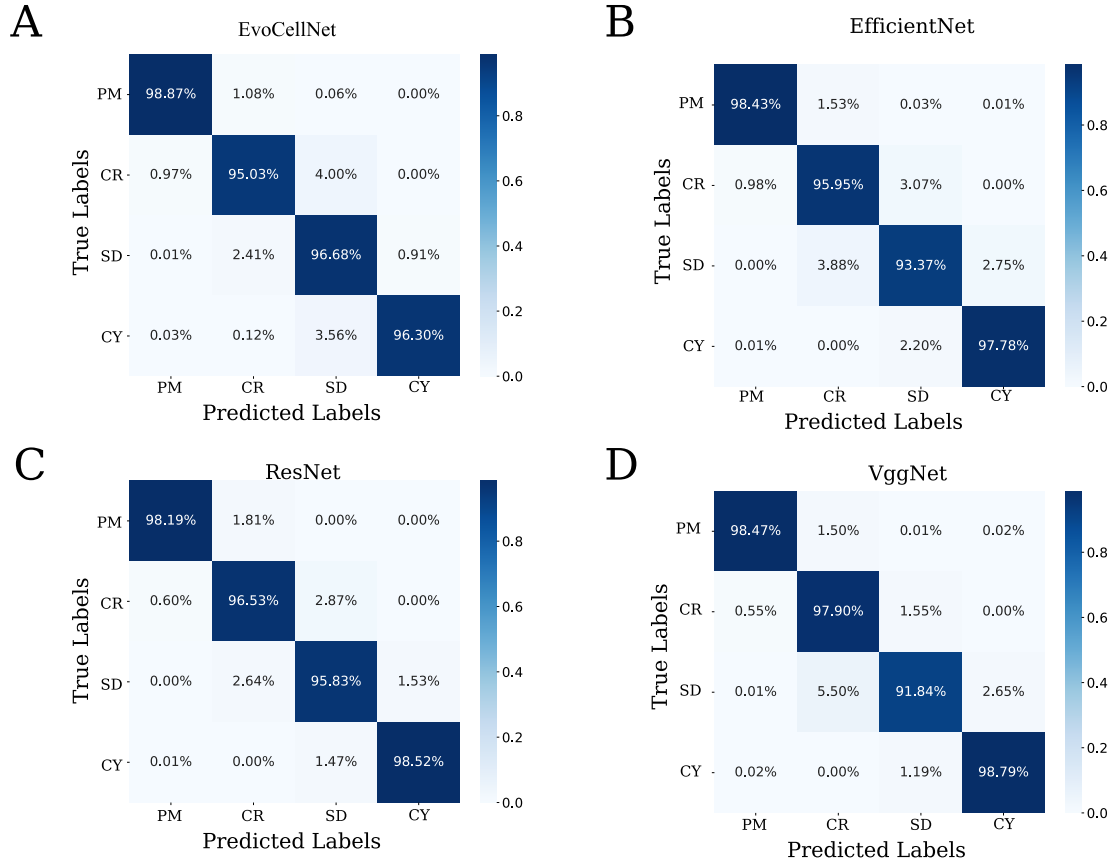


Figure 4.3: **Confusion matrix for the validation set.** Confusion matrix with true labels on the y-axis and predicted labels on the x-axis are shown for each model (A) EvoCellNet, (B) EfficientNet, (C) ResNet, (D) VggNet. The matrices are from the first fold of cross validation.

In summary, our results demonstrate the accurate classification performance of the models in identifying the stages of mitotic division. Although the SR stage presented a slightly greater challenge compared to other stages, the overall precision and recall metrics indicate the robustness of the models.

4.3.3 Uniform manifold approximation and projection (UMAP) of cell stage progression

It is evident from the validation metrics that most classification errors are concentrated at adjacent classes, which is consistent for all trained networks. This outcome was

expected, considering the strict time frame used to define class boundaries. However, there are difference among the trained models in their ability to confine these errors to the adjacent classes (Figure.4.3). To investigate these difference we utilized the uniform manifold approximation and projection (UMAP) technique [180] to represent the four-feature prediction of every input from the validation dataset vector in a two-dimensional space. Figure.4.4(A) illustrates this 2D space for the validation dataset (*fold* 1) for all four models. The color-coded 2D cloud represents the predicted classes and instances of misclassifications are highlighted with black outlines. As anticipated, the cloud system forms distinct clusters for each class, with the majority of misclassifications occurring in the transition region between adjacent clusters, and relatively few misclassifications occurring between non-adjacent classes.

We observed that our shallow model exhibited poorer segregation of classes, as compared to the deeper networks which showed better separation of clusters. Visually, VggNet performs the best in limiting errors to the boundaries of adjacent clusters. The higher error rates between the CR and SD stages for all models are also apparent in the UMAP space. Although UMAP is a stochastic algorithm, the observed differences were consistent and reproducible with a predefined seed. Overall, these results demonstrate that trained networks capture the latent relationship between adjacent classes without any prior information. Among the trained networks, VggNet exhibits the best segregation, in terms of limiting the errors to cluster boundaries of the classes in the UMAP space.

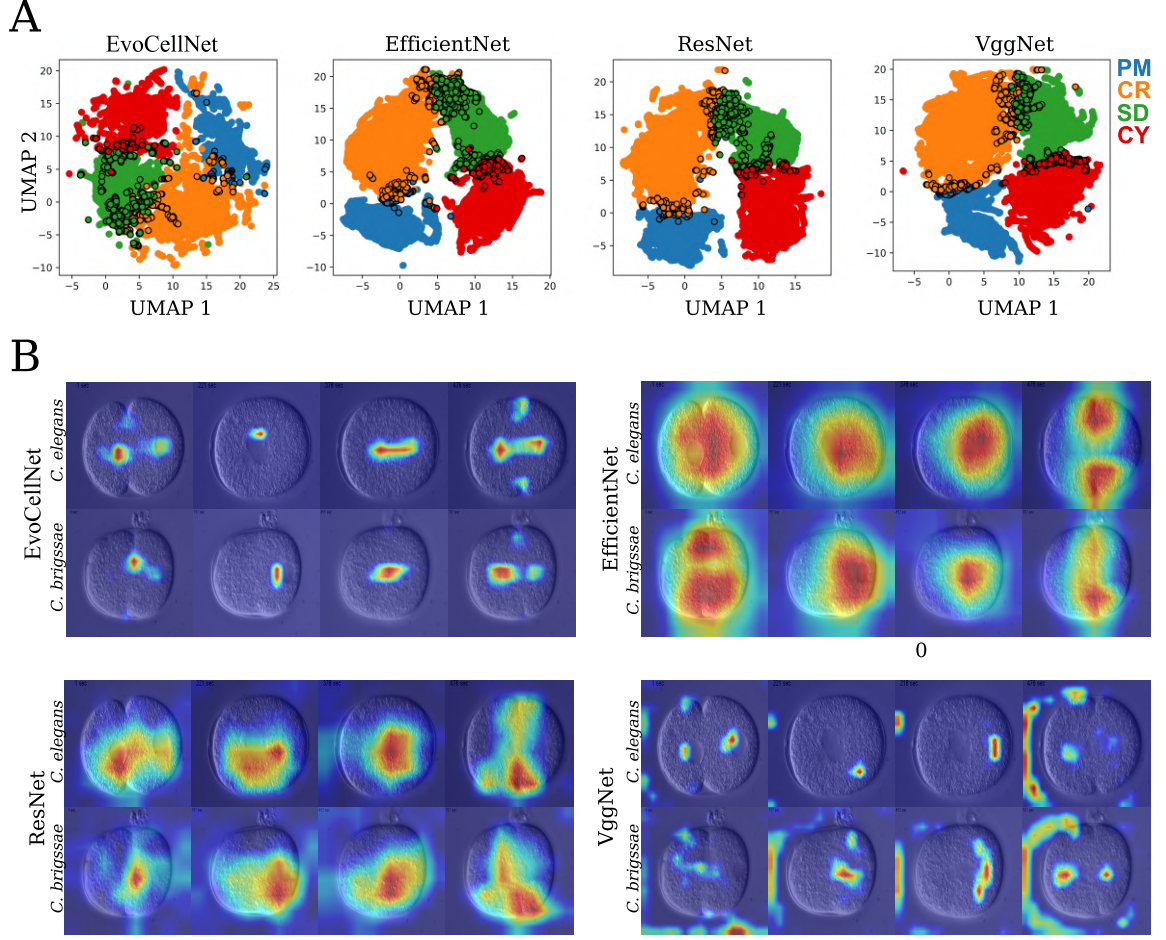


Figure 4.4: **UMAP projection of the validation dataset and gradient visualization using Grad-CAM.** (A) A 2D UMAP representation of the 4 value feature vector obtained from the last convolutional layer before applying softmax (*Methods*) is shown for all trained CNNs. The data points are color coded according to their predicted classes. Incorrect classifications are shown as solid circles with black boundaries. (C) A montage of 2 representative species with one randomly selected frame from each stage is shown. The images are overlaid with gradient activation values obtained using grad-CAM [238]. Activation values were rescaled to 0 and 255.

4.3.4 Network predictions are based on robust morphological features

During CNN training, convolutional filters are optimized to respond to image features that strongly correspond to a particular class. We use Grad-CAM [238] to visualize these features from the last convolutional layer of all models for sample embryos of *C. elegans* and *C. briggsae*. The results are shown in Figure.4.4(B). The filter activations were either directly on top of the visible cytoplasmic structures (e.g., pronuclei, spindle region) or concentrated around their boundaries. In both cases, the feature maps displayed stage-dependent patterns that were formed by sub-cellular organelles in each stage. The filters also responded strongly to the transition region between uniform and granular intensity profiles, which correspond to the boundaries of the embryo and cytoplasmic structures such as the pronucleus, spindle region, and cytoplasmic furrow. These feature maps validate that the trained models are detecting stage-dependent textures to differentiate between classes, as a manual approach would also do.

A complete time series movie with overlaid gradient maps for all the best-performing models is shown in Figure.SVC1. The activations of filters from the last layer varied among the four models, as depicted in Figure.4.4 (B). These variations can be attributed to the differences in the depth of the trained network architectures. Since Grad-CAM uses the feature maps from a specific convolutional layer, increasing the number of convolutional layers decreases the spatial resolution of the feature maps due to the successive pooling operations. Therefore with deeper networks the gradients tends to lose resolution and granularity. In line with this understanding, the final predictions in EvoCellNet focuses on local and fine grained features and gradients strongly overlap with sub-cellular organelles.

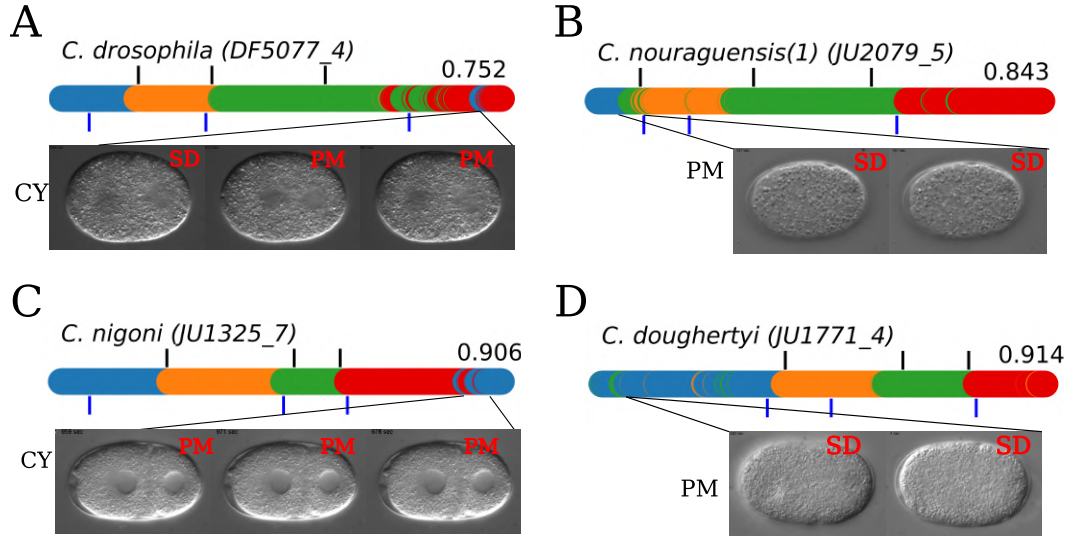


Figure 4.5: **Network performance drops in the absence of stage dependant cytoplasmic features..** Four time-series where the CNN classifier performed with less than 92 % accuracy are shown. Unique identifier from the original database is written in brackets. Each horizontal colored plot is a frame wise classification output from the ensemble predictor. The colors indicate specific stages, similar to Figure.4.2. Black vertical lines mark the manually labelled transition boundaries and blue lines mark the automatically calculated boundaries as explained in *Methods*. Red text inside the image frame indicates CNN classification label. Black text indicates correct label for each row. In the case of (A) *C. drosophila* and (C) *C. nigoni* the CNN incorrectly classifies cytoplasmic stage (CY) as pro-nuclear migration stage (PM). This is due to poor visibility of the cytoplasmic furrow. Also, the images contain features that resemble pronuclear migration stage. In the case of (B) *C. nouraguensis* and (D) *C. doughertyi* there are no visible cytoplasmic structures in the plane of acquisition. This results in miss-classification errors.

4.3.5 Automated cell stage checkpoint recognition across diverse *Caenorhabditis* embryos

The performance of the four CNN models was evaluated on a test dataset consisting of 44 time-series from 23 species of *Caenorhabditis* embryos. Predictions for each model were obtained by averaging the outputs of the best models from the four folds of cross-validation. Ensemble performance of each model for the four classes are summarized in TableC.3. The trend in errors is similar to that of the validation dataset. Figure.C3 (C) presents a comparison of the precision, recall, and F1 score for each model. Among these models, EfficientNet demonstrated the best performance with a classification accuracy of approximately 95%. The confusion matrix for the test dataset is shown in Figure.C4. Upon careful inspection, we identified four time series where all models exhibited poor performance. A closer examination revealed that these time-series lacked visible cytoplasmic structures for a significant portion of the time series, potentially due to changes in the plane of acquisition. These problematic time-series are shown in Figure.4.5. Excluding these time-series resulted in an increase of 2% in accuracy for all networks.

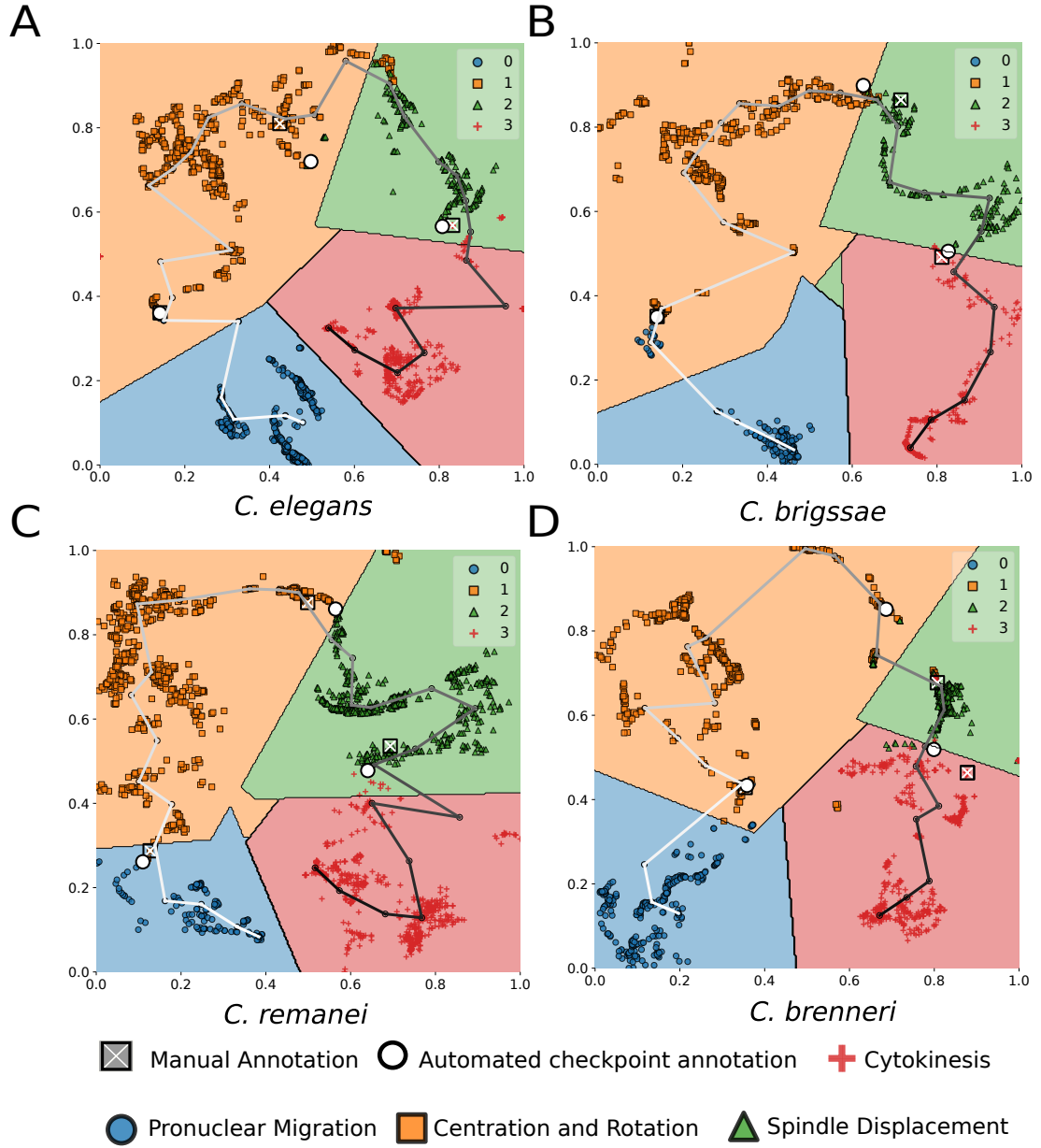


Figure 4.6: **Cell stage boundaries combined with radial density trace in VGGNet classified UMAP space.** (A-D) Normalized UMAP representations in (x-axis) UMAP1 and (y-axis) UMAP2 space from VGGNet classification were split into sectors corresponding to PM, CR, SD and Cy. Boundaries were determined using SVMs (as described in the Methods section) and overlaid with radial density traces, averaged every 12° . The radial traces were connected geometrically with 0° corresponding to the first frame and the direction set by the transition boundary between pronuclear migration (PM) and cytokinesis (Cy) for (A) *C. elegans*, (B) *C. briggsae*, (C) *C. remanei*, and (D) *C. brenneri*. $num_sectors$ used for this overlay is 30. Color gradient: start: white, end: black, x: ground truth of the annotated checkpoint, o: checkpoint from SVM.

With an accurate classification output, we can use the UMAP representation to mark the positions of key cellular events, which were previously manually annotated. By using a linear SVM kernel on the 2D UMAP output, we can automatically identify the positions of the defined checkpoints between adjacent classes. In addition, the UMAP cluster reconstructs the cell stage progression from PM to Cy stage. To aid in visualizing this representation a radial cluster averaging is done starting from PM stage ending at Cy stage. A representative UMAP output with the identified boundaries is shown for *C. elegans*, *C. brigssae*, *C. remanei* and *C. brenneri* time series in Figure.4.6 (A)-(D). For comparison a similar plot for EvoCellNet, ResNet and EfficientNet for the *C. elegans* embryo is shown in C5. The UMAP representation was done using VggNet which had visually the best segregation of classes in our testing. For the complete test dataset the absolute time difference between the manually annotated time boundary and the automatically found separation boundary (ΔT) is shown in Table 4.3.) Overall, the SVM algorithm applied to the UMAP-derived feature vector offers a robust approach for identifying transition boundaries and visualizing the cell cycle trajectory.

Model	Checkpoint 1		Checkpoint 2		Checkpoint 3	
	Mean (s)	Std	Mean (s)	Std	Mean (s)	Std
EvoCellNet	13.2	20.6	29.6	32.1	18.5	21.1
EfficientNet	12.1	27	25.4	18.9	13	17
ResNet	12	18.3	26.4	22.4	13.8	17
VggNet	8.6	14.6	24.9	16.9	13.2	22.5

Table 4.3: **Absolute mean error in checkpoint prediction.** Average and standard deviations of time difference (ΔT) between manual annotation boundary and automatically identified boundary are shown for the four trained models.

4.4 Discussion

Automated recognition of microscopy images is an intense field of research. Recent applications of DL-based classifiers have shown robust recognition of image features that demonstrate the applicability of DL for the complex task of categorizing cellular images [69, 100, 196, 297].

In our work, we have trained multiple CNNs to accurately recognize distinct cell stages of the first mitotic division in diverse *Caenorhabditis* embryos. The main objective was to automate the annotation of cell stages and assess the robustness of decision-making by the networks while making predictions. The ability to leverage cytoplasmic features for cell stage classification using DL has been demonstrated in multiple cell types [128, 196, 226]. However, most studies only focus on a specific cell type for their application. We demonstrate accurate recognition of cell stages in 21 different species of *Caenorhabditis* embryos of varying embryo sizes and strikingly different mechanics of spindle dynamics [272].

In addition to recognition of robust image features, CNNs have been demonstrated to capture the latent relationship between classes as shown by DeepCycle for kidney cells [226]. A similar example has been shown for reconstructing disease progression in diabetic retinopathy [75]. A key difference in our work is the use of label-free imaging data where no cell cycle markers were used. The reconstruction of the cell cycle in UMAP space was solely dependent on the presence or absence of cell cytoplasmic features.

With the Grad-CAM analysis, we demonstrate that the decision-making of our networks is strongly dependent on visible cytoplasmic features. By tuning the network depth, we were able to restrict the gradients only to the most important differences among classes. In EvoCellNet, the final classification is precisely overlaid with cytoplasmic features. This, however, does not mean that other networks' predictions are not

dependent on these features. By reducing the network depth, we were able to increase the focus of convolutional filter parameters only on the cytoplasmic structures of the pronucleus, mitotic spindle, and the cytoplasmic furrow. The other deeper networks, however, have the ability to learn more global features which lead to their superior performance in segregating the different classes, with VggNet demonstrating the best segregation of class clusters in UMAP space.

In the future, we could explore the integration of time-series data as input to the deep learning models, allowing for the consideration of temporal dynamics in cell stage recognition. We could further investigate more advanced time-series modeling techniques to improve the prediction scheme and the quality of the UMAP trajectory within a class sector. The recognition of cell stages based on cytoplasmic features holds the possibility of automating the comparison of the first mitotic division across diverse nematode embryos. The learned features of the trained networks, both deeper and simpler architectures, can be used to compare spindle mobility differences. The possibility to track or segment cytoplasmic features during the first mitotic division can be extended to include non-*Caenorhabditis* species of the nematode genera for a wider range of comparison of cellular events.

4.5 Conclusion

In conclusion, in this study we presented a classification pipeline utilizing a shallow convolutional neural network (CNN) to accurately identify different stages of the first mitotic division in diverse *Caenorhabditis* embryos. Time series recordings of nematode embryos from a published database were collected, and manual annotation was performed based on a morphology of the four stages. Random oversampling and data augmentation techniques were employed to address class imbalance to ensure robust network training. Four different CNN models were evaluated, including a custom-

designed network (EvoCellNet), ResNet, VggNet, and EfficientNet. All models demonstrated high accuracy (95%) in classifying cell stages, with the EvoCellNet achieving a comparable performance despite its simpler architecture. Dimensional reduction technique UMAP and gradient visualization using Grad-CAM, provided insights into the decision-making processes and feature detection capabilities of the CNN models. The image gradients visibly overlap with cell stage dependent features and UMAP representation of prediction vectors retained the latent relationship between the annotated cell stages. The accurate classification output allowed for the automatic identification of key cellular events using a linear SVM kernel in UMAP space. The temporal accuracy of the identified checkpoints was within a reasonable range, demonstrating the effectiveness of the classification pipeline in accurately recognizing and tracking cell cycle progression.

Chapter 5

In vitro microtubule oscillations
quantified using graph based
segmentation approach¹

¹Primary data acquisition by Shivani Yadav

5.1 Introduction

Microtubules (MTs) are essential components of the eukaryotic cytoskeleton, playing key roles in cellular functions such as cell division, migration, and the propulsion of cilia and flagella. MTs are composed of 13 protofilaments which are arranged in a hollow cylindrical geometry wherein each filament is composed of repeating heterodimers of α - β tubulin [68]. Their hollow tubular structure allows them to withstand compressive forces which provides mechanical support to the cell, thus maintaining its shape [115, 123]. Structurally, microtubules have outer and inner diameters of 25 nm and 17 nm, respectively, and their lengths can range from tens of nanometers to the micrometer scale. Despite their considerable persistence length of 0.2-9 mm [96, 126], microtubules do not appear uniformly straight in the cell. Experimental observations of microtubules showing periodic curves indicate their ability to sustain compressive forces as they buckle under compression [111]. Similar buckling observations have been made during their constrained polymerization near the cell periphery [23] and under axial compression driven by molecular motors [95, 276, 294]. Although *in vivo* buckling is damped due to the surrounding elastic and viscous cytoplasm [27, 155, 245], *in vitro* microtubule buckling is more pronounced, exhibiting high-curvature geometry and sustained oscillatory beats [158, 276, 294].

Numerous theoretical studies have explored the emergence of beating patterns in single filaments [41, 58, 83, 85]. Although, an extensive theoretical framework of buckling in slender rods exists, *in-vitro* observations of reproducible oscillatory filaments have been limited [95, 276]. In previous work from our lab Yadav et al. [294] demonstrated a dynein gliding assay setup where filament tips are fixed. This setup is capable of generating systematic beats of MTs of varying lengths and motor density [294]. This setup was used to reproduce oscillatory dynamics of tip immobilized MTs in a under controlled motor density. The experimental design generates large number of MT

movies where sustained oscillations can be observed. Although spindle oscillations are much more complex, our setup can be used to study differences in these oscillatory dynamics with changing motor numbers which is critical for spindle oscillations [109]. Additionally, quantification of this dynamics can help in understanding biophysical properties of MTs.

To quantify the dynamics of this buckling instability, accurate segmentation of filament geometry from *in-vitro* reconstitution assays is required. Extraction of a contour in sense of a continuous curve has been previously done using curve fitting approaches or with the use of levels sets [132, 232, 292, 293]. The segmentation of microtubules in buckling states, however, poses a significant challenge. The problem has two aspects: (a) filaments can demonstrate excessive looping and self-intersections, and (b) they can rapidly transition from curved to linear geometry when they relax spontaneously. Hence, an effective segmentation approach is needed which is robust to complex geometries and rapidly changing filament positions.

Graph-based approaches have previously been used in resolving objects that overlap and intersect. Notable examples in biology include *C. elegans* adult segmentation and cytoskeleton meshworks [26, 228, 292]. Indeed, Xu et al. (2019) have combined graph-based tracking and level sets to achieve good results in tracking of intersecting filaments. However, a limitation of their local matching, is apparant in the incorrect orientation of resolved microtubules (MTs) (*see Figure 5.7*). Therefore, there is a lack of single filament extraction methods that can reliably segment and maintain the correct geometry of a "beating" microtubule, particularly in instances where it exhibits self-intersections and overlaps.

In this chapter we present a novel semi-automated pipeline for tracing complex single filament contours observed during *in-vitro* MT reconstitution beating assays. The approach utilizes active contours seeded by an initial guess based on an intensity threshold for filament segmentation. The novelty of our approach lies in resolving multiple

crossovers and loops through a directed graph-based path-finding method combined with a template matching approach which uses dynamic time warping (DTW). We demonstrate accurate segmentation of filament contours while maintaining the overall geometry that is robust to filament intersections and rapidly changing geometry.

5.2 KnotResolver overview

The KnotResolver pipeline consists of two modules. The first module performs segmentation to extract skeletons of a single microtubule from a beating assay. To run the analysis a ".csv" file is first collated which contains the path information of all time series available for analysis. This file was used to optimize segmentation parameters interactively using *KnotSegmenter.m* script. The optimized parameters are used as input by *KnotResolver.m*, which is prompted at the end of the optimization script. The second module utilizes the proposed graph based approach to identify the order of pixel indices that preserves the correct filament geometry of the segmented skeleton. The overall workflow of the pipeline is illustrated in Figure.5.2. The components of the workflow are discussed in detailed.

5.2.1 Filament segmentation

Image preprocessing include image contrast enhancement which saturates the top and bottom 1% of all intensity values. Then median filtering with a 3 x 3 kernel is performed to smooth the intensity fluctuations along the filament contour. The pre-processed image stack is binarized using Otsu's method [207]. The segmented contour is used as the initial seed for an energy-based active contour model based on Chan and Vese's method [36]. This improves the segmentation by optimizing for filament boundary. Finally, the optimized binary contour is skeletonized using the medial axis transform [153] to obtain one-pixel width filament contours, which are stored in a structured array

for all frames in the image stack. The segmentation process is visualized in Figure. 5.2(A). The different parameters for segmentation can be interactively tuned using the KnotSegmenter.m script.

5.2.2 Graph representation of self-intersecting filaments

Skeletons obtained in the previous step are checked for the presence of branch points in a 3x3 window neighborhood at every binary pixel. The condition verifies whether the sum of the elements in the window neighborhood is greater than or equal to 4, with the central element (position 5) set to 1. If this condition is satisfied, the pixel is labeled as a branch point. In cases where there are no branch points in a segmented skeleton, the correct orientation of the filament indices is determined using the geodesic distance algorithm [246]. This algorithm fixes one of the end points of the filament as a seed and organizes the remaining indices in order of increasing distance from the selected end point. This however will not work for branched structures. We therefore proceed to generate a graph representation that allows us to navigate this complex geometry.

We represent a skeleton containing branch points as a directed graph $G=(V,E)$, made of vertices V and edges E , with the direction given by a chosen start point. The vertices fall into two categories: branch vertices V_b and filament segment vertices V_f Figure.5.2(B). A branch vertex is defined as a group of connected pixels that satisfy the branching criteria described earlier and the remaining regions are labeled as filament vertices. Each individual pixel is labeled based on a 4-neighborhood connectivity criterion, ensuring connectivity among all neighboring pixels. The graph's edges E result from connectivity between unique label groups, specifically forming edges between every pair of vertices, one being a V_b vertex and the other a V_f vertex. This guarantees edges exclusively between neighboring pairs of V_b and V_f vertices Figure. 5.2(C). Instances where a filament merges with itself are represented as loops in the graph. All edges are bidirectional, enabling traversal in either direction, except for start and end

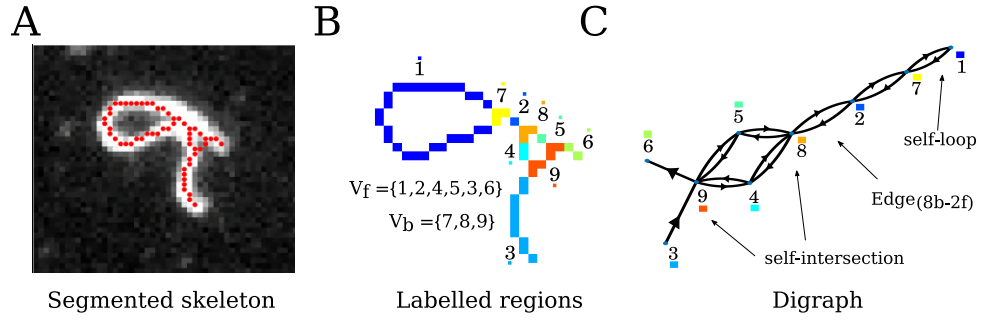


Figure 5.1: **Filament graph representation.** (A) A representative image (gray) with the segmented contour overlaid (red) is converted into (*middle*) labelled regions between the points of branching (colors, numbers). The regions are used to construct a directed graph of the connected components (nodes) and paths (edges).

vertices where traversal is limited to a single direction. A representative example of a branched skeleton, its diverse connected components, and the corresponding graph representation are illustrated in Figure. 5.2(A-C).

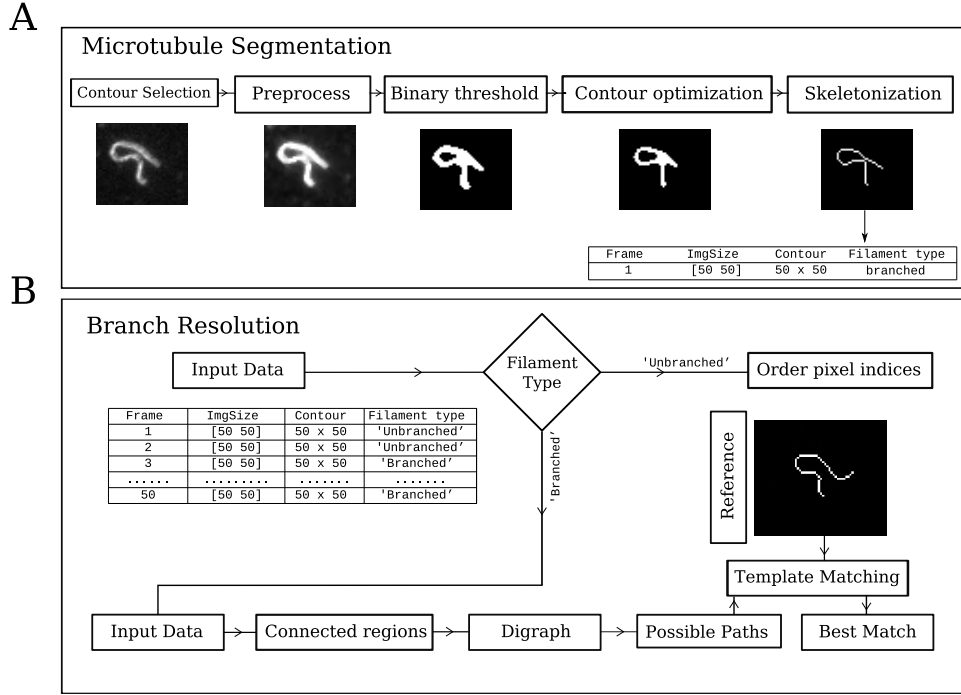


Figure 5.2: Filament crossover resolution by graph representation and minimization of curve dissimilarities. (A) Individual frames from a selected time-series of rhodamine-labeled MTs are segmented to result in a binary image. The contour of interest is selected interactively and the nearest tip to this point defines the fixed end of the filament. The binary image is used to seed the active contour optimization for improved accuracy of contour detection. The optimized binary image is skeletonized resulting in a contour with a 1 pixel line width using medial axis transform. Contours consisting of pixels with neighbours ≥ 4 are labeled as branched, suggesting a crossover event. The individual steps are written in sequence and a representative output is shown below each step. For every frame details of image size, The resulting segmented contour and filament type is stored in a structured array. (B) All contours are identified in the time-series and when branches are detected, they are mapped to a directed graph (digraph) where connected regions are the edges and the branch points (crossovers) are the nodes. All possible paths are calculated for the digraphs and the distance minimization with the last unbranched contour used to find the best match.

Graph based template matching using dynamic time warping (DTW)

A graph $G = (V, E)$ represents a branched skeleton, where $V = v_i$ denotes the set of vertices and $E = e_{i,j}$ denotes the set of edges. The graph structure is determined by the number of unique branch points, which either indicate intersections with other filaments or an instance where the filament forms a loop. We generate possible path combinations between the start and endpoint by traversing every filament segment vertex V_f and through its connected branch vertex V_b . For graphs that lack an endpoint, the path ends at the last visited node. Since the number of paths can increase exponentially with the number of nodes, we limit the number of possibilities generated at each step to 10. This could potentially return a sub-optimal path as evident for complex filaments from our simulation dataset.

To avoid this we implemented a contour template matching step to find sequence of labels that are most likely to fall in order. Essentially, we try to find larger filament regions that have not moved from the previous plane. To do this, we perform a rough template match to obtain an initial estimate of the optimal path through the graph. In this process we first generate a Euclidean-based template between the current and previously resolved skeleton. By doing so, we can pinpoint regions that have remained unchanged, particularly the fixed-tip in our beating assay. We select three labels which have deviated the least from the previous frame, these are annotated as $V_{m=1to3}$. We then check whether the indices of each of these labels fall in sequence. Identified indices for each label should have more than N (N set to 8 for our analysis, the parameter can be tuned according to filament length) indices falling within a threshold distance d ($d = 2$ pixel). Subsequently, we again use the brute force search, focusing on the most critical vertices. By limiting the search on these key vertices, we significantly reduce the number of paths generated for more complex graph structures. Once all the probable paths are generated, we employ dynamic time warping to select the best match.

The problem of resolving intersections is solved by finding the correct walk through the graph network among the generated possibilities. The correct walk is selected as the best match to a reference contour using a template matching algorithm called dynamic time warping (DTW). The goal is to first find the optimal alignment between the reference $R(x_j, y_j)$ and the selected sequence $S(x_i, y_i)$, where i and j are the indices of x and y , respectively, minimizing the cumulative cost function D [234]:

$$D(i, j) = |S(x_i, y_i) - R(x_j, y_j)| + \min(D(i-1, j-1), D(i-1, j), D(i, j-1)) \quad (5.1)$$

This cost is compared among the possible paths, and the path with the minimal cost is selected. Pixel indices are then iteratively ordered, along the best found path using the geodesic distance approach used earlier [211]. Figure.5.2(B) (I-V) illustrates possible paths for branched contour with the reference shown in Figure.5.2(A). The cost of possible paths is shown in Figure.5.2(C). A montage of selected frames with overaly showing complete dynamics is shown in Figure.D1.

5.3 Materials and Methods

5.3.1 Generating simulated image of oscillatory filaments

To model a beating filament scenario, we utilized the stochastic simulation engine Cytosim [199]. This tool has been previously used for simulating the dynamics of microtubule-motor systems and the buckling dynamics of individual clamped microtubules (MTs) [294]. Using the parameters from these simulations, we systematically varied microtubule lengths and the density of minus-end directed motors, using dynein-like parameters as detailed in prior studies. Simulated microtubule lengths ranged from 5 to 50 μm , and motor densities included 12, 25, 50, and 100 motors μm^{-2} . Each sim-

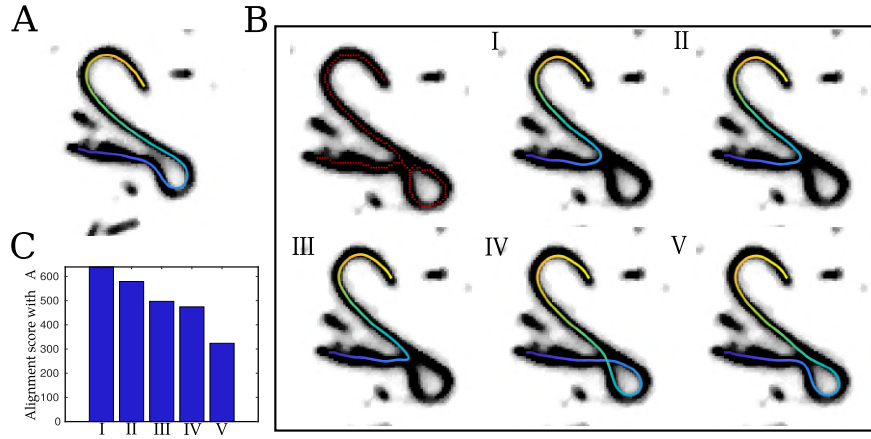


Figure 5.3: **Identifying the optimal contour by minimizing deviation from an unbranched contour.** (A) A representative image of a bend microtubule at time t , when it does not have any knots (unbranched) is with the segmented contour. Colors: position index along the contour, blue: start, yellow: end. (B) The microtubule in frame $t+\delta t$ showing a knot (branched) as seen from the contour (red line). (I-V) The possible paths from the start to the other end are overlaid as contours with the position-index in color. Blue: start, yellow: end. δt : interval between frames. (C) The alignment score based on dynamic time wrapping for all paths is calculated by comparing with the unbranched path from the previous time-frame, and the path with the lowest score is the optimal.

ulation was performed for 5 minutes and 120 frames were saved. The objective was to generate skeleton contours of varying lengths, resembling tip-immobilized filaments observed in a fixed tip gliding assays. The data range captures range of buckling dynamics with multiple loop formation which are rarely observed in experiments. Nonetheless the dataset allows us to interrogate the limitation of our method.

Adding noise to simulated images

The simulated skeletons were dilated size-2 diamond-shaped structuring element and convolved the dilated skeletons with a Gaussian point spread function (size: 3, sigma: 2). We introduced noise of varying signal-to-noise ratio (SNR_{db}) to the simulated data. The noise model is Gaussian with a specified noise power level P_{noise} . The desired P_{noise} is related to SNR by the relation

$$P_{noise} = P_{signal}/SNR \quad (5.2)$$

where P_{signal} is equal to the variance of intensity of the original image. SNR in linear scale is related to SNR_{db} by the following relation

$$SNR = 10^{\frac{SNR_{db}}{10}} \quad (5.3)$$

The SNR_{db} values used were 5,10,20 and 30. The generated Gaussian noise is computed as $Noise$,

$$Noise \sim \sigma_{noise} * \mathcal{N}(0, 1) \quad (5.4)$$

where, $\mathcal{N}$ is a normal distribution with zero mean and the σ_{noise} is equal to

$$\sigma_{noise} = \sqrt{P_{noise}} \quad (5.5)$$

The *Noise* is added to the original image to get a noisy image with specific P_{noise} .

5.3.2 Segmentation accuracy metrics

The simulated dataset generated with Cytosim is our benchmark to assess the proposed filament tracking method [199]. Segmentation accuracy is quantified using Dice score. While the ability to correctly arrange contour indices in branched contours is quantified using the Fréchet distance.

Dice score

It is computed using the following equation:

$$\text{Dice Score} = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (5.6)$$

where TP: True positives, FP: false positives, FN: False negatives and TN: True negatives. True positives are defined as segmented pixel indices that are within a threshold distance (2 pixels) from the ground truth annotations. The choice of the threshold was based on the point spread function (PSF) used to convolve the filament skeletons to resemble experimental images. False positives are segmented pixels which are not within this threshold. Ground truth pixels that were not detected as TP are counted as false negatives.

Fréchet Distance

Fréchet distance computes the minimum distance between two curves which is sufficient to traverse their separate paths. This provides the minimum of the maximum separation in a suitable distance metric between the two curves. The metric, here, is used for comparing the overlap of two curves while taking into account the order of indices between the two curves. For a pair of identical curves Fréchet distance should be

zero, which indicates a perfect overlap and identical ordering of pixel indices between two curves. Fréchet distance is computed between the indices of ground truth curve $I = [(I_1, I_1), (I_2, I_2), (I_m, I_m)]$ and the resolved curve $R = [(R_1, R_1), (R_2, R_2), (R_n, R_n)]$ obtained from the KnotResolver pipeline. Let D be the Euclidean distance matrix between the points of curves I and R . The computation of the Fréchet distance, $F(I, R)$ involves filling in a memoization table M , where, $M_{m,n}$ represents the Fréchet distance between the prefixes I and Q . $F(I, R)$ is obtained from $M_{m,n}$, where m and n are the lengths of curves I and R respectively. The recurrence relation for filling in M can be represented as follows:

$$M_{i,j} = \max(\min(M_{i-1,j}, M_{i,j-1}, M_{i-1,j-1}), D_{i,j}) \quad (5.7)$$

Where: $M_{i,j}$ is the value at cell (i, j) in the memoization table. $D_{i,j}$ is the Euclidean distance between the i -th point of curve I and the j -th point of curve R .

5.3.3 Analysis of filament contours

All analysis was performed in MATLAB R2020b (Mathworks Inc., Natick, MA, USA). We calculate the tangent, normal, and curvature vectors for each smoothed contour by using the Frenet-Serret formulas [177]. Smoothing of the x and y coordinates is performed using the Savitzky-Golay filter [235]. The end-to-end distance of the curve is computed as Euclidean distance between the tip and end of each filament. All results are visualized by plotting the curve with color interpolation which indicates the arrangement of contour indices from tip to end.

5.3.4 Manual correction of contours

The user can manually correct the path on problematic frames or when the initial frame is branched to initiate tracking. An example of manual editing on a branched

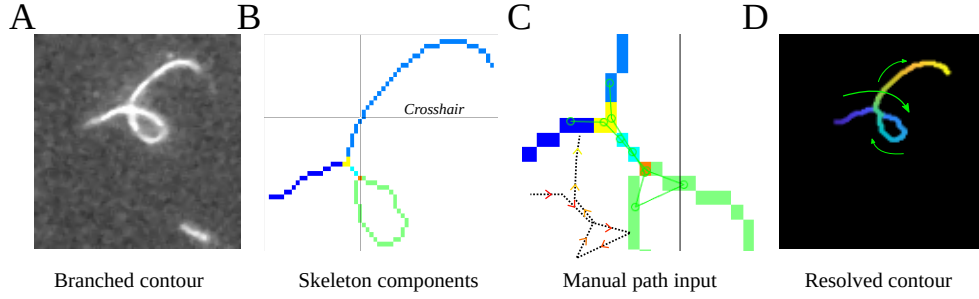


Figure 5.4: **Interactive resolution of complex knots from experimental data.** (A) Time-series of MT filament buckling used as the input to KnotResolver. The selected frame shows a filament with a self-loop. (B) The skeleton of the input frame in (A) with color-coded separate components of the skeleton. These are identified based on the discussed branch criterion. (C) Manual path connections are provided interactively over (B) using the crosshair. (D) The final segmented output with the input path showing the correct arrangement of pixel coordinates along the provided path.

skeleton after segmentation is shown in 5.4 (A-D). To perform manual edit the user can mention the frame number for edit in the **optimized.csv* file under the column header *manualCheck*. The final **optimized.csv* file can be run on the MATLAB *cmd* with the function *KnotResolver*. Additional details on code usage are available here.

5.3.5 Code Availability

The tool was written in MATLAB2020b (Mathworks, Natick, MA, USA) and uses functions called from the 'Signal Processing Toolbox', 'Image Processing Toolbox' 'Statistics and Machine Learning Toolbox' and an optional 'Parallel Computing Toolbox'. The code for the filament bending analysis has been made OpenSource and can be accessed for testing and future extension from Github at <https://github.com/CyCelsLab/MTLoopResolve>. Along with the code an example time-series with optimized segmentation parameters is provided for testing.

5.4 Results and Discussion

In this section, we validate our analysis pipeline and demonstrate its application in the analysis of microtubule segmentation in a beating assay. The evaluation involves assessing the performance of microtubule segmentation and branch resolution using simulated dataset of filament beating. Subsequently, we conduct a qualitative comparison with existing methods, followed by a demonstration of the pipeline’s effectiveness in analyzing single microtubule beating assays.

5.4.1 Performance validation of microtubule segmentation on simulated data

To assess the segmentation accuracy of our method we simulated multiple time series of tip immobilized microtubules undergoing buckling similar to our previous work [294]. We compared our combined approach of initial seed estimate using intensity threshold followed by active contour based optimization with cases where the two methods are used individually. The performance of these methods was compared at different signal to noise ratios (SNR). The results of this assessment is quantified using the Dice score and is shown in Figure.5.5. The raw data for this analysis is shown in Figure.D2.

Dice Score quantifies the overlap between the ground truth skeleton and the segmented skeleton. At SNR 5, threshold-only method exhibits a lower average dice score of 0.75 ± 0.02 compared to 0.94 ± 0.08 and 0.97 ± 0.02 for active contour and KnotResolver respectively. This difference is due to the poor estimate of the filament skeleton at low SNR conditions. In contrast, the active contour-based optimization shows robust performance even with high background noise, as demonstrated by a representative segmented output shown in Figure.5.5 (A). At higher SNR, the threshold-based method performs better due to the higher variance between the foreground and background which is ideal for segmentation. This, however, is rarely achieved in experiments. The

marginal drop between threshold and KnotResolver is due to loss of few pixels at the filament tip during contour optimization.

Although both the active contour based methods show stable performance with different SNR levels, KnotResolver has a higher Dice score. When only using active contours, the initial seed for each frame is determined by the previous optimized contour. This can lead to incomplete seeding of the microtubule contour when the filament rapidly changes its position. This difference can be exaggerated when the gap (Δf) between successive frames is increased as shown in Figure.5.5(C). The average dice score over the complete time series decreases for the active contour-based method with increasing Δf , while the performance of our combined approach is less affected by increasing Δf .

Our approach demonstrates robustness to lower SNR levels and rapidly changing microtubule positions, which is essential for accurately segmenting microtubules that exhibit beating patterns. The distribution of pixel distances between the segmented and ground truth contour, used for Dice score calculations, is shown in Figure.5.5(B).

5.4.2 Performance evaluation of branch resolution on simulated data

Although high segmentation accuracy is achievable using our approach, a segmented skeleton lacks any information about the geometry of the curve. To solve this we implemented a branch resolution module in our pipeline. The accuracy of this module is assessed using Fréchet distance which is sensitive to the difference between geometric shape of the ground truth skeleton and the output from the module. We compute this metric for every frame for our simulated dataset which also allows us to identify possible limitations of our method. The results of this evaluation are presented in Figure.5.6 (A-B) as a 2D matrix of Dice and Fréchet distance scores.

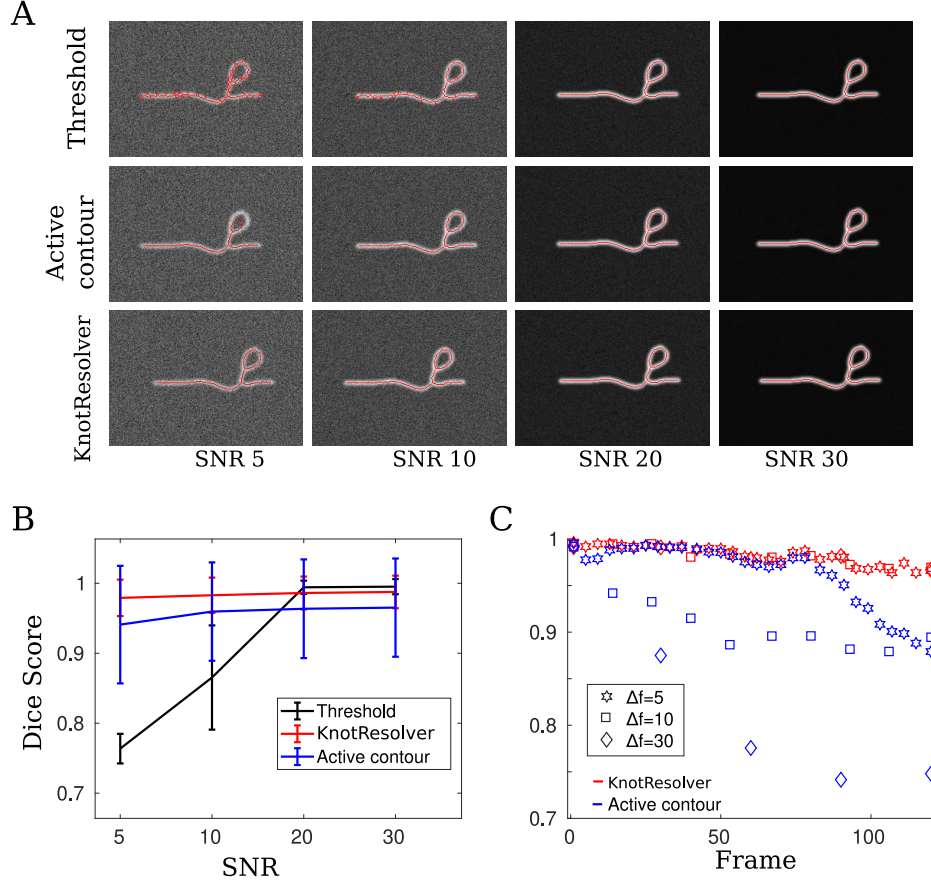


Figure 5.5: **Quantifying segmentation performance of KnotResolver and alternative segmentation methods with changing ‘noise’.** (A) Montage of segmentation outputs (shown in red) overlaid on the input image for the three compared approaches at different SNR. (B) Filament segmentation accuracy is compared using the Dice Score. Three alternative approaches were tested: threshold-based (black), active contour (blue) and KnotResolver (red). Error bars indicate the standard deviation around the mean. (C) Impact of increasing time step between adjacent frames on segmentation accuracy. Segmentation accuracy is compared between KnotResolver (red) and the active contour (blue) method for three time differences ("star": 5, "square": 10, "diamond": 30). Mean Dice scores are computed for every frame across all time series. All the analysis in (B and C) involves a total of $n = 24$ time series, each consisting of 120 frames. The points are average scores from each time series.

We find that during tracking of microtubules with lengths $> 20 \mu m$ there is a sharp decrease in the Dice score after ~ 60 frames Figure.5.6(A-B). Similarly the Fréchet distance increases rapidly indicating a deviation from correct geometry of the segmented curve. One such time series corresponding to the asterisk in Figure.5.6(A) is shown in Figure.5.6 (C). A closer analysis of all such time series revealed a key limitation of our method. The template matching fails in scenarios where the portion of the filament crosses the main filament body, causing it to re-emerge on the other side. Although the possibility of the correct path is considered a lower alignment score is assigned to the path with the overlap region. This results in breakdown of the correct geometry. Another problem is the loss of the free filament end, when it merges with a part of the main filament body. In such cases, we terminate the best path to the last visited node in the graph, resulting in a shorter filament, although with the correct pixel arrangement.

To address these limitations, we implemented a manual restart approach (*see materials and methods*). When encountering problematic frames, we can provide a manual path input interactively over the graph structure. This allowed us to capture the overall filament dynamics, as demonstrated in lower inset of Figure.5.6(C). This restart is effective but time consuming as the output contour is correctly arranged. While such challenging instances are rarely observed in our experimental dataset, the limitations underscore the difficulty of accurately resolving intersecting filaments when there is loss of information on filament tips or when there is a complete overlap of filament regions. Nonetheless the incorporation of an optional manual restart approach for difficult scenarios is an effective method for complete filament analysis.

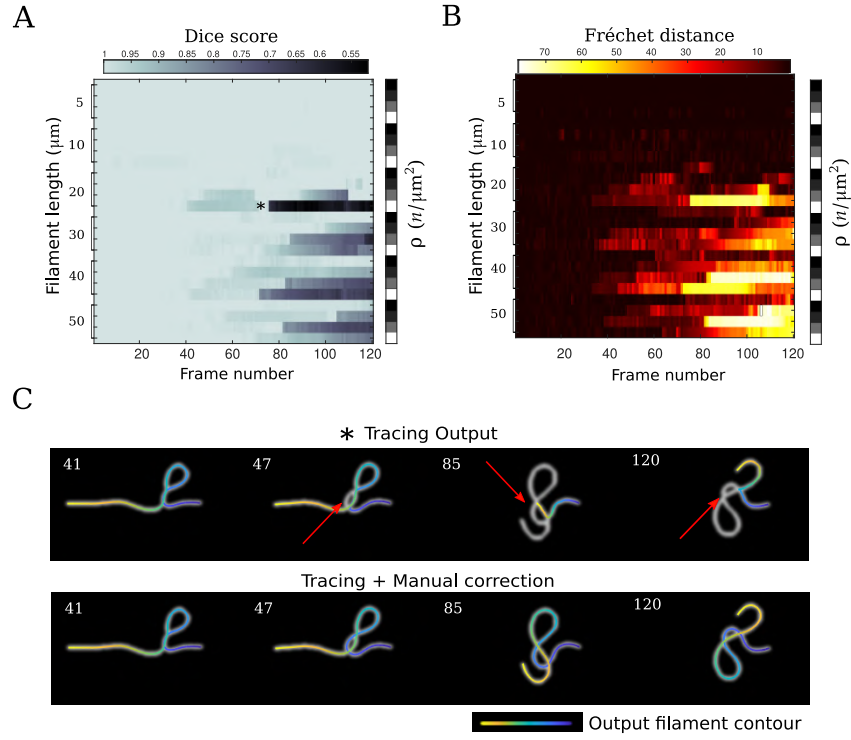


Figure 5.6: **Identifying the limits of detection accuracy from simulated data and manual path-input.** (A) Simulated time-series of filament buckling and bending are used as inputs to KnotResolver and the dice score (grayscale colorbar) estimates the detection. Simulations were run for increasing MT lengths from 5 to 50 μm (y-axis) and frame-numbers from 1 to 120 (x-axis). Each value of MT length was simulated for four motor densities 12 (black), 25 (dark gray), 50 (light gray) and 100 motors μm^{-2} (white). Asterisks mark the time series selected for further analysis in C. (B) A similar matrix to (A) with Fréchet distance (C) Branch resolution output using manual restarts at problem areas showing correct arrangement of the filament contour in comparison with (B).

5.4.3 Qualitative comparison with existing methods

We conducted a qualitative comparison of our filament segmentation and tracking method with two methods, FIESTA [232] and TSOAX [292]. FIESTA is a semi-automated software that uses intensity thresholding to segment filament contours. It employs specific Gaussian fitting terms to achieve sub-pixel localization. Filament intersections are handled through cross point identification followed by linking based on a minimum angle criterion. TSOAX is based on open active contours [293]. It automatically initializes multiple open curves during the detection stage, which elongate and stop at filament intersections and tips. Curve intersections are addressed by identifying T-junctions which are linked through a temporal local matching step. The segmentation comparison when self-loop and intersections are present for an example experimental time series is shown in Figure.5.7.

FIESTA output is over segmented for curved contours. The key limitation is curve fitting method which prefers linear geometry. TSOAX performed well in segmenting complex filament geometries. However, we observed incorrect orientation of the tracked filament in our analysis. This was evident in the color overlay representation of the contour arrangement, where gradient colors indicate the linking Figure. 5.7. The loop region of the filament had reversed its orientation. In comparison, our method was able to accurately follow the filament geometry in both the simulated and experimental time series. No manual corrections were made for TSOAX or in our method for this time series.

5.4.4 Semi-automated tracing of filaments in a pinned gliding assay

We demonstrate the application of our pipeline on an experimental dataset of fluorescence-based MT beating assays. In this modified assay, we achieve microtubule fixation at

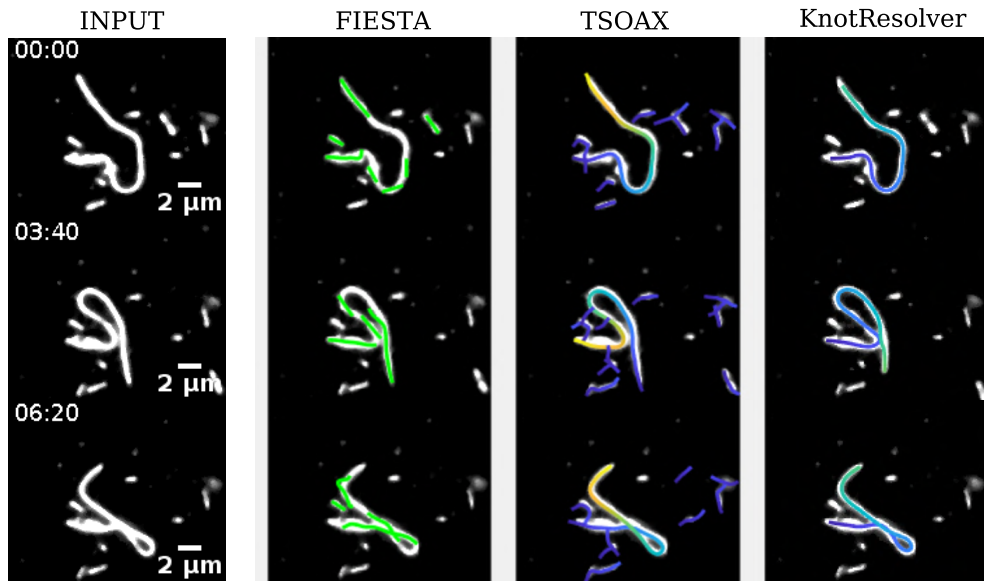


Figure 5.7: **Qualitative comparison with of KnotResolver with threshold-fitting and active contour based approaches.** (left) A representative time-series of labeled MTs undergoing bending and buckling, was used as input to test existing tools for filament tracking. FIESTA segmentation result is shown in green [232]. Parameters used were FWHM = 1000 nm, proportion curved filaments = 25%, minimum track length = 1 frame, maximum break = 4. TSOAX is based on active contours [292]. Default parameters were used for TSOAX except of the manual input of the minimum and foreground intensity value. Parameters used for KnotResolver were intensity threshold: 0.46, number of iterations for active contours: 10, contraction bias and smoothing factor for active contours: 0.4 and 0.2 respectively. No manual corrections were made for this time series. Scale bar: 2 μm . Time: mm:ss.

one end, leading to an increased frequency of microtubules exhibiting beating-type phenomena and resulting in complex microtubule geometries.

Segmentation and tracking results for four selected time series are shown in Fig 5.8. Our method is able to segment and track single filaments with good accuracy. Segmentation parameters were tuned for each time series separately. We utilized the pipeline to visualize the beating dynamics in terms of curvature, tangent angle kymograph, and a metric referred to as d_r , which represents the ratio of filament length to the end-to-end distance Figure.5.8.

The complete time series with segmentation overlay is shown for three time series. Supplementary Video SVD1 corresponds to the complete data shown in Figure.D1. Supplementary Video SVD2 demonstrates complex buckling dynamics of a single MT where the limitation of our method as found by the simulation dataset can be identified. Manual corrections were made three times in this analysis. The segmentation is overlaid with Frenet–Serret vectors. Supplementary Video SVD3 emphasizes the robustness of our method to the presence of additional filaments and rapidly changing MT geometry. Overall, these results demonstrate optimal tracking of intersecting filaments which can be used to study the geometric properties of these curved filaments.

5.5 Conclusions

In this study, the proposed pipeline for filament tracing was first validated on a simulated dataset. The pipeline was then applied to experimental data obtained from a fluorescence-based gliding assay, which was modified to a beating assay by fixing one end of the microtubule. This modification resulted in higher frequency of microtubule beating, which was analyzed using the proposed pipeline. A total of 10 time series of microtubule beating were processed using the pipeline, out of which only two time series required a manual restart at similar scenarios as identified in the simulated time

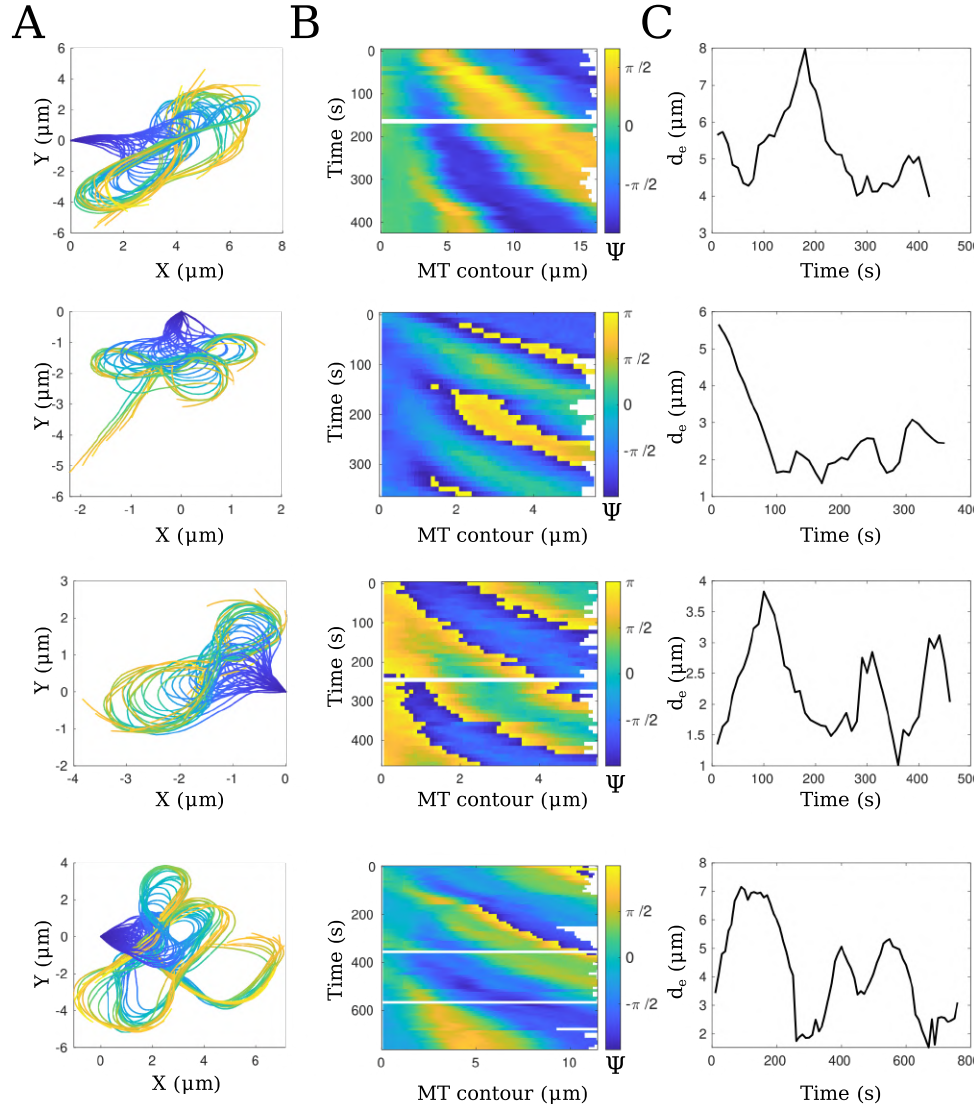


Figure 5.8: **Contour analysis in time of representative experimental filaments.** Four filaments showing beating mobility patterns are shown. (A) The skeletal line in the montage is color coded along its length (from tip to end). (B) 2D plot of tangent angle along the microtubule length (μm) is shown. (C) End to end distances along the time series is shown.

series. These time series were highlighted in the analysis. Key merits of the branch resolution pipeline were also demonstrated in representative examples of filament loop formation and multi-filament intersections.

The pipeline was also used to visualize beating dynamics in terms of curvature, tangent angle kymograph, and d_r , which is a ratio of filament length to the end-to-end distance. The analysis of multiple time series of microtubule gliding, in which the tip of the microtubule was fixed, resulted in the identification of the filament contour, which was color-coded along its length. Finally, the complete dynamics of bending were displayed, providing a comprehensive understanding of the microtubule beating phenomenon. The findings from this analysis were utilized to quantify the emergence of wave-like oscillations under different configurations of MT geometry and density. Additional details can be found in [295].

Future improvements to our method can be implemented in the segmentation pipeline. Currently, the accuracy of tip localization is compromised as the segmented contour shrinks during optimization. A solution proposed in previous work involves extending endpoints while considering filament intensity discussed here [208]. More advanced methods of deep learning-based tip detection frameworks could also complement our approach. In the branch resolution module, enhancements can be introduced in the path generation segment. Although a brute force approach has provided satisfactory results so far, it could be substituted with an iterative approach that minimizes curve differences as they curve develops. Such an approach has the potential to address current limitations of our method, particularly in scenarios involving branch merging and re-emergence.

Chapter 6

Conclusion and Outlook

The main objective of this thesis was to develop a computational image analysis tool for quantifying comparative differences between diverse nematode species. Our central problem was to investigate apparent differences between spindle mobility as observed among selected species of *P. pacificus*, *O. tipulae*, *C. elegans*, *C. remanei*, and *D.sp.1* JU359. The initial observation of qualitative differences between the mobility of endogenous granules among the species motivated us to investigate the differences in cytoplasmic viscosity across the different embryos. We proceeded to develop a computational tool DICOT to accurately quantify intracellular particle flows from label-free differential interference contrast (DIC) microscopy images and time-series data [39]. In DICOT, a novel image filtering approach called scaling of Gaussian (SoG) is proposed, that enhances the contrast of granules in DIC images. Combined with automated thresholding and minimal distance-based tracking, our tool can accurately detect and track the motion of intracellular yolk granules in *C. elegans* embryos. We quantified the oscillatory dynamics of these granules during anaphase and found that the frequency matches well with the reported oscillation frequency of the mitotic spindle.

The ability to quantify intracellular flows and mechanics from label-free microscopy data has significant implications for understanding cellular processes with minimal perturbation. The viscosity metric quantified for *C. elegans* is in close agreement with that identified by Daniels et al. (2006) and Grill et. al. (2001) [55, 107]. Indeed, the basic assumption of our quantification is the passive nature of these yolk granules. The geometry and configuration of these granules is similar to that of exogenous injected particles that are used for microrheological analysis. Yamada et.al (2000) has demonstrated the possibility of comparing MSD profile of endogenous granules to that of exogenous particles [296]. Nonetheless, strong cytoplasmic flows have been observed which dominate the pseudo-cleavage stage during membrane ruffling and during the spindle oscillation and elongation stage [120]. Therefore, the involvement of active forces that may influence our quantification cannot be ruled out. Our first sanity

check was to quantify these granule mobility away from the spindle and during early anaphase. During this stage, cytoplasmic flows are minimal, and the spindle is mostly nascent. Our ATP knockdown treatment in *C. elegans* gives us confidence that indeed the granules are diffusive during tracking. We still observe slight ballistic movement near the cell cortex in some of our time series, but these flows are transient and localized. This is more pronounced on the posterior side of *P. pacificus*. The analysis of MSD is done only in the linear region of the mean MSD profile to limit the effects of these transient flows.

The work presents a general method for segmentation and tracking of both *in vitro* and *in vivo* DIC microscopy data. The tool we developed has wider relevance for biophysical measurements of cellular mobility and subcellular mechanics using label-free microscopy. While our method performs well on the datasets tested, it may have limitations in extremely crowded or noisy environments. Additionally, the accuracy of the method will depend on the proper selection of filter parameters, which could be challenging for users without domain expertise. Future work could explore combining our approach with deep learning methods for improved detection and analysis, as well as extending the tool to three-dimensional data and other label-free imaging modalities. Incorporating additional features, such as object shape or intensity profiles, into the detection and tracking algorithms could potentially improve the accuracy and robustness of the method. Our implementation of the tool as a GUI with interactive parameter optimization as well as choice of multiple filters is likely to allow for use on multiple data sets [39]. In conclusion, our approach represents a general method for segmentation and tracking of *in vivo* DIC microscopy data. *In-vitro* analysis of bead mobility was also performed and is share in more details in the paper [39].

Coming back to our initial question on cellular trait divergence, the differences in viscosity measurement proved to be an insufficient explanator of the changes observed during spindle transverse and longitudinal motion among the studied nematodes. It

does, however, offer a simple explanation for the absence of spindle elongation in *C. monodelphis*, which has abnormally high cellular viscosity. To further interrogate the differences among the species, we proceeded to interrogate the force acting on the spindle during the anaphase stage. Our collaborators conducted spindle recoil experiments from which we were able to estimate cortical forces and cytoplasmic elasticity (contributed by microtubules) using our estimates of cellular viscosity. We propose that a trade-off between cortical pulling forces and cytoplasmic viscosity results in spindle oscillations when the ratio of pulling forces to viscosity is high for a given embryo size [136]. Reduction of pulling forces can still lead to oscillations provided it is compensated for by reduced cytoplasmic viscosity, as seen in *C. remanei*. This provides a framework that could be tested with more species of nematodes, as well as generalized to other cellular systems.

One interesting future direction is the exploration of the molecular complex responsible for cortical pulling forces during spindle motion. The absence of certain anchoring proteins in the genomes of some species suggests rapid molecular divergence within closely related species. Investigating the specific proteins involved and their functional differences could shed light on the evolution of these crucial cellular processes. Establishing transgenic lines in these species for live microtubule tracking would be a vital step toward quantitatively measuring parameters like motor attachment and detachment rates, providing a deeper understanding of the molecular mechanisms at play. Another promising avenue is the in-depth study of the effective elasticity parameter, which is a combined measure of cytoplasmic stiffness and microtubule components. Investigating the molecular components responsible for this parameter and their role in spindle dynamics could reveal novel insights into the cellular mechanics of spindle behavior.

We also note that species that have high viscosity also have a longer cell cycle duration (correlation coefficient $R = 0.84$). This result suggests that the only combi-

nations that have been retained by natural selection are compensatory changes, where high viscosity is compensated for by a slowdown of the cell cycle or vice versa. Over the course of evolution, slow species could have afforded an increase in cytoplasmic viscosity because even though objects are slowed down by the viscous drag, they will have time to reach their final positions. Conversely, a low viscous drag may have preconditioned the emergence of a fast cell cycle. In the analysis of cell divisional time between *C. briggsae* and *C. elegans* it was found that these species show a difference of 2°C in the thermal range of fertility limits [12]. Interestingly, the range of first mitotic cell divisional time for a given species could be varied from 30 to 4 minutes in a thermal range of 5-30°C, demonstrating highly robust evolutionary mechanisms to handle changing embryo environmental conditions. Regardless of the orientation of changes, this interesting correlation raises the question of the selective pressure responsible for species-specific viscosity values. Also, quantification of cell divisional time was done manually and spindle trait variability within this thermal range was not quantified.

Motivated by this, we aimed to develop a classification scheme that could distinguish between different stages of the first mitotic division. Such a model could be used to automatically quantify cell cycle times across diverse nematode embryos. This will allow us to check for the correctness of our hypothesis on a more diverse range of nematodes. Classifiers built on AI such as CNN have also been shown to learn robust cytoplasmic features associated with a particular class [227]. Our motivation was to construct a classifier that could also learn spatial features associated with a corresponding class, which could aid in a more reproducible comparison of features associated with the first mitotic division. This can also be complemented with a comparison of spindle mobility changes with increasing temperatures.

We trained multiple CNNs and achieved solid performance on the held-out test dataset (95%) across four trained networks. Three of these were pre-existing networks: VGGNet, EfficientNet, and ResNet. Being deeper architectures, the features associ-

ated with their predictions were spread out. Although intermediate features would be more localized and refined, we asked if we could rely only on localized spindle features while making class predictions. To this end, we created EVOCELLNET, a shallow CNN whose activation are highly localized and depend only on cytoplasmic features, as validated by gradient visualizations [238]. This network could potentially be used in the future for a more thorough comparison to judge similarities in spindle dynamics across related nematode species. However, localized activation do not mean that the predictions are more robust. Indeed, having a larger receptive field allows learning of more robust and global features while making class predictions. Similar to previous work on reconstructing cell stage progression using Network predictions [226], we were able to reconstruct cell cycle trajectories of the first mitotic division across diverse nematode embryos. Using our network, we demonstrate automatic annotation of key cellular checkpoints with good temporal accuracy. Our network, specifically VGGNet, can annotate these checkpoints with 8, 25, 13 s accuracy for a time series which is on average 10 mins long, Table.4.3. Our trained networks, therefore, offer automated annotation and checkpoint recognition of four distinct stages of the first mitotic division, validated and tested on 21 different nematode species. Moreover, the integration of visualization techniques such as UMAP and Grad-CAM has provided a deeper understanding of the CNN models’ decision-making processes [181, 238]. Investigating how CNNs identify and prioritize features during classification can lead to the discovery of novel cellular markers and biological cues, potentially advancing our knowledge of cell cycle regulation and related cellular events.

The prediction scheme in the future could be improved to use time-series data instead of single frame input. Since the optimization is done to separate frames around the key checkpoints, the sequence or quality of the UMAP trajectory is compromised within a sector while being robust around its edges. This can be improved using more complex time series modeling studies. Overall, the combination of DL-based

annotation and automatic tracking of endogenous granules could help us build a dataset of intracellular parameters that could shed light on the evolution of cellular traits, more importantly, the limits and biophysical constraints that evolution faces.

A key component relevant to spindle oscillation and elongation are microtubules. Our analysis of spindle recoil showed minor differences between MT properties of elasticity. Nonetheless, a more quantitative approach using *in vitro* could be performed for a closer look. The oscillatory dynamics of MT driven by MT rigidity is not a novel hypothesis. Many simulation studies have shown sustained oscillation in linear filaments under self-propulsion [41, 109]. However, the simulations have lacked experimental evidence of any such relation. Our lab proposed an *in vitro* reconstitution system that would represent sustained MT oscillations [294]. This system promises to study such relationships in a controlled manner. However, the experimental system came with its own challenges.

The complex MT geometry proved to be difficult to analyze. We proposed a graph-based approach combined with a template-matching scheme that showed good performance in quantifying the oscillatory dynamics. The tool combines graph-based reconstruction of filaments combined with dynamic time warping to find the correct orientation of complex filaments. Through a thorough analysis of simulated buckling with our tool, we were able to highlight key limitations of our approach and added a manual correction that allows tracing of buckling filaments in highly complex structures. The development and successful validation of the contour analysis pipeline for filament tracing, along with its application to studying microtubule beating dynamics, represent a significant step forward for filament analysis. This novel approach based on graph representation opens the door to several future research directions and potential improvements.

Future work on MT analysis can focus on enhancing the accuracy and efficiency of filament tracing by developing more accurate segmentation algorithms tailored specif-

ically for curved filaments, along with removing the need for automated correction methods in contour loss scenarios. Integration of machine learning techniques, like deep learning, holds promise for more accurate and automated tracking, while extending the pipeline’s capabilities to 3D analysis and multi-filament tracking would provide a more comprehensive understanding of motor-microtubule interactions. Additionally, validating the pipeline in vivo scenarios can further drive innovation and address complex challenges in cellular bio-mechanics research, ultimately contributing to a deeper understanding of biological processes and cellular mechanics.

In addition to the work related to the above, I have also built a segmentation and analysis pipeline for bacterial cells and an RBC detection framework that is applied to detect RBCs in microfluidic chambers. The work on bacterial morphology analysis was used to understand the relationship between surface area (SA) and volume scaling in *E. coli* cells, which is an important area of research. From an image analysis perspective, the study highlights the challenge of accurately measuring cell dimensions, especially in cases of drug-induced filamentation or varying growth rates, where cells can exhibit non-standard shapes and sizes. This underscores the importance of robust image segmentation algorithms that can handle diverse cellular morphologies. Additionally, the study’s insights into the dynamics of single-cell growth and how it relates to scaling exponents provide valuable information for designing image analysis pipelines that can capture and quantify these dynamics [130]. The RBC detection pipeline was used to quantify the variable effect of different antibody treatments on adhesion of RBC to endothelial cells lined over a glass chamber [60]. This novel microfluidic chamber simulated *in vivo* conditions during malaria infection cycles. The detection pipeline can detect stationary RBCs as they come in contact with layered cells. The two studies have been added to the appendix with a short description of the respective algorithms.

The discussed chapters demonstrate the potential of image analysis in understanding biological aspects of cell division, cytoplasmic microrheology, and microtubule be-

havior, offering deeper insights into the world of cellular biology. Moreover, the interdisciplinary applications of these advanced computational methods span the fields of cell biology, biophysics, and evolutionary biology. Overall, the research contributes to development of computational image analysis techniques for biological research.

Appendix A

Supporting material to Chapter 2

A.1 Supporting Figures

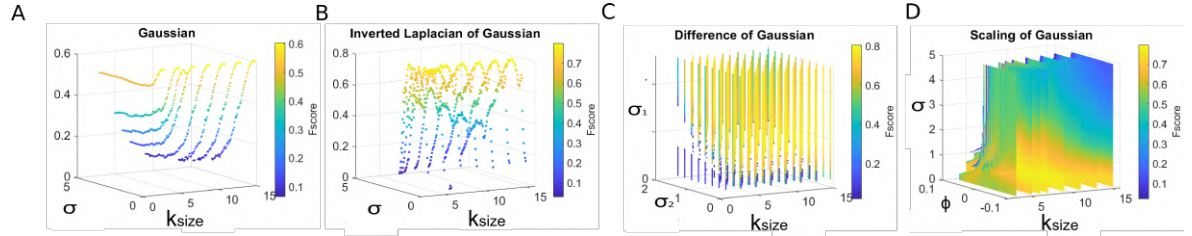


Figure A1: **Detection accuracy estimated by F1-score to compare filters.** (A) Gaussian, (B) inverted Laplacian of Gaussian (iLoG), (C) difference of Gaussian (DoG) and (D) scaling of Gaussian (SoG). For all filters the range of ksize was 3 to 15, σ was scanned over a range 0.1 to 5 with steps of 0.0495 while for DoG $\sigma_1 = \sigma$ while σ_2 was varied between 0.1 and 2 with steps of 0.211. For SoG the sensitivity factor ϕ was sampled between -0.1 and 0.1 with steps of 0.002. Colorbar indicates the F1-score.

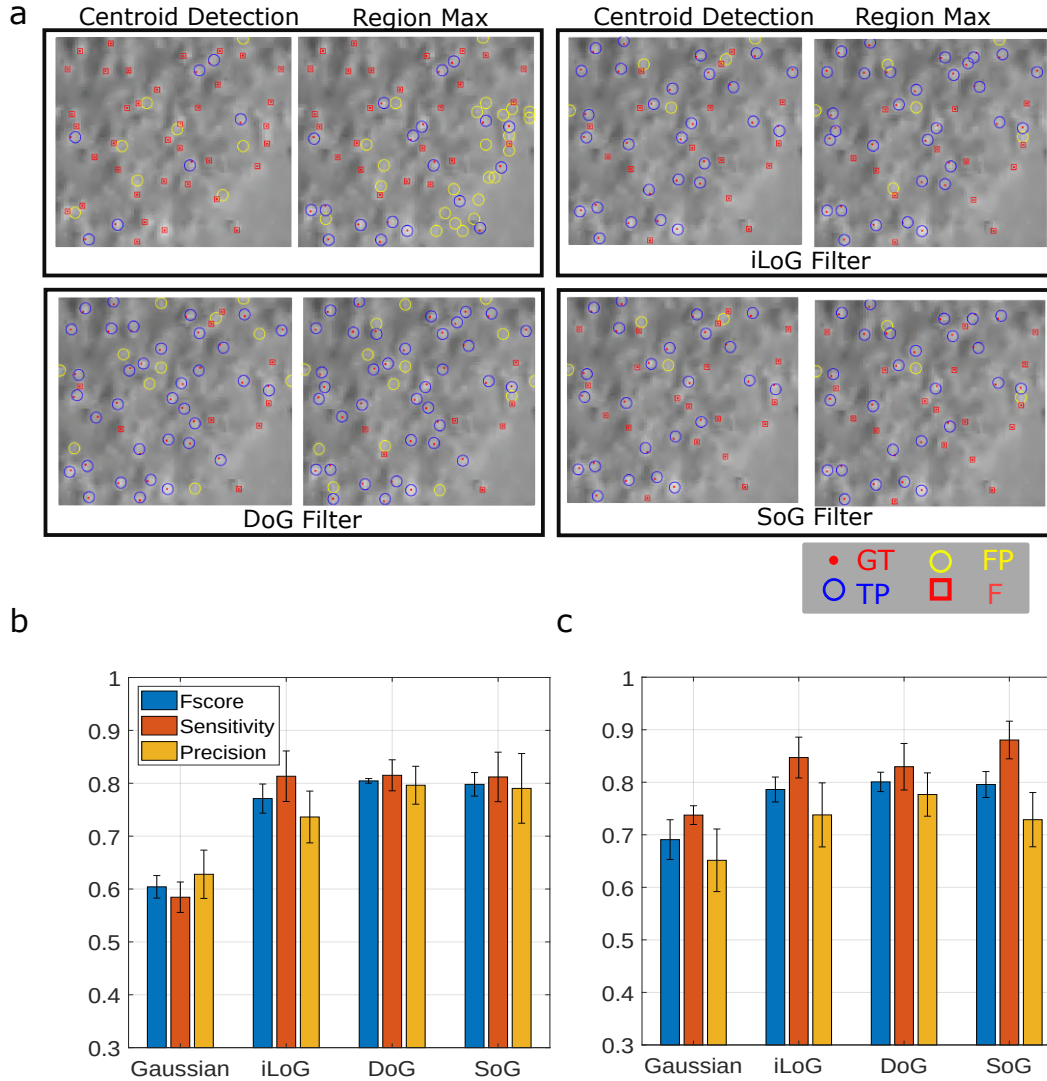


Figure A2: **Filters compared for sensitivity, precision and F1-score.** (A) An ROI from a mid-plane DIC image of a *C. elegans* embryo was manually annotated to mark lipid granules to serve as the ground truth GT (red dots) and compared to image-filtering with Gaussian, inverted Laplacian of Gaussian (iLoG), difference of Gaussian (DoG) and scaling of Gaussian (SoG). The respective true positives TP (blue circles), false positives FP (yellow circles) and false negatives FN (red squares) were determined as described in the methods section, comparing centroid and region-max based object detection. (B, C) The F-score, sensitivity and precision calculated from these comparisons for (B) centroids and (C) region-max based identification of objects are plotted.

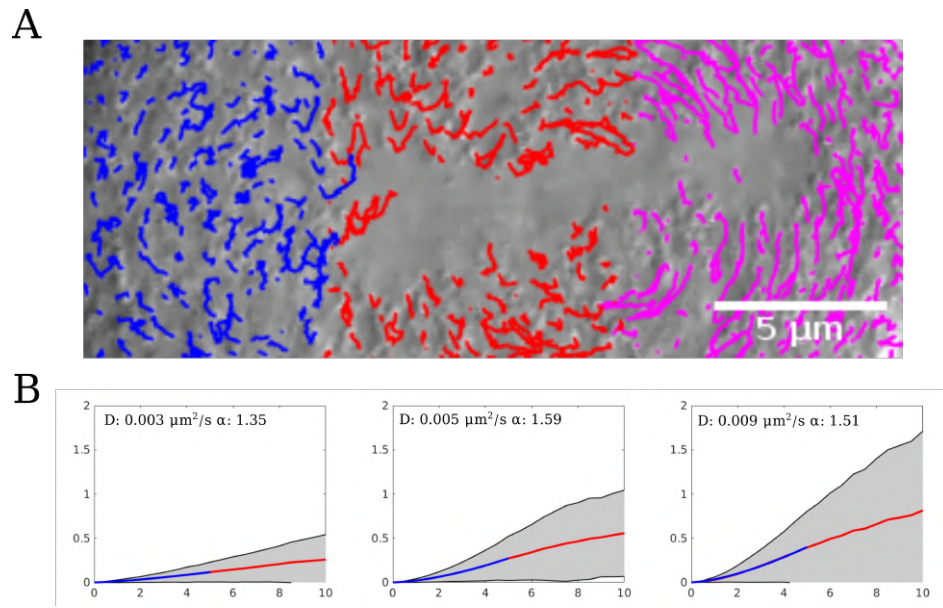


Figure A3: **MSD of granule from the embryo mid-plane.** (A) Tracked granules overlaid on a representative cropped region from their respective plane of acquisition (B) MSD measurement from the midplane, at anterior mid and posterior region (Blue: anterior, Red: posterior).

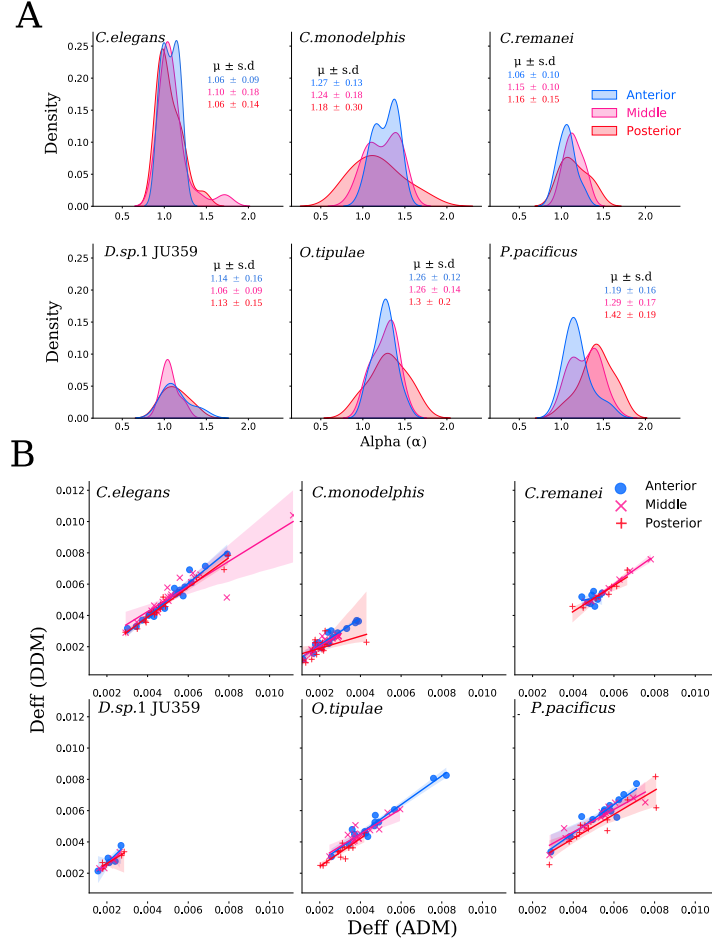


Figure A4: **Comparison between drift diffusion model (DDM) and anomalous diffusion model (ADM).** (A) Anomalous coefficient (α) density distribution for all six species in anterior (Blue), posterior (Red) and middle (Pink) region. Legend values indicate distribution mean with standard deviation of respective color-coded regions. (B) Regression plot of diffusion coefficient estimates from the drift diffusion model (DDM) and anomalous diffusion model (ADM) plotted against each other for the anterior (Blue), posterior (Red) and middle (Pink) regions. KS statistics show no significant distribution of the two estimates across regions. KS statistic for ADM vs DDM coefficient; Species (KS statistic) : P-value; *C. elegans* (0.066): 0.99; *C. remanei* (0.093):0.99; *D. sp. 1 JU359*(0.2):0.83, *C. monodelphis* (0.117): 0.81, *O. tipulae* (0.071): 0.99; *P. pacificus*(0.192):0.29.

Appendix B

Supporting material to Chapter 3

B.1 Supporting Figures

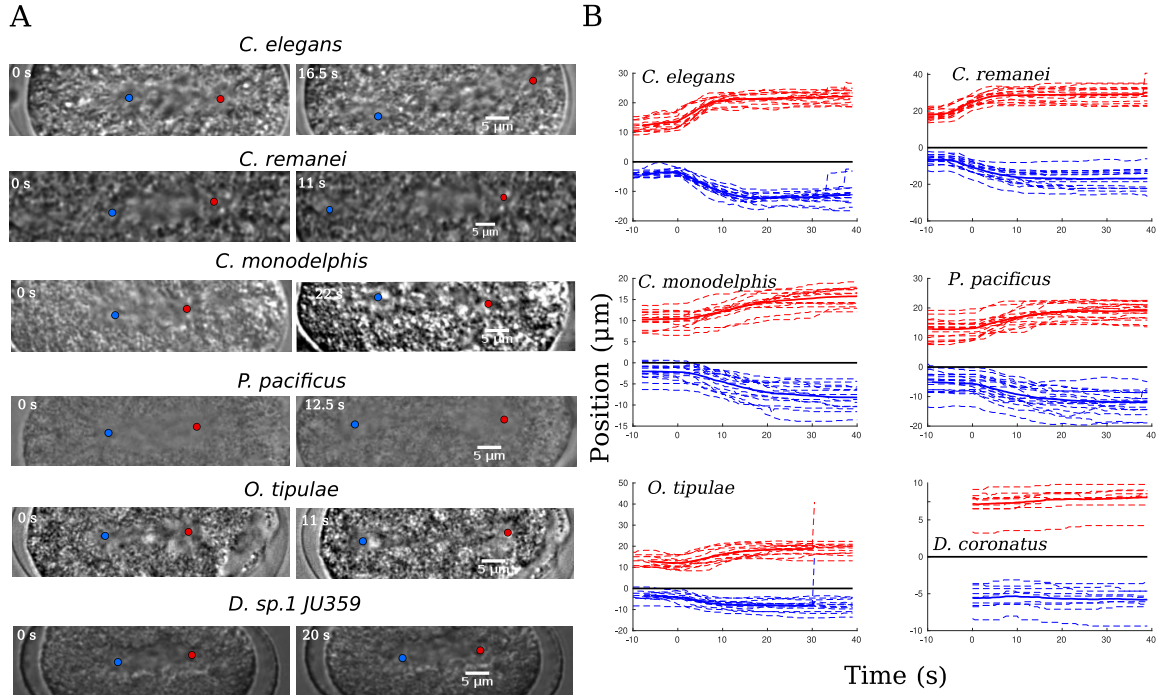


Figure B1: **Effect of spindle cutting on centrosome recoil.** (A) Representative images from immediately before (left column: $t = 0$ s) and after spindle cutting (right column) for different species are shown. The images after cutting are selected based on the approximate time required for the centrosomes to achieve their maximal displacement from the initial position. Scale bar $5 \mu m$. (B) The position of the spindle pole along the AP axis is plotted as a function of time. $t = 0$ signifies onset of anaphase, i.e. spindle elongation (which also corresponds to the time of spindle severing). Position = $0 \mu m$ indicates the geometric centre of the cell along the AP axis.

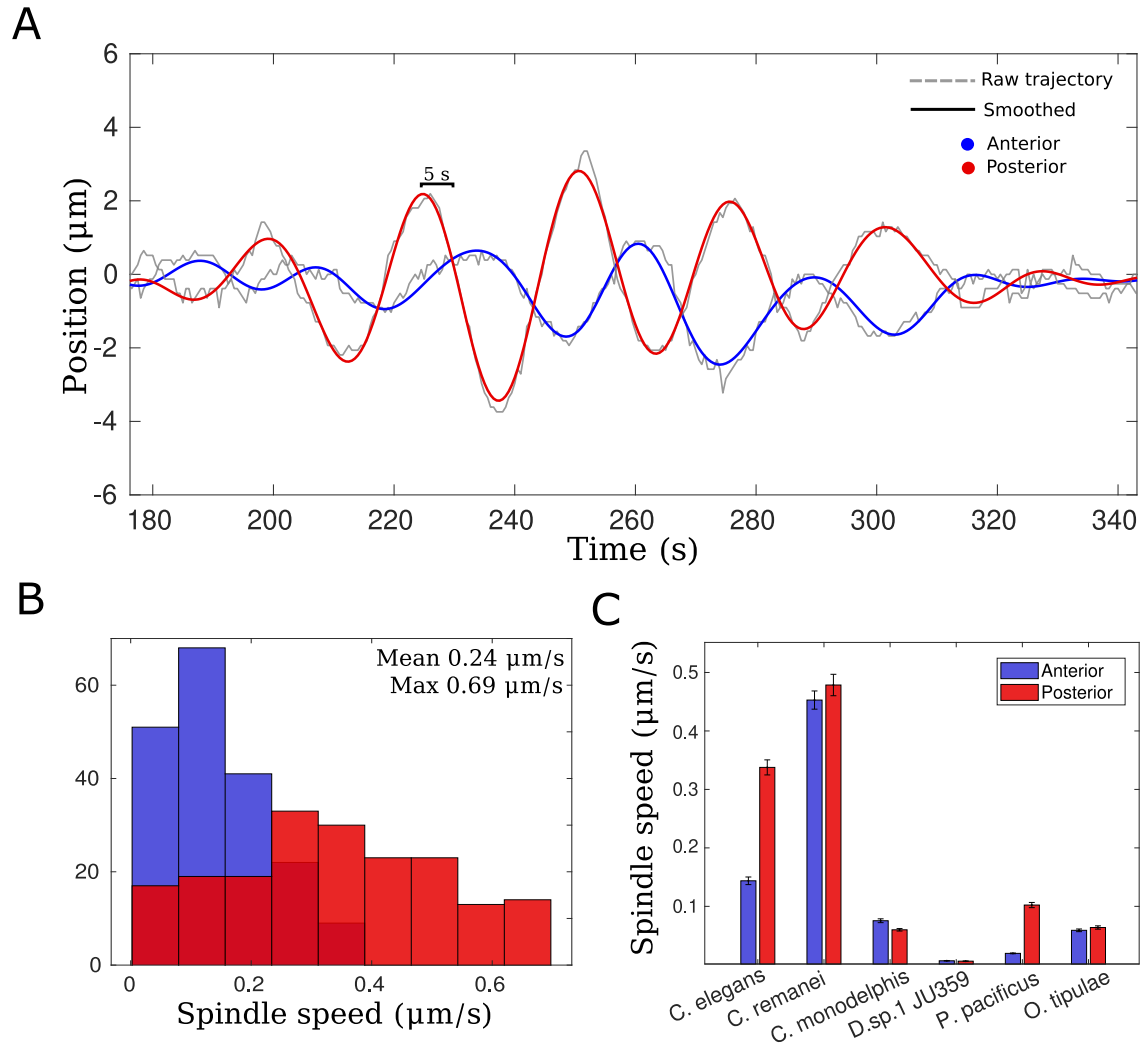


Figure B2: **Spindle oscillation speed.** (A) The spindle pole positions of a representative *C. elegans* embryo are plotted relative to the centerline as a function of time (black line) for anterior and posterior poles. Small fluctuations are removed by smoothing the data (blue: anterior, red: posterior). Oscillation speed is estimated as a change in position in a 5 s window successively throughout the trajectory from the smoothed trajectory. (B) The frequency distribution of anterior (blue) and posterior (red) centrosomes is plotted. (C) Speeds (mean $\pm$ SEM) are compared between the 6 species.

Appendix C

Supporting material to Chapter 4

C.1 Supporting Tables

Network	Precision	Recall	F1 Score
EvoCellNet	0.95 ± 0.002	0.96 ± 0.002	0.95 ± 0.002
ResNet	0.97 ± 0.005	0.97 ± 0.004	0.97 ± 0.004
VggNet	0.97 ± 0.001	0.97 ± 0.005	0.97 ± 0.003
EfficientNet	0.96 ± 0.005	0.96 ± 0.003	0.96 ± 0.004

Table C.1: **Training metrics.** The average precision, recall, and F1-score metrics with standard deviation for the training dataset were calculated across the four folds of the k-fold split.

Network	Precision	Recall	F1 Score
EvoCellNet	0.95 ± 0.006	0.96 ± 0.004	0.96 ± 0.005
ResNet	0.97 ± 0.006	0.97 ± 0.005	0.97 ± 0.005
VggNet	0.97 ± 0.005	0.97 ± 0.006	0.97 ± 0.005
EfficientNet	0.96 ± 0.002	0.96 ± 0.000	0.96 ± 0.001

Table C.2: **Validation metrics.** The average precision, recall, and F1-score metrics with standard deviation for the validation dataset were calculated across the four folds of the k-fold split.

	Pronuclear migration	Centration rotation	Spindle displacement	Cytokinesis
Test				
EvoCellNet				
Precision	0.96	0.95	0.84	0.97
Recall	0.99	0.91	0.91	0.94
ResNet				
Precision	0.98	0.93	0.89	0.96
Recall	0.96	0.94	0.89	0.96
VggNet				
Precision	0.96	0.93	0.89	0.97
Recall	0.99	0.93	0.88	0.96
EfficientNet				
Precision	0.98	0.94	0.89	0.97
Recall	0.98	0.94	0.89	0.96

Table C.3: **Network performance on each stage for test dataset.** The average precision and recall metrics with standard deviation for individual stages for all the four models used in this study. The value are calculated from the models used in the first fold only.

Network	CPU (ms)	GPU (ms)
EvoCellNet	15±0.6	0.3±0.05
ResNet	28±0.5	0.4±0.02
EfficientNet	30±0.5	0.5±0.01
VggNet	170±10	1.8±0.15

Table C.4: **Average inference time of trained on a single image.**

C.2 Supporting Figures

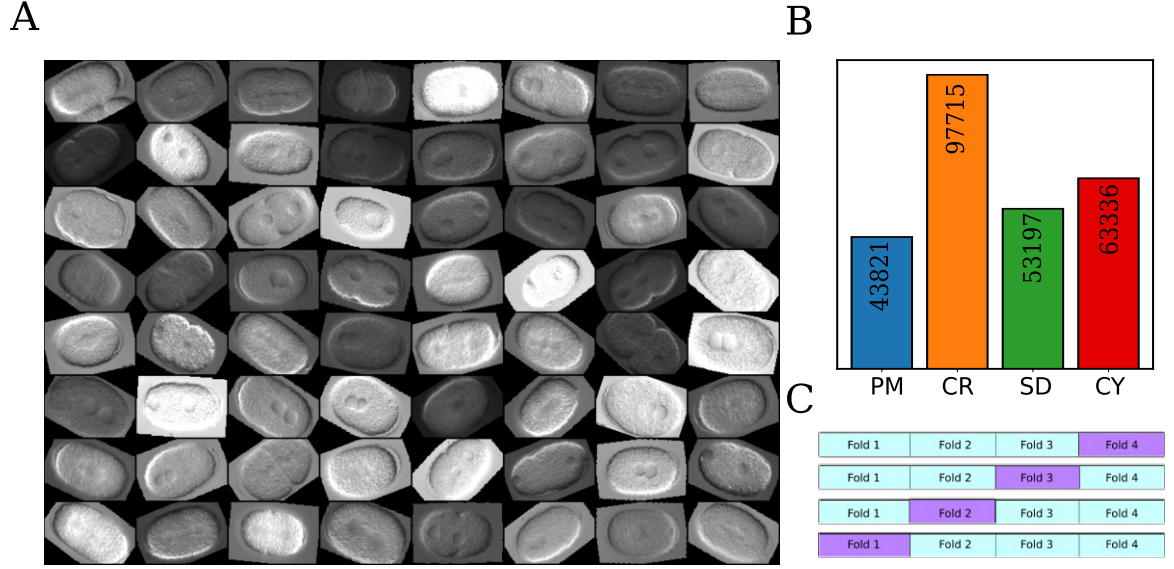


Figure C1: **Training dataset with image augmentation.** **(A)** The image montage consists of a batch of images (64) taken from the training dataset. Different augmentations such as image rotation, flipping (vertical and horizontal) and affine transformations have been applied randomly. **(B)** The bar chart shows the total number of frames obtained after annotation for each class. Average size of each frame is 170 Kb, the size of whole dataset after distributing the frames into four classes is ~ 45 GB. Color indicate annotated stages; PM (blue), CR (orange), SD (green) and CY (red). **(C)** The schematic represents k-fold cross validation scheme using which the training dataset was divided into four equal proportions maintaining equal proportion (25%) of each class.

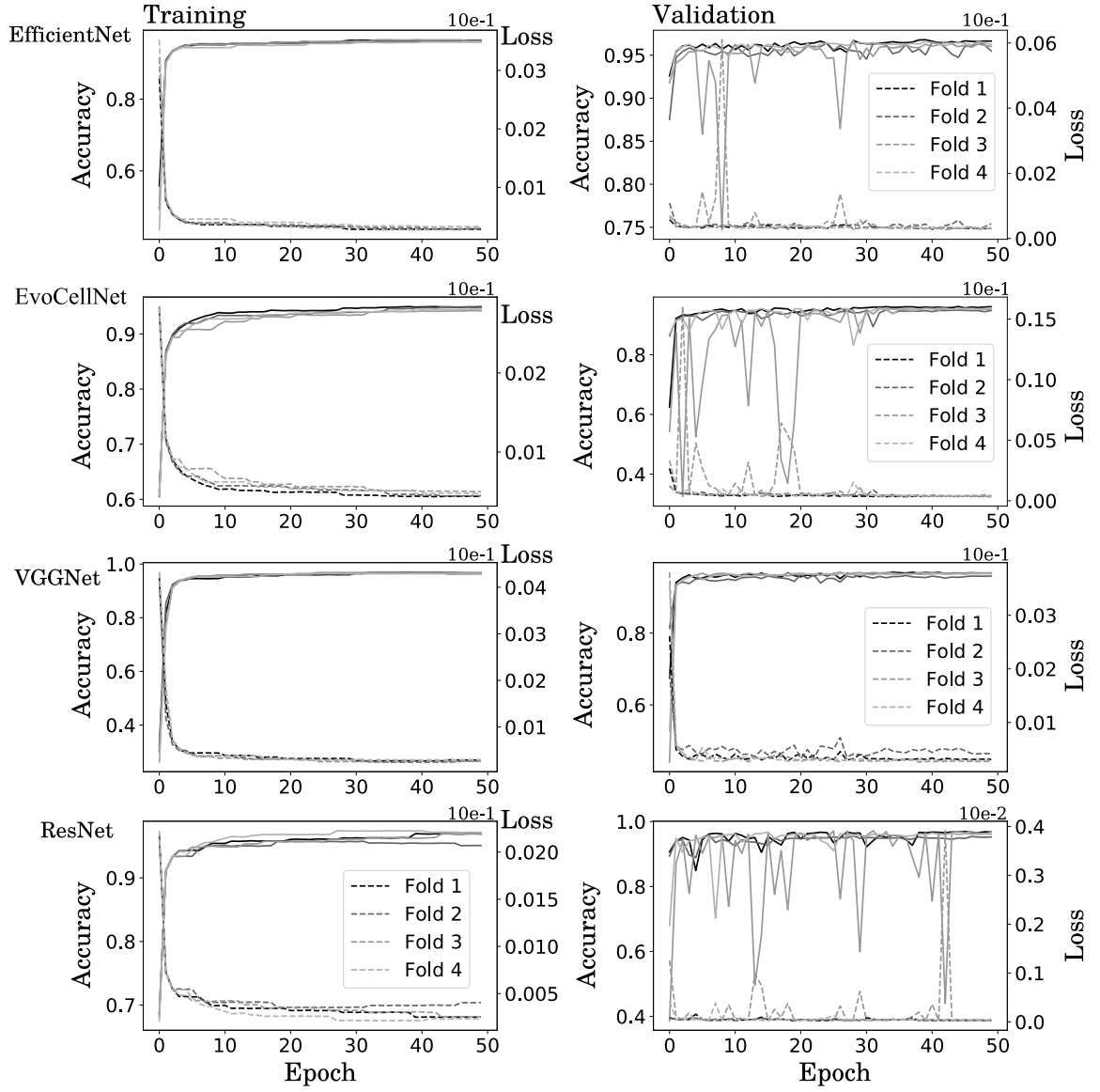


Figure C2: **Training and validation metrics for loss and accuracy.** Training curves of accuracy (left: left y-axis) and loss (right: right y-axis) are shown with epoch number (0-50) for all four folds of training. Similarly, validation curves are shown on the right.

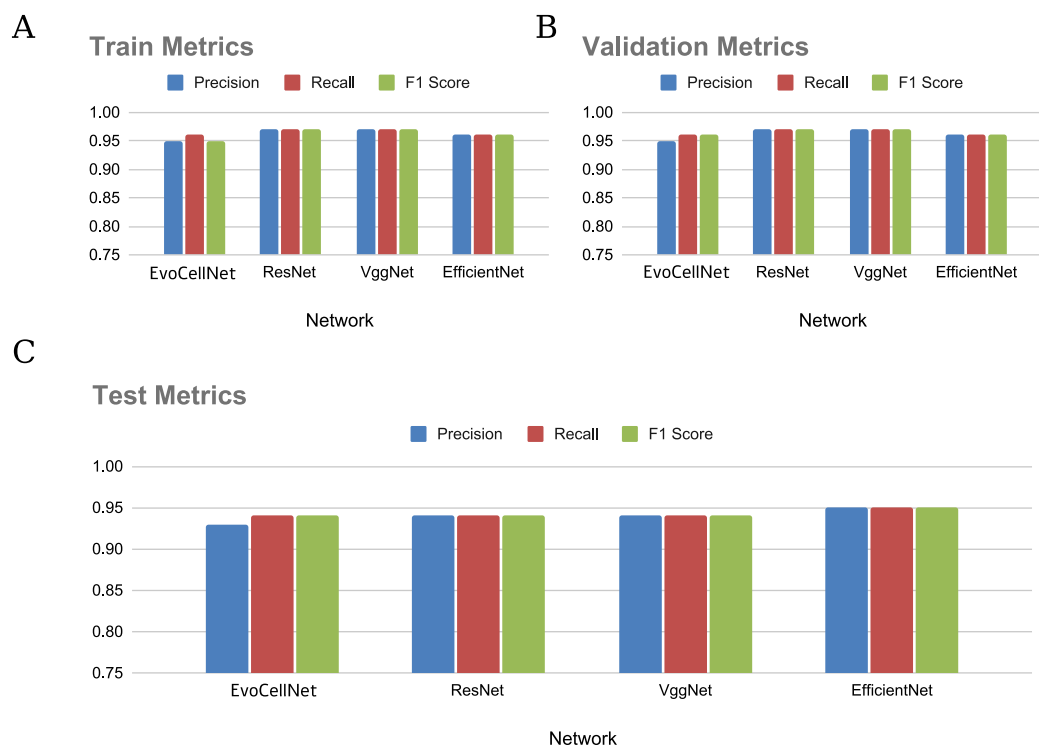


Figure C3: **Performance comparison of the trained models in terms of precision, recall and F1 score.** Bar plot of precision, recall and F1 score for the four trained networks: CustomNet, ResNet, VggNet and EfficientNet. The bar plots are shown for the (A) training (B) validation and (C) test Dataset.

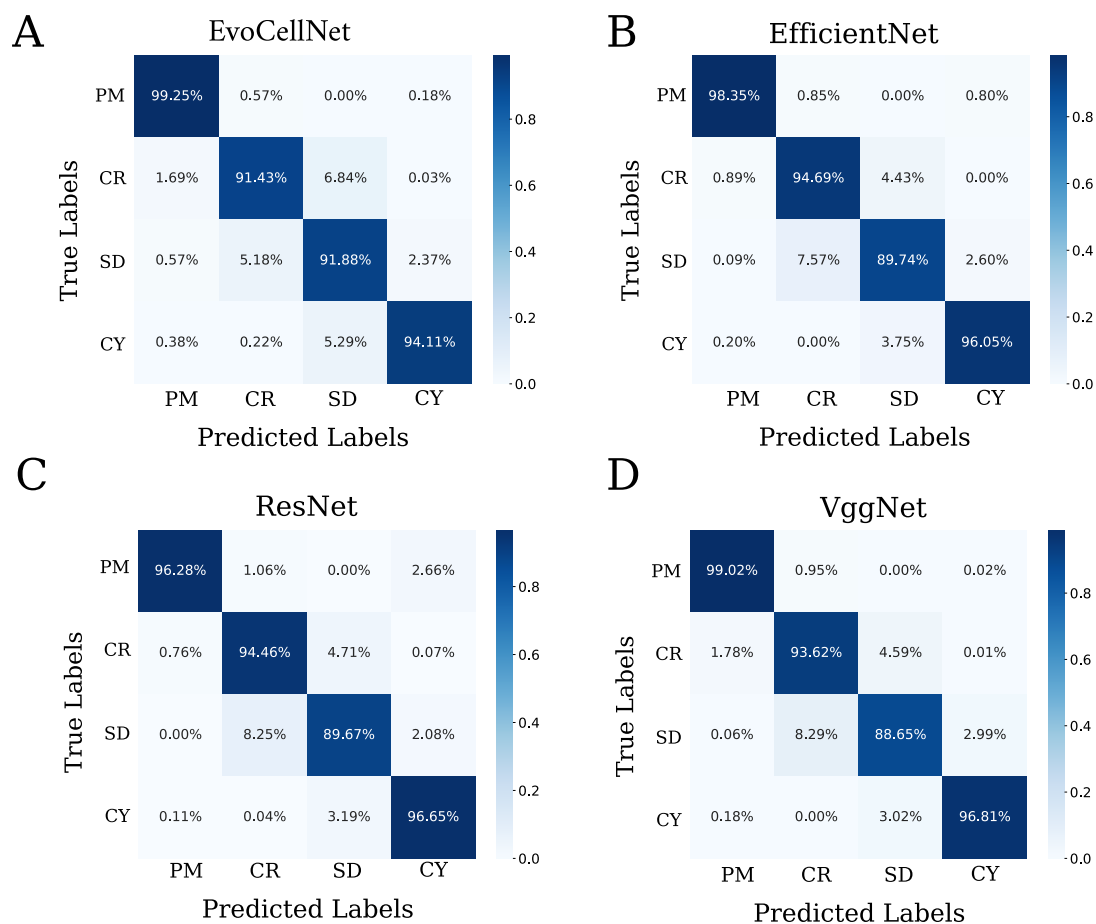


Figure C4: **Confusion matrix for the test dataset.** Confusion matrix with true labels on the y-axis and predicted labels on the x-axis are shown for each model (A) EvoCellNet, (B) EfficientNet, (C) ResNet, (D) VggNet. The matrices are an calculated from an ensemble of four models trained in each cross validation split.

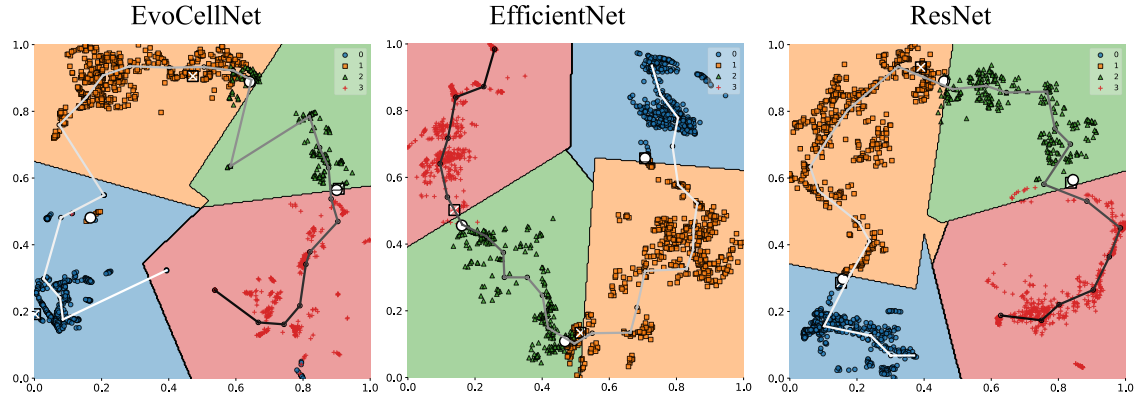
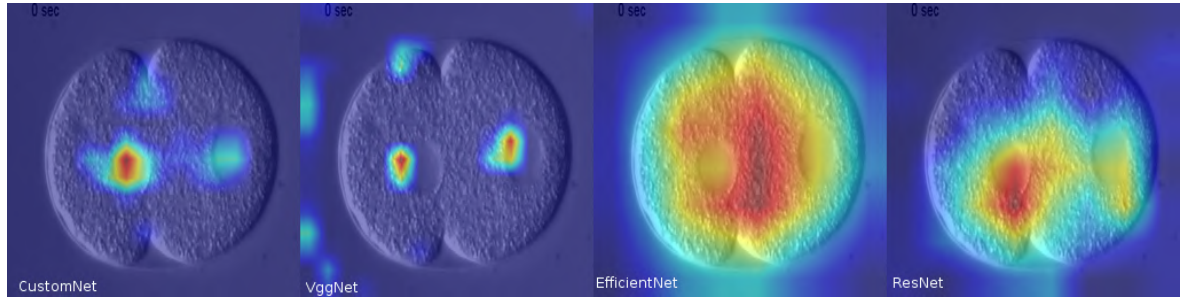


Figure C5: **Cell stage boundaries combined with radial density trace for EvoCellNet, ResNet and EfficientNet in UMAP space.** Normalized UMAP representations in (x-axis) UMAP1 and (y-axis) UMAP2 space split into sectors corresponding to PM, CR, SD and Cy is shown. Boundaries were determined using SVMs (as described in the Methods section) and overlaid with radial density traces, averaged every 12° . The radial traces were connected geometrically with 0° corresponding to the first frame and the direction set by the transition boundary between pronuclear migration (PM) and cytokinesis (Cy) for *C. elegans*. $num_{sectors}$ used for this overlay is 30. Color gradient: start: white, end: black, x: ground truth of the annotated checkpoint, o: checkpoint from SVM.



Video SVC1: **Grad-CAM comparison between four models.** Representative filter activation from the 4 models superimposed on the time-series of a *C. elegans* embryo. Activation values are obtained from the last convolutional layer of each model. Colors indicate the activation strength (blue: low, red: high). The models were selected on the basis of lowest validation loss across the four folds of cross-validation.

Appendix D

Supporting material to Chapter 5

D.1 Figures

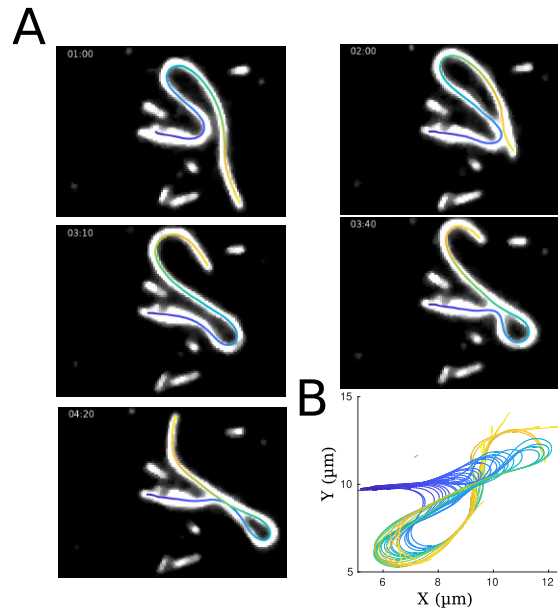


Figure D1: **Representative time-series of filament with tracked contours.** (A) The fluorescence time-series of bent and crossing over filaments are overlaid contour segmented using proposed pipeline. The skeleton is color coded along its length (form tip to end). Time: mm:ss. (B) The filament contours from successive frames are projected showing the traveling wave of curvature.

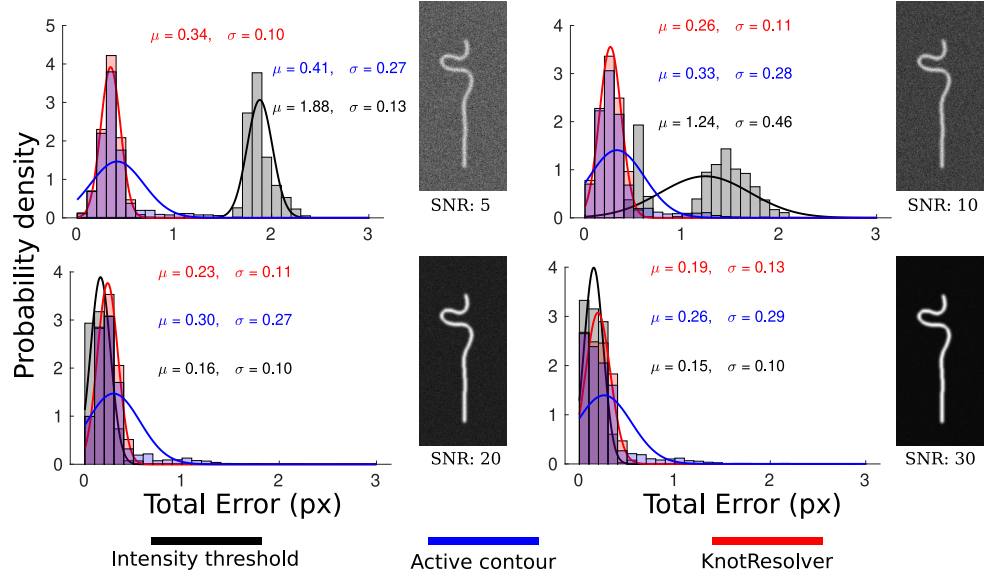
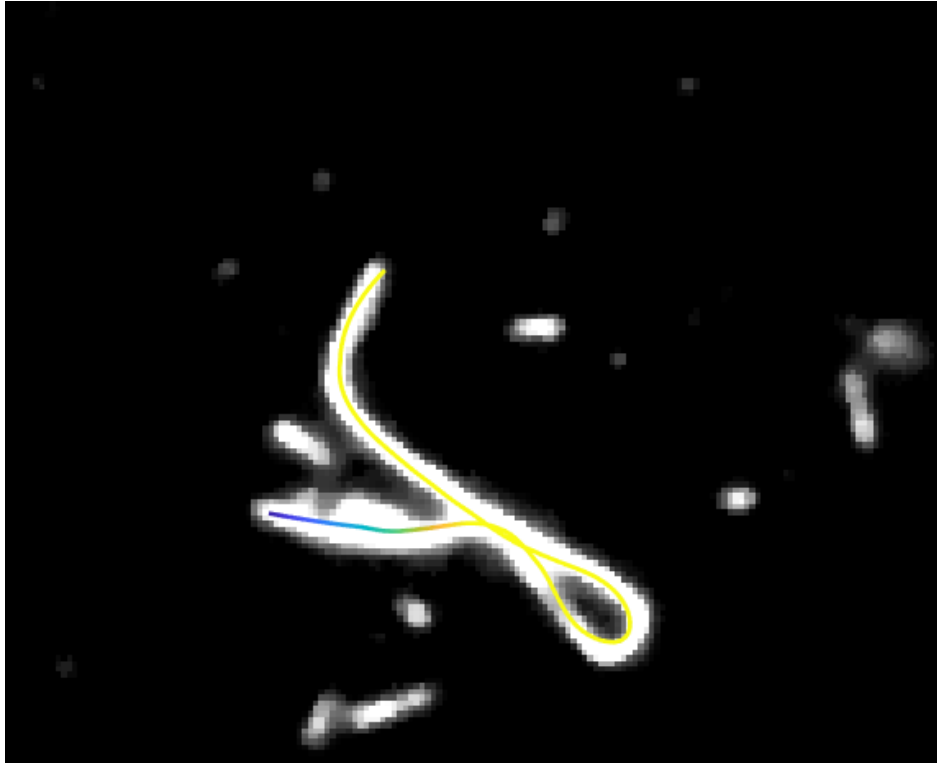
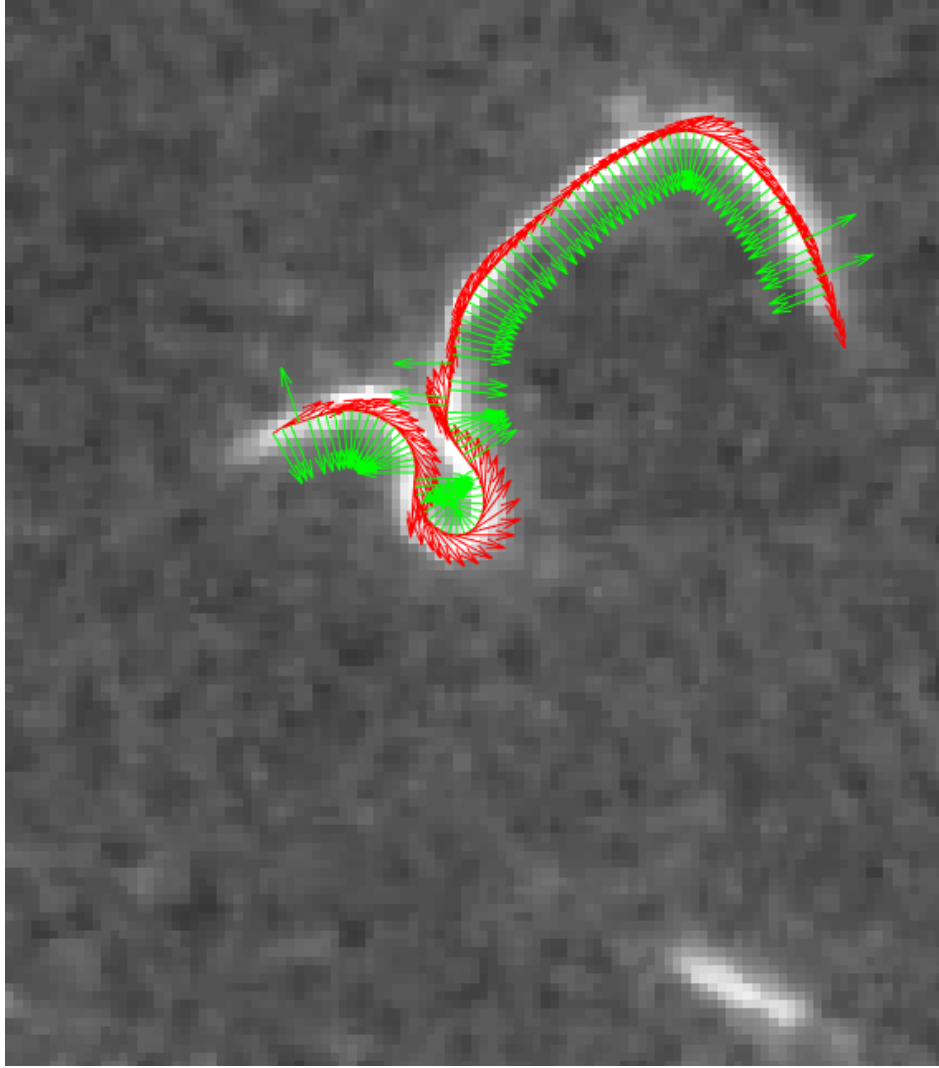


Figure D2: **Simulated filament images with increasing SNR used to measure position detection error comparing segmentation methods.** Segmentation accuracy of compared methods at different signal-to-noise ratios (SNR). Simulated filaments were segmented at varying SNR levels using three different methods: the Threshold method (black), the Active contour method (blue), and our proposed method (red). The total error in pixels for each segmented pixel with respect to the ground truth is plotted as a histogram of probability density function. The inset displays the average and sigma of a Gaussian fit to each histogram ($n = 2880$). A representative image of an input filament at the specified SNR is shown adjacent to each histogram for visual reference.

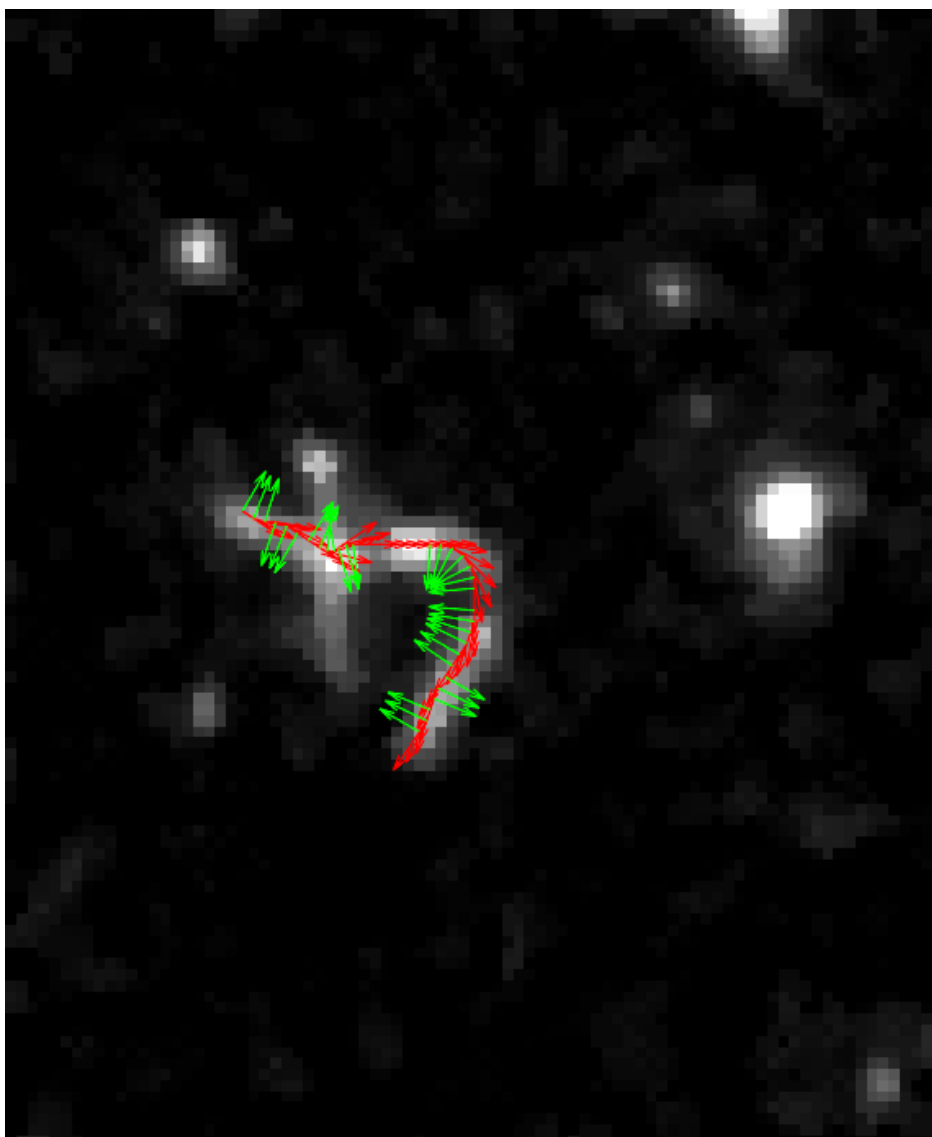
D.1.1 Supporting videos



Video SVD1: **Representative time-series of filament knot resolution with tracked contour.** The fluorescence microscopy time-series of a filament undergoing self-intersection (gray) overlaid with the segmented and tracked contours. KnotResolver parameters used to automatically track the series without manual intervention were: intensity threshold = 0.46, contour iteration = 10, contraction bias = 0.4, smooth factor = 0.2. Scalebar: $2\ \mu m$, interval between frames (δt): 10 s.



Video SVD2: **Representative time-series of filament knot resolution with Frenet–Serret vectors overlaid.** The fluorescence time-series of bent and crossing-over filaments are overlaid with tangential vectors calculated along the segmented skeleton using proposed pipeline. KnotResolver parameters used for segmentation: intensity threshold: 0.5, contour iteration: 10, contraction bias: 0.4, smooth factor: 0.2, manual correction: frames (17, 67, 69). Scale: $2\ \mu\text{m}$, $\Delta t = 10\ \text{s}$.



Video SVD3: **Representative time-series of filament knot resolution with multiple intersecting filaments.** The fluorescence time-series of filament showing sustained oscillations. A second crossing filament can be seen in the background which does not affects the segmentation of the main filament. KnotResolver parameters used for segmentation: intensity threshold: 0.3, contour iteration: 50, contraction bias: 0.25, smooth factor: 0.5, manual correction: None. Scale: $2\ \mu m$, $\Delta t = 10\ s$.

Appendix E

Bacterial cell morphology analysis¹

E.1 Abstract

The regulation of scaling relationship between surface area and volume in bacteria is a fundamental question in cellular physiology. Quantifying it requires accurate measurement of cell dimensions and its change during growth and division. Here, we present a baseline MATLAB segmentation pipeline for quantitative analysis and tracking of bacterial cell shape and morphology. This pipeline combines basic image analysis algorithms to extract precise measurements of cell length, width, and other morphological parameters from microscopy images. The pipeline can be fine-tuned for various imaging modalities such as phase contrast, fluorescence, and DIC microscopy.

E.2 Introduction

The size and shape of bacterial cells are fundamental characteristics that influence various aspects of their physiology and adaptation [299]. Quantifying and understanding how cell morphology scales with increasing size is crucial for understanding the mechanisms governing growth, division, and nutrient uptake [260]. *Escherichia coli*, a model

¹Primary data acquisition by Tanvi Kale

rod-shaped bacterium, has been extensively studied as a model system to study the homeostatic systems regulating its stereotypical size and shape [31, 130, 205, 260].

In this work we developed a flexible tool that combines basic image processing algorithms to precisely quantify and analyze bacterial cell morphology. This quantitative data generated using this tool was used to study scaling relationship between surface area and volume in the model organism *Escherichia coli* [130]. The tool was also used for tracking individual cells through successive frames and to reconstruct growth trajectories and division events, allowing for the investigation of morphological changes throughout the cell cycle [130]. The tracking module was imported from DICOT as discussed in chapter 2.2. The major limitation of this tool is poor separation of cell when they are close to each other. A possible solution is to combine cellpose annotation as it is robust in separating closely placed cells with our output [251].

E.3 Methods

In order to detect cell outlines, we developed an analysis pipeline in MATLAB (Mathworks Inc. Natick, MA, USA) using functions provided through the Image Processing Toolbox. The function names in *italics* are MATLAB functions. Sample image from each microscopy modality used is shown in Figure.D1(A). The workflow is as follows:

1. **Image preprocessing** : All input images were intensity adjusted using *imadjust* before segmentation.
2. **Seed Identification** : Cell outlines were identified by a combination of edge-detection by the Canny method [32] and gap filling (*imfill*). This binary output formed our initial estimate of the cell position. A representative output is shown in Figure.D1(B).
3. **Active contour optimization** : Cells were filtered based on a minimum so-

lidity criterion between 0.3 and 1 to remove small and irregular objects. Then the cell boundaries were optimized using an edge-based active contour method (*activecontour*) [134], with a smoothing factor 1 and a contraction bias of -0.1 . The alternative "edge" optimization is "Chan-vese" which is based on an energy model [36]. The switch between these two can be made by the user.

4. **Manual filtering** : Optimized contour boundaries were overlaid on the original input image and artifacts were interactively deselected to avoid cell debris and out-of-focus cells that have escaped our filtering criterion.
5. **Cell morphometry** : The remaining cells with contours were analyzed for length and width by first detecting the midline. The midline is found by vertically aligning the cell major axis and then iteratively finding the midpoint along the vertical axis. This approach results in sub-optimal results when the cell is long and not oriented at right angles. The widths are then taken by averaging the 3rd and 4th quartiles of the orthogonal distance between the midlines and the contour boundaries. By taking only the top 50% values, we exclude narrow widths at the ends of the cell from the analysis. The binary skeleton of a sample *E. coli* cell with mid-line (red) and widths (green) overlaid is shown in Figure.D1(D).

The MATLAB code was used in Kale et. al 2023 [130] with minor modifications. The code is available as open source and can be downloaded from (here)

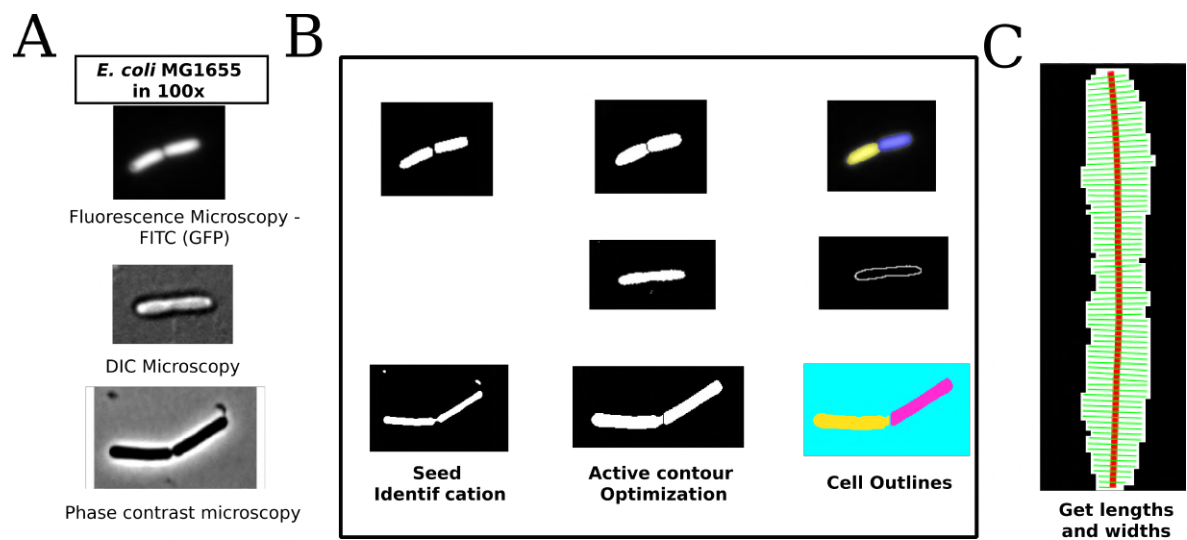


Figure D1: *E.coli* cell analysis pipeline for DIC, phase and fluorescence images. (A) Example images from Fluorescence, DIC, and Phase-Contrast microscopy. (B) Workflow of segmentation steps starting from the left, seed identification using Otsu's threshold [207], followed by active contour-based optimization using the intensity threshold image as the initial seed. The final outlines are shown in the last montage. Neighboring cells are removed either based on the convexity of the contour or depression in the boundary intensity level. (C) Shows the final outline contour with the midline in red and the perpendicular lines from the midline intersecting the boundary as green lines.

Appendix F

RBC cytoadhesion detection in microfluidics chamber¹

F.1 Abstract

Malaria is a significant global health threat. Understanding of the parasite’s complex life cycle is therefore, necessary to develop effective interventions. The infection of red blood cells (RBCs) is a crucial aspect of the parasite’s infection cycle. However, detecting such an interaction in blood vessel environments and monitoring any variation following vaccine treatment is difficult. To address this issue, a novel strategy combining antigenic peptides and microfluidic chambers was proposed [60]. In the study, infected RBC (iRBC) can be observed to make contact with epithelial cells lining the microfluidic chamber which simulates real blood vessel conditions. To facilitate the measurement of iRBC contact and potentially track these occurrences, we introduce a Deep Learning method to detect all RBC flowing within the microfluidic chamber. This method was used to accurately quantifies RBC adhesion dynamics within microfluidic chambers, demonstrating changes in cytoadhesion frequency with different antibody

¹Primary data acquisition and study design by Bhagyashree Deshmukh, Plasmodium Epigenetics Lab, IISER Pune

treatment [60]. This tool can be used for studying cytoadhesion mechanisms which could improve our understanding of infection dynamics, contributing to the development of effective malaria vaccines.

F.2 Introduction

Malaria remains a significant public health concern in developing nations. During the blood stage, merozoites attach to erythrocytes and express virulence proteins which facilitates cytoadhesion, and sequestration of infected RBCs (iRBC) in vital organs [52, 204]. Antibodies against these proteins exhibit neutralizing effects, and have tremendous potential as vaccine candidates [42].

Traditional static assays have been instrumental in evaluating the efficacy of drugs, antibodies, and pharmacological agents in inhibiting the cytoadhesion of RBCs to target cells [8]. However, these assays lack the ability to replicate the complex *in-vivo* conditions that influence cell-to-cell interactions [16, 57]. This is crucial as dynamic flow conditions in blood vessels result in varying iRBC (infected RBC) behaviors, such as rolling motion, tumbling, or firm adhesion, depending on receptor interactions which has important implications on infection efficacy [57]. To mimic these physiological conditions, our collaborator combined an organ-on-chip system with a microfluidic device [60]. In this system primed HBEC-5i (cerebral microvascular endothelial cell [222]) cells were seeded onto the capillary, and the parasite culture (diluted to 0.15% hematocrit) was passed through the capillary at 25 mbar pressure. These conditions resemble deep capillary beds and in our initial observation iRBCs indeed exhibited a slow-rolling phenotype and firm cytoadhesion, while healthy RBCs did not interact with primed HBEC-5i cells. However, manual detection of these events proved to be difficult with very long time datasets and fast-flowing background noise. Therefore, we aim to develop an automatic detection system for quantifying cytoadhesion dynamics

within microfluidic chambers which could offer insights into the infection dynamics of malaria and aid in the study of vaccine efficacy.

F.3 Methods

To quantify cytoadhesion, we created an image analysis pipeline to accurately detect stationary RBCs within the experimental system. Additional details on the experimental setup, antibody development and study design can be found here [60]. An initial exploratory analysis of the video data generated revealed several challenges, these were, significant background noise caused by RBCs flowing in the background and the presence on endothelial cells, and detection inconsistencies due to the changing orientation of RBCs as they flowed in the microfluidic chamber. A representative frame displaying these challenges is shown in Figure F1(A). To address these challenges, we proposed a two-step detection framework. First step involves image filtering with a Laplacian of Gaussian (LoG) filter and a second, post-processing classification of detections into in-focus and background which includes out-of-focus RBCs and epithelial cells.

The filtered RBCs are then tracked and checked if stationary. The approach of first detecting all intensity peaks from the time frames is motivated by apparent intensity peaks which correspond to RBC locations as shown in Figure F1(B). The detection output is overlaid for a separate ROI and is shown in Figure F1(C). The steps are detailed next.

1. **Intensity Peak Detection:** A Laplacian of Gaussian (LoG) filter is utilized to detect intensity peaks indicating potential RBC locations. The filter parameters used were σ , which is the standard deviation set at 1.5 and filter size k of 11.
2. **CNN based classification of "In-Focus" RBCs:** We used a shallow convolutional network to check for "in-focus" RBCs. To create a robust dataset that

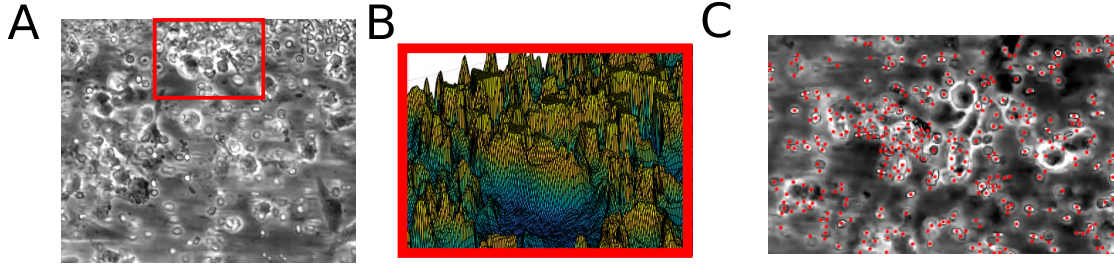


Figure F1: **RBC detection using LoG filter.** (A) Cropped ROI region from a single frame of the time series of the data acquired in the microfluidic chamber. The selected ROI is rescaled and converted to grayscale. (B) Intensity mesh plot from ROI shown in (A) is displayed, and intensity peaks in yellow can be seen corresponding to the RBCs and background, which includes the endothelial cells. (C) Output of blob detection pipeline using LoG filter ($\sigma: 1.5$, $k:11$) is shown as red dots overlaid on a sample ROI taken from the same time series.

represents all possible geometries of RBCs in the microfluidics chamber, we took a noisy student annotation approach [289]. In this approach, we first manually annotated RBCs ($n=100$) and background ($n=2000$) and trained the network to separate them, representative images from the two categories are shown in Figure.F2(A). The image data is a 32×32 window crop around each LoG detection. The network was then used to automatically classify LoG detections into pseudo-labels of RBC and non-RBC categories. The RBC class detections were then manually checked to quickly create a dataset of nearly 2300 accurate detections of in-focus RBCs. The rest of the detections were moved to the background class ($n=15000$). The network was re-trained from scratch using this larger dataset. The network training is shown in Figure.F2(B). We used random class oversampling before network training. The network was trained with a learning rate of 0.001 with the stochastic gradient descent algorithm Figure.F2(C).

3. **Classifier Accuracy Assessment** : The network validation accuracy on the oversampled dataest is 95.75%. We additionally applied a threshold probability where detections with classification probability greater than 0.93 are used for further analysis.

4. **Stationary detection Filtering:** All detections were first filtered based a "Stationary" criterion to refine detected RBCs based on spatial and temporal stability. This is defined as the presence of another detection within a threshold distance of 10 pixels ($\sim$ diameter of RBCs) for 5 consecutive frames. Output of the classification and stationary detection filtering is shown in Figure.F2(D).

Visually, we could confirm that our analysis results in accurate selection of stationary RBCs in microfluidics experiments. The detection method can be further developed to distinguish between different adhesion dynamics such as rolling and tumbling. In the future, we would like to address this limitation and perform a more thorough evaluation of RBC adhesion trajectory.

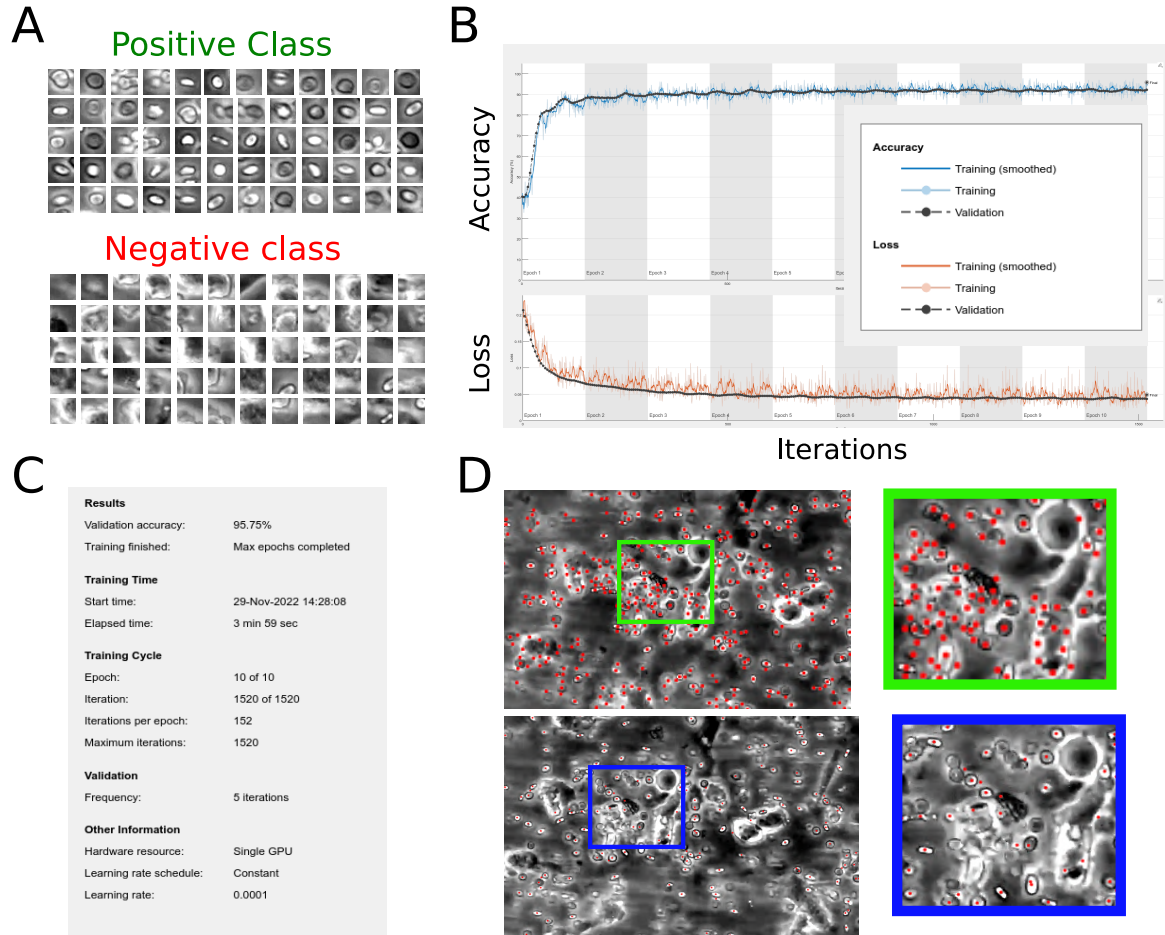


Figure F2: **LoG detection filtering to detect "in-focus" RBCs using CNN.** (A) Example images from the positive class and negative class. Positive class contains LoG detections where RBC images are in focus and clearly visible. Negative class contains LoG detections that belong to the background or out-of-focus RBCs. (B) CNN training for binary classification between positive and negative class. Accuracy (blue) and loss (orange) metrics for the training (solid lines) and validation (dotted dashed lines) split are shown. (C) Final validation accuracy (95.75%) and training hyperparameters are shown. (D) LoG detections are shown in the top montage and post-processed CNN classifications labelled as the positive class are shown in the bottom montage. Cropped ROIs show the zoomed-in regions of the respective ROIs.

Appendix G

Microtubule tip detection and tracking in gliding assays using deep learning

G.1 Abstract

Quantitative information on motor proteins and microtubules can be extracted through *in vitro* gliding assays. However, existing tools often suffer from limitations in detection accuracy and tracking performance, especially in crowded environments where multiple filaments intersect. In this study, we have developed an MT tip detection pipeline for microtubule tip tracking, which can be combined with *KnotResolver* for a thorough analysis of MT gliding mobility. In the tip detection model, we implemented a RetinaNet model architecture, with a ResNet-50 backbone and a feature pyramid network (FPN), for accurate tip detection in gliding assays. In this preliminary work, we showcase the application of this network on a simulated dataset of MT gliding assays. For improved tracking, post-detection linking of microtubule tips and skeleton is performed using a multiple object tracking algorithm. The pipeline demonstrates

robust microtubule tracking performance, particularly in high-density image scenarios. In the future, we aim to fine-tune this baseline network to experimental time series.

G.2 Introduction

Microtubules (MTs) play a crucial role in various cellular processes, and their dynamic behavior is of significant interest in cell biology research. Quantitative analysis of MT dynamics, particularly through in vitro gliding assays, offers valuable insights into the underlying molecular mechanisms governing cellular functions. Several tools for MT tracking and analysis exist [208].

TSOAX, introduced by Xu et al., is a notable open-source software tool designed for tracking biopolymer growth and network deformation from 2D and 3D time-lapse microscopy images [292]. TSOAX extends the capabilities of the SOAX software to handle time-lapse movies, employing stretching open active contours and a combined local and global graph matching framework to establish temporal correspondence between growing and intersecting filaments across frames [293]. This allows tracking of intersecting filaments. This tool has been utilized to track individual filaments or network branches. FiberApp, developed by Usov and Mezzenga, is another open-source software tool designed for tracking and analyzing polymers, filaments, biomacromolecules, and fibrous objects from microscopy images [271]. FiberApp provides automated detection and tracking of individual fibers using adaptive thresholding and skeletonization. This tool supports input from various microscopy modalities and has been applied to study the assembly and dynamics of diverse fibrous systems. Mosby et al. introduced a Python-based automated tracking routine for myosin II filaments, enabling tracking of individual filaments in fluorescence microscopy images of in vitro motility assays [193]. This routine utilizes similar image processing techniques to FiberApp. The software outputs a range of metrics for each tracked filament, facilitating high-throughput anal-

ysis of myosin II filament behavior. Ruhnnow et al. developed an algorithm for tracking single particles and elongated filaments with nanometer precision, improving the detection accuracy of MTs [232]. This algorithm precisely localizes the centerlines and tips of curved filaments with nanometer precision using multiple Gaussian models. Recently, Kapoor et al. introduced MTrack, a software tool for automated detection, tracking, and analysis of dynamic microtubules with high precision [132]. This tool automates the quantification of microtubule seeds by using Maximally Stable Extremal regions (MSER) algorithm, to each identified seed a sum of 2D Gaussians is fit to determine seed location with subpixel accuracy. Additionally, 2D Gaussian polynomial models are employed to track growing filaments that is robust to filament intersections.

Overall, these software tools contribute significantly to advancing our understanding of biological processes by automating the analysis of filamentous structures in microscopy images, enabling high-throughput and quantitative studies of cytoskeletal dynamics and other related phenomena. However, existing tools for MT tip detection and tracking often exhibit limitations in accuracy and performance, especially in crowded environments where multiple filaments intersect, presenting challenges for reliable analysis. A key limitation is either stringent criterion of filament geometry which assumes straight geometry or inability to track highly curved MT which are observed in *in vivo* scenarios [23]. We have already made significant advances in tracking self-intersecting and buckled filaments in Chapter 5. We complement *KnotResolver* with an automated tip detection framework that can be used to improve tip localization and tracking of microtubules in complex environments.

G.3 Methods

G.3.1 A gliding assay database

We construct a gliding assay database using Cytosim [199] and MATLAB [191]. The simulation parameters are detailed in Table.G.1. The MATLAB script *GenerateAnnotation.m* was used to simulate gliding assays multiple times, in a loop for n iterations. During each iteration, it executes the command `./sim GlidingAssay.cym` to simulate the gliding assay using cytosim and saves the output using the `./play movie object` command.

Parameter	Value	Parameter	Value
time_step	0.1	viscosity	1
precondition	0	display	zoom=1.55, size=512
shape	periodic	length	10 x 10
rigidity	0.25	segmentation	0.02
display	line_width=2	binding_rate	10
binding_range	0.01	unbinding_rate	0.3
unbinding_force	5	activity	move
unloaded_speed	2	stall_force	6
display	visible=0	stiffness	100

Table G.1: **Gliding assay simulation parameters.**

1. **Extracting tip and contour coordinates from simulation output:** After each simulation, the script retrieves all created *ppm* files and processes them. For each *ppm* file (image output of Cytosim), we extract the green channel, remove any non-green pixels, and convert the remaining image into a binary format. From the binary image filament coordinates are extracted and saved along with the frame ID and simulation ID in a structured array which is saved as a *.csv* file.
2. **Simulating experimental gliding assays:** To create gliding assays similar to experimental datasets, we first dilate the filament contour with a diamond struc-

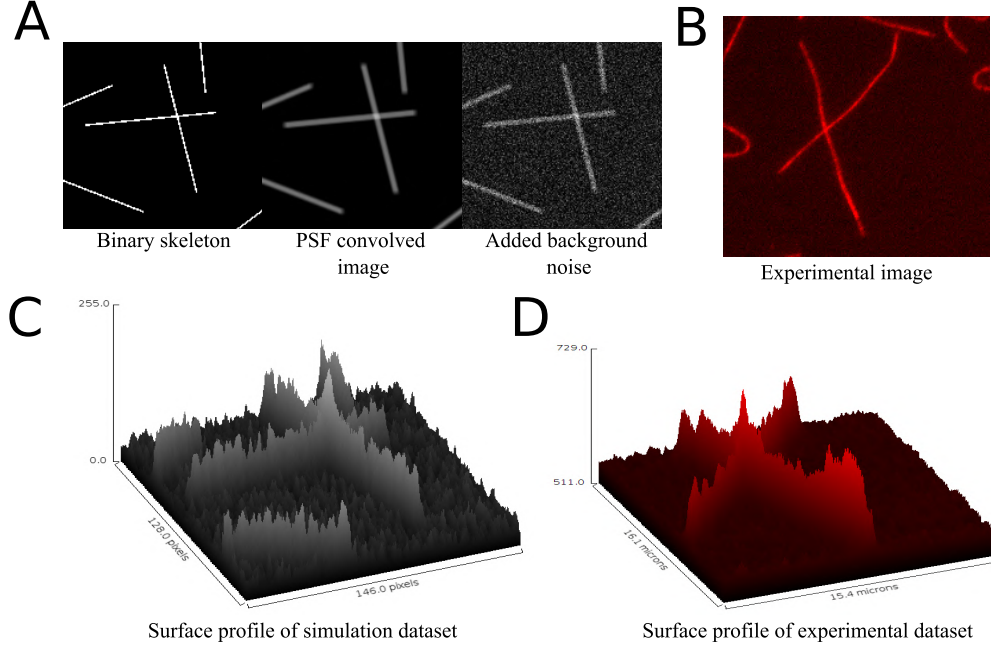


Figure G1: **Simulating experimental time series using cytosim.** (A) Show image processing from left to right: binary skeleton output from the cytosim simulation, PSF convolved image and final image with background noise. (B) A representative ROI from experimental datasets where MTs are labelled with fluorescent markers. (C)-(D) Surface plot showing intensity profile in 3D for simulated dataset in (C) and experimental dataset in (D).

ture element of pixel size 2, Figure.G1(A). This can be increased to create thicker MTs. To the dilated contour, we convolve the filament with a Gaussian kernel to replicate the point spread function (PSF) of a fluorescent molecule. Finally, background shot noise is added using a Poisson distribution [121] , Figure.G1(A). Additionally, the end coordinates of the filaments are saved to create a ground truth dataset. The ouput image surface profile is compared to a experimental time series in FigureG1(C-D).

G.3.2 Filament tip detection network training

The detection network implements a RetinaNet model for tip detection in images [157]. The model backbone is based on ResNet-50 architecture with a feature pyramid network (FPN) for multi-scale feature extraction [117].

1. **Network training:** The network is trained using stochastic gradient descent (SGD) optimizer on a COCO-style dataset prepared from tip annotations as detailed in the previous section. The learning rate (lr) was set to 0.005, with $momentum=0.9$ and $weight\ decay$ of $=0.0005$. The dataset is split into training and validation sets to monitor the model’s performance during training. Pretrained weights of *resnet50_fpn* from the *torchvision* library were used [174].
2. **Data augmentation:** Input images were augmented using random horizontal flips with a probability of 5. Network was trained for 5 epochs with *batch_size* of 2 using Nvidia P1000 GPU.
3. **Tip tracking:** The output coordinates predicted by the network are tracked using *CentroidTracker* provided by *motrackers* with default parameters [61].

In future iterations of this study, several enhancements could be considered to improve the effectiveness of the developed pipeline. Firstly, fine tuning the model on experimental time series and establishing baseline metrics will potentially improve the performance of this network. Different baseline network architecture can be explored or advanced feature extraction techniques can be implemented [253]. Additionally, data augmentation strategies tailored specifically for microtubule tip detection could help improve the model’s generalization ability and robustness to variations in input images. Furthermore, integrating additional features into the tracking process, such as geometric priors or physical parameters, may lead to more accurate results, particularly for experimental time series data.

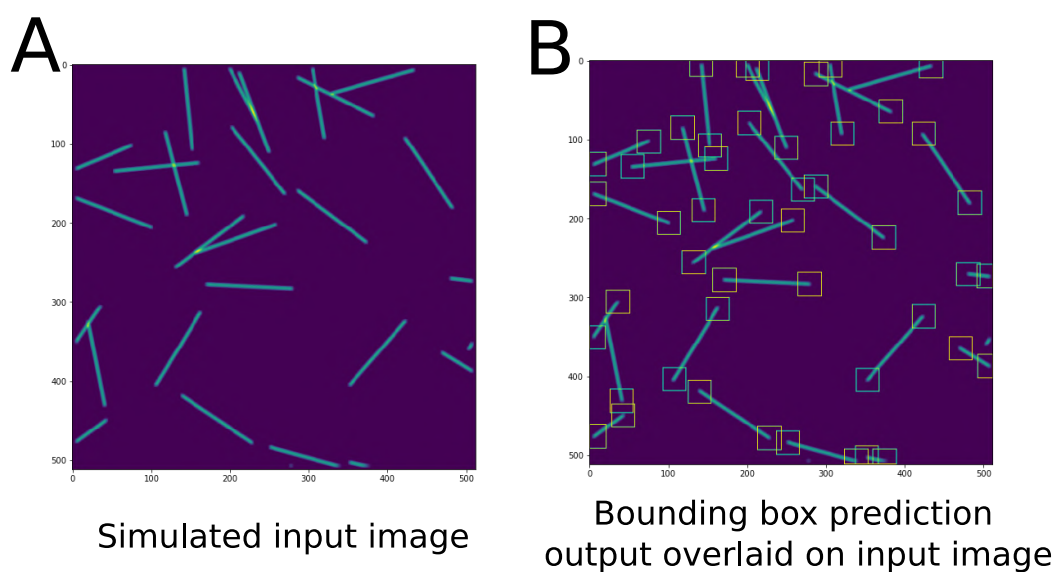


Figure G2: **Tip detection inference on simulated dataset.** **(A)** Image shows the input for the tip detection network. The image was converted to a tensor object before input. **(B)** Image in (A) is overlaid with predicted tip bounding box detected with a >0.95 confidence.

Publications

Peer reviewed publications

- Chaphalkar, Anushree R., Yash K. Jawale, **Khatri, Dhruv**, and Chaitanya A. Athale. "Quantifying intracellular particle flows by DIC object tracking." Biophysical Journal 120, no. 3 (2021): 393-401.
- **Dhruv Khatri**, Thibault Brugière, Chaitanya A. Athale, and Marie Delattre. "Evolutionary divergence of anaphase spindle mechanics in nematode embryos constrained by antagonistic pulling and viscous forces." Molecular Biology of the Cell 33, no. 6 (2022): ar61.
- Kale, Tanvi, **Khatri, Dhruv**, and Chaitanya A. Athale. "Allometry of escherichia coli surface area with volume: effect of size variability, filamentation and division dynamics." Physical Biology (2023).
- Wave-like oscillations of clamped microtubules driven by collective dynein transport. SA Yadav, D Khatri, A Soni, N Khetan, CA Athale. Biophysical Journal 123 (4), 509-524
- **Dhruv Khatri**, Tanvi Kale, Jashaswi Basu, Shivani A. Yadav, Chaitanya A. Athale. "Quantification of cell shape, intracellular flows and transport based on DIC object detection and tracking." Journal of Microscopy (2024).

In preparation & preprints

- **Khatri, Dhruv**, Shivani A. Yadav, and Chaitanya A. Athale. "KnotResolver: Tracking self-intersecting filaments in microscopy using directed graphs." arXiv

preprint arXiv:2404.12029 (2024).

- **Dhruv Khatri**, Chaitanya A. Athale. "Deep learning based cell stage detection of first embryonic division across evolutionarily diverse nematodes"

Conference papers

- **Dhruv Khatri** and Chaitanya A Athale. "Deep learning for cell stage classification during first embryonic division across diverse nematodes". Cell Bio Virtual 2021 Meeting
- **Dhruv Khatri**, Shivani A. Yadav, Prince Sah and Chaitanya A. Athale. "Tracking self-intersecting filaments using a graph based approach". The Microtubules, Motors, Transport and Trafficking (M2T2 2023)
- **Dhruv Khatri**, Shivani A. Yadav, Prince Sah and Chaitanya A. Athale. "Contour tracing of single filaments in pinned gliding assays". Imaging Cell Dynamics 2023 (*Online attendance*).

Bibliography

- [1] Ernst Abbe. Beiträge zur theorie des mikroskops und der mikroskopischen wahrnehmung. *Archiv für mikroskopische Anatomie*, 9(1):413–468, 1873.
- [2] Abien Fred Agarap. Deep learning using rectified linear units (relu). *arXiv preprint arXiv:1803.08375*, 2018.
- [3] Omar Al-Kofahi, Richard J Radke, Susan K Goderie, Qin Shen, Sally Temple, and Badrinath Roysam. Automated cell lineage construction: a rapid method to analyze clonal development established with murine neural progenitor cells. *Cell cycle*, 5(3):327–335, 2006.
- [4] Stephen P Alexander and Conly L Rieder. Chromosome motion during attachment to the vertebrate spindle: initial saltatory-like behavior of chromosomes and quantitative analysis of force production by nascent kinetochore fibers. *The Journal of cell biology*, 113(4):805–815, 1991.
- [5] Maria Almonacid, Marie-Émilie Terret, and Marie-Hélène Verlhac. Actin-based spindle positioning: new insights from female gametes. *Journal of cell science*, 127(3):477–483, 2014.
- [6] María Isabel Arjona, Javad Najafi, and Nicolas Minc. Cytoplasm mechanics and cellular organization. *Current Opinion in Cell Biology*, 85:102278, 2023.
- [7] Lucas Assirati, Núbia Rosa da Silva, Lilian Berton, Alneu de A Lopes, and

- Odemir M Bruno. Performing edge detection by difference of gaussians using q-gaussian kernels. In *Journal of Physics: Conference Series*, volume 490, page 012020. IOP Publishing, 2014.
- [8] Yvonne Azasi. Assay of static adhesion of plasmodium falciparum-infected erythrocytes to cells, including inhibition of the adhesion. In *Malaria Immunology: Targeting the Surface of Infected Erythrocytes*, pages 515–525. Springer, 2022.
- [9] Shoubhik Chandan Banerjee, Khursheed Ahmad Khan, and Rati Sharma. Deep-worm-tracker: Deep learning methods for accurate detection and tracking for behavioral studies in c. elegans. *Applied Animal Behaviour Science*, 266:106024, 2023.
- [10] Nawaf Hazim Barnouti et al. Improve face recognition rate using different image pre-processing techniques. *American Journal of Engineering Research (AJER)*, 5(4):46–53, 2016.
- [11] Kathleen Bates, Kim N Le, and Hang Lu. Deep learning for robust and flexible tracking in behavioral studies for c. elegans. *PLOS Computational Biology*, 18(4):e1009942, 2022.
- [12] Nathalie Beaujean, Juliette Salvaing, Nur Annies Abd Hadi, and Sari Pennings. Antibody-based detection of global nuclear dna methylation in cells, tissue sections, and mammalian embryos. *Methods Mol Biol*, 1708:59–80, 2018.
- [13] Maria L. Begasse, Mark Leaver, Federico Vazquez, Stephan W. Grill, and Anthony A. Hyman. Temperature dependence of cell division timing accounts for a shift in the thermal limits of c.elegans and C.briggsae. *Cell Reports*, 10(5):647–653, feb 2015.
- [14] Howard C Berg. *Random walks in biology*. Princeton University Press, 1993.

- [15] Stuart Berg, Dominik Kutra, Thorben Kroeger, Christoph N Straehle, Bernhard X Kausler, Carsten Haubold, Martin Schiegg, Janez Ales, Thorsten Beier, Markus Rudy, et al. Ilastik: interactive machine learning for (bio) image analysis. *Nature methods*, 16(12):1226–1232, 2019.
- [16] Maria Bernabeu, Caitlin Howard, Ying Zheng, and Joseph D Smith. Bioengineered 3d microvessels for investigating plasmodium falciparum pathogenesis. *Trends in parasitology*, 37(5):401–413, 2021.
- [17] J-F Berret. Local viscoelasticity of living cells measured by rotational magnetic spectroscopy. *Nature communications*, 7(1):10134, 2016.
- [18] Alex Bewley, Zongyuan Ge, Lionel Ott, Fabio Ramos, and Ben Upcroft. Simple online and realtime tracking. In *2016 IEEE international conference on image processing (ICIP)*, pages 3464–3468. IEEE, 2016.
- [19] P. Birch, Bhargav K. Mitra, Nagachetan Bangalore, S. Rehman, R. Young, and C. Chatwin. Approximate bandpass and frequency response models of the difference of gaussian filter. *Optics Communications*, 283:4942–4948, 2010.
- [20] Simon Blanchoud, Yemima Budirahardja, Félix Naef, and Pierre Gönczy. Asset: a robust algorithm for the automated segmentation and standardization of early caenorhabditis elegans embryos. *Developmental Dynamics*, 239(12):3285–3296, 2010.
- [21] Simon Blanchoud, Yemima Budirahardja, Félix Naef, and Pierre Gönczy. Asset: a robust algorithm for the automated segmentation and standardization of early caenorhabditis elegans embryos. *Dev Dyn*, 239(12):3285–96, Dec 2010.
- [22] Léon Bottou. Large-scale machine learning with stochastic gradient descent. In *Proceedings of COMPSTAT’2010*, pages 177–186. Springer, 2010.

- [23] Clifford P Brangwynne, Frederick C MacKintosh, Sanjay Kumar, Nicholas A Geisse, Jennifer Talbot, L Mahadevan, Kevin K Parker, Donald E Ingber, and David A Weitz. Microtubules can bear enhanced compressive loads in living cells because of lateral reinforcement. *The Journal of cell biology*, 173(5):733–741, 2006.
- [24] Michael Brauchle, Karin Kiontke, Philip MacMenamin, David HA Fitch, and Fabio Piano. Evolution of early embryogenesis in rhabditid nematodes. *Developmental Biology*, 335(1):253–262, 2009.
- [25] Sydney Brenner. The genetics of *caenorhabditis elegans*. *Genetics*, 77(1):71–94, 1974.
- [26] David Breuer and Zoran Nikoloski. Define: an optimisation-based method for robust disentangling of filamentous networks. *Scientific reports*, 5(1):18267, 2015.
- [27] GW Brodland and R Gordon. Intermediate filaments may prevent buckling of compressively loaded microtubules. 1990.
- [28] Itay Budin, Tristan de Rond, Yan Chen, Leanne Jade G Chan, Christopher J Petzold, and Jay D Keasling. Viscous control of cellular respiration by membrane lipid composition. *Science*, 362(6419):1186–1189, 2018.
- [29] CR Burch and JPP Stock. Phase-contrast microscopy. *Journal of Scientific Instruments*, 19(5):71, 1942.
- [30] R Buscall, P D’Haene, and Joannes Mewis. Maximum density for flow of dispersions of near monodisperse spherical particles. *Langmuir*, 10(5):1439–1441, 1994.
- [31] Manuel Campos, Ivan V Surovtsev, Setsu Kato, Ahmad Paintdakhi, Bruno Bel-

- tran, Sarah E Ebmeier, and Christine Jacobs-Wagner. A constant size extension drives bacterial cell size homeostasis. *Cell*, 159(6):1433–1446, 2014.
- [32] John Canny. A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence*, (6):679–698, 1986.
- [33] Jianfeng Cao, Guoye Guan, Vincy Wing Sze Ho, Ming-Kin Wong, Lu-Yan Chan, Chao Tang, Zhongying Zhao, and Hong Yan. Establishment of a morphological atlas of the *caenorhabditis elegans* embryo using deep-learning-based 4d segmentation. *Nature communications*, 11(1):6254, 2020.
- [34] Anne E Carpenter, Thouis R Jones, Michael R Lamprecht, Colin Clarke, In Han Kang, Ola Friman, David A Guertin, Joo Han Chang, Robert A Lindquist, Jason Moffat, Polina Golland, and David M Sabatini. CellProfiler : image analysis software for identifying and quantifying cell phenotypes. 2006.
- [35] Brian C Carter, George T Shubeita, and Steven P Gross. Tracking single particles: a user-friendly quantitative evaluation. *Physical biology*, 2(1):60, 2005.
- [36] Tony F Chan and Luminita A Vese. Active contours without edges. *IEEE Transactions on image processing*, 10(2):266–277, 2001.
- [37] Yee-Hung M Chan and Wallace F Marshall. Scaling properties of cell and organelle size. *Organogenesis*, 6(2):88–96, 2010.
- [38] Ruchika Chandel and Gaurav Gupta. Image filtering algorithms and techniques: A review. *International Journal of Advanced Research in Computer Science and Software Engineering*, 3(10), 2013.
- [39] Anushree R Chaphalkar, Yash K Jawale, Dhruv Khatri, and Chaitanya A Athale. Quantifying intracellular particle flows by dic object tracking. *Biophysical Journal*, 120(3):393–401, 2021.

- [40] Michael K Cheezum, William F Walker, and William H Guilford. Quantitative comparison of algorithms for tracking single fluorescent particles. *Biophysical journal*, 81(4):2378–2388, 2001.
- [41] Raghunath Chelakkot, Arvind Gopinath, Lakshminarayanan Mahadevan, and Michael F Hagan. Flagellar dynamics of a connected chain of active, polar, brownian particles. *Journal of The Royal Society Interface*, 11(92):20130884, 2014.
- [42] Qijun Chen, Fredrik Pettersson, Anna M Vogt, Berit Schmidt, Sanjay Ahuja, Peter Liljeström, and Mats Wahlgren. Immunization with pfemp1-dbl1 α generates antibodies that disrupt rosettes and protect against the sequestration of plasmodium falciparum-infected erythrocytes. *Vaccine*, 22(21-22):2701–2712, 2004.
- [43] Xiaowei Chen, Xiaobo Zhou, and Stephen TC Wong. Automated segmentation, classification, and tracking of cancer cell nuclei in time-lapse microscopy. *IEEE transactions on biomedical engineering*, 53(4):762–766, 2006.
- [44] Hui-Jun Cheng, Ching-Hsien Hsu, Che-Lun Hung, and Chun-Yuan Lin. A review for cell and particle tracking on microscopy images using algorithms and deep learning technologies. *biomedical journal*, 45(3):465–471, 2022.
- [45] Marcelo Cicconet, Michelle Gutwein, Kristin C Gunsalus, and Davi Geiger. Label free cell-tracking and division detection based on 2d time-lapse images for lineage analysis of early embryo development. *Computers in biology and medicine*, 51:24–34, 2014.
- [46] George V Clokey and Lewis A Jacobson. The autofluorescent “lipofuscin granules” in the intestinal cells of caenorhabditis elegans are secondary lysosomes. *Mechanisms of ageing and development*, 35(1):79–94, 1986.

- [47] David Cluet, Pierre Nicolas Stéb , Soizic Riche, Martin Spichty, and Marie Delattre. Automated high-throughput quantification of mitotic spindle positioning from DIC movies of caenorhabditis embryos. *PLoS ONE*, 9(4), 2014.
- [48] Guy Barrett Coleman and Harry C Andrews. Image segmentation by clustering. *Proceedings of the IEEE*, 67(5):773–785, 1979.
- [49] Tony J Collins. Imagej for microscopy. *Biotechniques*, 43(S1):S25–S30, 2007.
- [50] Ryan Conrad and Kedar Narayan. Cem500k, a large-scale heterogeneous unlabeled cellular electron microscopy image dataset for deep learning. *Elife*, 10:e65894, 2021.
- [51] Elise A Corbin, Fang Kong, Chwee Teck Lim, William P King, and Rashid Bashir. Biophysical properties of human breast cancer cells measured using silicon mems resonators and atomic force microscopy. *Lab on a Chip*, 15(3):839–847, 2015.
- [52] Alan F Cowman, Julie Healer, Danushka Marapana, and Kevin Marsh. Malaria: biology and disease. *Cell*, 167(3):610–624, 2016.
- [53] Marina E Crowder, Magdalena Strzelecka, Jeremy D Wilbur, Matthew C Good, George Von Dassow, and Rebecca Heald. A comparative analysis of spindle morphometrics across metazoans. *Current Biology*, 25(11):1542–1550, 2015.
- [54] Si n Culley, Alicia Cuber Caballero, Jemima J Burden, and Virginie Uhlmann. Made to measure: An introduction to quantifying microscopy data in the life sciences. *Journal of Microscopy*, 2023.
- [55] Brian R Daniels, Byron C Masi, and Denis Wirtz. Probing single-cell micromechanics in vivo: the microrheology of c. elegans developing embryos. *Biophysical journal*, 90(12):4712–4719, 2006.
- [56] Gaudenz Danuser. Computer vision in cell biology. *Cell*, 147(5):973–978, 2011.

- [57] Anil K Dasanna, Christine Lansche, Michael Lanzer, and Ulrich S Schwarz. Rolling adhesion of schizont stage malaria-infected red blood cells in shear flow. *Biophysical journal*, 112(9):1908–1919, 2017.
- [58] Gabriele De, Eric Lauga, and Raymond E Goldstein. Spontaneous oscillations of elastic filaments induced by molecular motors. *Journal of The Royal Society Interface*, 14(136):20170491, 2017.
- [59] Marie Delattre and Nathan W Goehring. The first steps in the life of a worm: themes and variations in asymmetric division in *c. elegans* and other nematodes. In *Current Topics in Developmental Biology*, volume 144, pages 269–308. Elsevier, 2021.
- [60] Bhagyashree Deshmukh, Dhruv Khatri, Sanjay Kumar Kochar, Chaitanya A Athale, and Krishanpal Karmodiya. Multi-protein chimeric antigens, a novel combined approach for efficiently targeting and blocking the blood stage of *plasmodium falciparum*. *bioRxiv*, pages 2023–11, 2023.
- [61] Aditya M. Deshpande. Multi-object trackers in python. <https://github.com/adipandas/multi-object-tracker>, 2020.
- [62] Damien P Devos, Ralph Gräf, and Mark C Field. Evolution of the nucleus. *Current opinion in cell biology*, 28:8–15, 2014.
- [63] Nameirakpam Dhanachandra, Khumanthem Manglem, and Yambem Jina Chanu. Image segmentation using k-means clustering algorithm and subtractive clustering algorithm. *Procedia Computer Science*, 54:764–771, 2015.
- [64] Andrew Dillin, Ao-Lin Hsu, Nuno Arantes-Oliveira, Joshua Lehrer-Graiwer, Honor Hsin, Andrew G Fraser, Ravi S Kamath, Julie Ahringer, and Cynthia Kenyon. Rates of behavior and aging specified by mitochondrial function during development. *Science*, 298(5602):2398–2401, 2002.

- [65] Nizam Ud Din and Ji Yu. Training a deep learning model for single-cell segmentation without manual annotation. *Scientific reports*, 11(1):23995, 2021.
- [66] Lingraj Dora, Sanjay Agrawal, Rutuparna Panda, and Ajith Abraham. State-of-the-art methods for brain tissue segmentation: A review. *IEEE reviews in biomedical engineering*, 10:235–249, 2017.
- [67] Jonas F Dorn, Gaudenz Danuser, and Ge Yang. Computational processing and analysis of dynamic fluorescence image data. *Methods in cell biology*, 85:497–538, 2008.
- [68] Kenneth H Downing and Eva Nogales. Tubulin and microtubule structure. *Current opinion in cell biology*, 10(1):16–22, 1998.
- [69] Oliver Dürr and Beate Sick. Single-cell phenotype classification using deep convolutional neural networks. *Journal of biomolecular screening*, 21(9):998–1003, 2016.
- [70] Roland Eils and Chaitanya Athale. Computational imaging in cell biology. *The Journal of cell biology*, 161(3):477–481, 2003.
- [71] Albert Einstein. On the movement of small particles suspended in a stationary liquid demanded by the molecular-kinetic theory of heat. *Annalen der physik*, 17:549–560, 1905.
- [72] Albert Einstein et al. On the electrodynamics of moving bodies. *Annalen der physik*, 17(10):891–921, 1905.
- [73] Sandra E Encalada and Lawrence SB Goldstein. Biophysical challenges to axonal transport: motor-cargo deficiencies and neurodegeneration. *Annual review of biophysics*, 43:141–169, 2014.

- [74] Fred Etoc, Elie Balloul, Chiara Vicario, Davide Normanno, Domenik Liße, Assa Sittner, Jacob Piehler, Maxime Dahan, and Mathieu Coppey. Non-specific interactions govern cytosolic diffusion of nanosized objects in mammalian cells. *Nature materials*, 17(8):740–746, 2018.
- [75] Philipp Eulenberg, Niklas Köhler, Thomas Blasi, Andrew Filby, Anne E Carpenter, Paul Rees, Fabian J Theis, and F Alexander Wolf. Reconstructing cell cycle and disease progression using deep learning. *Nature communications*, 8(1):463, 2017.
- [76] Ben Fabry, Geoffrey N Maksym, James P Butler, Michael Glogauer, Daniel Navajas, and Jeffrey J Fredberg. Scaling the microrheology of living cells. *Physical review letters*, 87(14):148102, 2001.
- [77] Frank Fan and Keith V Wood. Bioluminescent assays for high-throughput screening. *Assay and drug development technologies*, 5(1):127–136, 2007.
- [78] Aly A Farag. Edge-based image segmentation. *Remote sensing reviews*, 6(1):95–121, 1992.
- [79] R. Farhadifar and D. Needleman. Automated segmentation of the first mitotic spindle in differential interference contrast microscopy images of *C. elegans* embryos. *Methods Mol. Biol.*, 1136:41–45, 2014.
- [80] Reza Farhadifar, Charles F Baer, Aurore-Cécile Valfort, Erik C Andersen, Thomas Müller-Reichert, Marie Delattre, and Daniel J Needleman. Scaling, selection, and evolutionary dynamics of the mitotic spindle. *Current biology*, 25(6):732–740, 2015.
- [81] Reza Farhadifar and Daniel Needleman. Automated segmentation of the first mitotic spindle in differential interference contrast microscopy images of *c. elegans* embryos. *Mitosis: Methods and protocols*, pages 41–45, 2014.

- [82] Reza Farhadifar, Che-Hang Yu, Gunar Fabig, Hai-Yin Wu, David B Stein, Matthew Rockman, Thomas Müller-Reichert, Michael J Shelley, and Daniel J Needleman. Stoichiometric interactions explain spindle dynamics and scaling across 100 million years of nematode evolution. *Elife*, 9:e55877, 2020.
- [83] Soheil Fatehiboroujeni, Arvind Gopinath, and Sachin Goyal. Three-dimensional nonlinear dynamics of prestressed active filaments: flapping, swirling, and flipping. *Physical Review E*, 103(1):013005, 2021.
- [84] Luis Ferro, Pedro Leal, Marco Marques, Joana Maciel, Marta I Oliveira, Mario A Barbosa, and Pedro Quelhas. Multinuclear cell analysis using laplacian of gaussian and delaunay graphs. In *Pattern Recognition and Image Analysis: 6th Iberian Conference, IbPRIA 2013, Funchal, Madeira, Portugal, June 5-7, 2013. Proceedings 6*, pages 441–449. Springer, 2013.
- [85] Yaouen Fily, Priya Subramanian, Tobias M Schneider, Raghunath Chelakkot, and Arvind Gopinath. Buckling instabilities and spatio-temporal dynamics of active elastic filaments. *Journal of the Royal Society Interface*, 17(165):20190794, 2020.
- [86] Itzhak Fishov and Conrad L Woldringh. Visualization of membrane domains in escherichia coli. *Molecular microbiology*, 32(6):1166–1172, 1999.
- [87] GM Foody, MB McCulloch, and WB Yates. The effect of training set size and composition on artificial neural network classification. *International Journal of Remote Sensing*, 16(9):1707–1723, 1995.
- [88] Manuel G Forero, Luis H Rodriguez, and Sergio L Miranda. Evaluation of segmentation techniques for cell tracking in confocal microscopy images. In *Applications of Digital Image Processing XLIV*, volume 11842, pages 342–351. SPIE, 2021.

- [89] Christian W Freudiger, Wei Min, Brian G Saar, Sijia Lu, Gary R Holtom, Chengwei He, Jason C Tsai, Jing X Kang, and X Sunney Xie. Label-free biomedical imaging with high sensitivity by stimulated raman scattering microscopy. *Science*, 322(5909):1857–1861, 2008.
- [90] Naama Gal, Diana Lechtman-Goldstein, and Daphne Weihs. Particle tracking in living cells: a review of the mean square displacement method and beyond. *Rheologica Acta*, 52:425–443, 2013.
- [91] Yuval Garini, Bart J Vermolen, and Ian T Young. From micro to nano: recent advances in high-resolution microscopy. *Current opinion in biotechnology*, 16(1):3–12, 2005.
- [92] Carlos Garzon-Coral, Horatiu A. Fantana, and Jonathon Howard. A force-generating machinery maintains the spindle at the cell center during mitosis. *Science*, 352(6289):1124–1127, may 2016.
- [93] Carlos Garzon-Coral, Horatiu A Fantana, and Jonathon Howard. A force-generating machinery maintains the spindle at the cell center during mitosis. *Science*, 352(6289):1124–1127, 2016.
- [94] Arkadiusz Gertych, Zhaoxuan Ma, Jian Tajbakhsh, Adriana Velásquez-Vacca, and Beatrice S Knudsen. Rapid 3-d delineation of cell nuclei for high-content screening platforms. *Computers in biology and medicine*, 69:328–338, 2016.
- [95] F. Gittes, E. fer, S. Baek, and J. Howard. Directional loading of the kinesin motor molecule as it buckles a microtubule. *Biophys J*, 70(1):418–429, Jan 1996.
- [96] F. Gittes, B. Mickey, J. Nettleton, and J. Howard. Flexural rigidity of microtubules and actin filaments measured from thermal fluctuations in shape. *J Cell Biol*, 120(4):923–934, Feb 1993.

- [97] Claudiu A Giurumescu and Andrew D Chisholm. Cell identification and cell lineage analysis. *Methods Cell Biol*, 106:325–41, 2011.
- [98] Brian Glancy. Mitonet: A generalizable model for segmentation of individual mitochondria within electron microscopy datasets. *Cell Systems*, 14(1):7–8, 2023.
- [99] Primož Godec, Matjaž Pančur, Nejc Ilenič, Andrej Čopar, Martin Stražar, Aleš Erjavec, Ajda Pretnar, Janez Demšar, Anže Starič, Marko Toplak, et al. Democratized image analytics by visual programming through integration of deep models and small-scale machine learning. *Nature communications*, 10(1):4551, 2019.
- [100] William J Godinez, Imtiaz Hossain, Stanley E Lazic, John W Davies, and Xian Zhang. A multi-scale convolutional neural network for phenotyping high-content cellular images. *Bioinformatics*, 33(13):2010–2019, 2017.
- [101] Bob Goldstein. On the evolution of early development in the nematoda. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 356(1414):1521–1531, 2001.
- [102] P Gönczy, H Schnabel, T Kaletta, A D Amores, T Hyman, and R Schnabel. Dissection of cell division processes in the one cell stage caenorhabditis elegans embryo by mutational analysis. *J Cell Biol*, 144(5):927–46, Mar 1999.
- [103] Pierre Gönczy, Christophe Echeverri, Karen Oegema, Alan Coulson, Steven JM Jones, Richard R Copley, John Duperon, Jeff Oegema, Michael Brehm, Etienne Cassin, et al. Functional genomic analysis of cell division in c. elegans using rnai of genes on chromosome iii. *Nature*, 408(6810):331–336, 2000.
- [104] Pierre Gönczy, Heinke Schnabel, Titus Kaletta, Ana Duran Amores, Tony Hyman, and Ralf Schnabel. Dissection of cell division processes in the one cell

- stage *caenorhabditis elegans* embryo by mutational analysis. *The Journal of cell biology*, 144(5):927–946, 1999.
- [105] Rafael C Gonzalez. *Digital image processing*. Pearson education india, 2009.
- [106] Noemi Gozzi, Edoardo Giacomello, Martina Sollini, Margarita Kirienko, Angela Ammirabile, Pierluca Lanzi, Daniele Loiacono, and Arturo Chiti. Image embeddings extracted from cnns outperform other transfer learning approaches in classification of chest radiographs. *Diagnostics*, 12(9):2084, 2022.
- [107] Stephan W Grill, Pierre GoËnczy, Ernst HK Stelzer, and Anthony A Hyman. Polarity controls forces governing asymmetric spindle positioning in the *caenorhabditis elegans* embryo. *Nature*, 409(6820):630–633, 2001.
- [108] Stephan W Grill, Jonathon Howard, Erik Schaffer, Ernst HK Stelzer, and Anthony A Hyman. The distribution of active force generators controls mitotic spindle position. *Science*, 301(5632):518–521, 2003.
- [109] Stephan W Grill, Karsten Kruse, and Frank Jülicher. Theory of mitotic spindle oscillations. *Physical review letters*, 94(10):108104, 2005.
- [110] S. Gunn. On the discrete representation of the laplacian of gaussian. *Pattern Recognit.*, 32:1463–1472, 1999.
- [111] Stephanie L Gupton, Wendy C Salmon, and Clare M Waterman-Storer. Converging populations of f-actin promote breakage of associated microtubules to spatially regulate microtubule turnover in migrating cells. *Current biology*, 12(22):1891–1899, 2002.
- [112] Eric S Haag, David HA Fitch, and Marie Delattre. From “the worm” to “the worms” and back again: the evolutionary developmental biology of nematodes. *Genetics*, 210(2):397–433, 2018.

- [113] Shugo Hamahashi, Shuichi Onami, and Hiroaki Kitano. Detection of nuclei in 4d nomarski dic microscope images of early *caenorhabditis elegans* embryos using local image entropy and object tracking. *BMC bioinformatics*, 6:1–15, 2005.
- [114] Hiroyasu Hatakeyama, Yoshihito Nakahata, Hirokazu Yarimizu, and Makoto Kanzaki. Live-cell single-molecule labeling and analysis of myosin motors with quantum dots. *Molecular biology of the cell*, 28(1):173–181, 2017.
- [115] Taviare Hawkins, Matthew Mirigian, M Selcuk Yasar, and Jennifer L Ross. Mechanics of microtubules. *Journal of biomechanics*, 43(1):23–30, 2010.
- [116] K Hayashi, CG Pack, MK Sato, K Mouri, K Kaizu, K Takahashi, and Y Okada. Viscosity and drag force involved in organelle transport: investigation of the fluctuation dissipation theorem. *The European Physical Journal E*, 36:1–7, 2013.
- [117] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
- [118] Rainer Heintzmann and Thomas Huser. Super-resolution structured illumination microscopy. *Chemical reviews*, 117(23):13890–13908, 2017.
- [119] Greg J Hermann, Lena K Schroeder, Caroline A Hieb, Aaron M Kershner, Beverley M Rabbitts, Paul Fonarev, Barth D Grant, and James R Priess. Genetic analysis of lysosomal trafficking in *caenorhabditis elegans*. *Molecular biology of the cell*, 16(7):3273–3288, 2005.
- [120] Steven N Hird and John G White. Cortical and cytoplasmic flow polarity in early embryonic cells of *caenorhabditis elegans*. *The Journal of cell biology*, 121(6):1343–1355, 1993.

- [121] Hao Hu. Poisson distribution and application. *A Course in Department of Physics and Astronomy; University of Tennessee at Knoxville: Knoxville, TN, USA*, 2008.
- [122] Jan Huisken, Jim Swoger, Filippo Del Bene, Joachim Wittbrodt, and Ernst HK Stelzer. Optical sectioning deep inside live embryos by selective plane illumination microscopy. *Science*, 305(5686):1007–1009, 2004.
- [123] Donald E Ingber. Tensegrity: the architectural basis of cellular mechanotransduction. *Annual review of physiology*, 59(1):575–599, 1997.
- [124] Kunalika Jain, Neha Khetan, Shivani A Yadav, Saravanan Palani, and Chaitanya A Athale. Collective dynein transport of the nucleus by pulling on astral microtubules during *saccharomyces cerevisiae* mitosis. *Yeast*, 38(6):352–366, 2021.
- [125] Chris Janetopoulos, Tian Jin, and Peter Devreotes. Receptor-mediated activation of heterotrimeric g-proteins in living cells. *Science*, 291(5512):2408–2411, 2001.
- [126] Marcel E Janson and Marileen Dogterom. A bending mode analysis for growing microtubules: evidence for a velocity-dependent rigidity. *Biophysical journal*, 87(4):2723–2736, 2004.
- [127] Hongyuan Jiang. Cell size modulates oscillation, positioning and length of mitotic spindles. *Scientific reports*, 5(1):10504, 2015.
- [128] Xin Jin, Yuanwen Zou, and Zhongbing Huang. An imbalanced image classification method for the cell cycle phase. *Information*, 12(6):249, 2021.
- [129] Jia-Yang Juang, Pin-Yi Chen, Da-Chang Yang, Shang-Ping Wu, An Yen, and Hsin-I Hsieh. The avian egg exhibits general allometric invariances in mechanical design. *Scientific Reports*, 7(1):14205, 2017.

- [130] Tanvi Kale, Dhruv Khatri, and Chaitanya A Athale. Allometry of escherichia coli surface area with volume: effect of size variability, filamentation and division dynamics. *Physical Biology*, 20(4):046007, 2023.
- [131] Nick Kanopoulos, Nagesh Vasanthavada, and Robert L Baker. Design of an image edge detection filter using the sobel operator. *IEEE Journal of solid-state circuits*, 23(2):358–367, 1988.
- [132] Varun Kapoor, William G Hirst, Christoph Hentschel, Stephan Preibisch, and Simone Reber. Mtrack: automated detection, tracking, and analysis of dynamic microtubules. *Scientific Reports*, 9(1):3794, 2019.
- [133] Richard Kasprovicz, Rakesh Suman, and Peter O’Toole. Characterising live cell behaviour: Traditional label-free and quantitative phase imaging approaches. *The international journal of biochemistry & cell biology*, 84:89–95, 2017.
- [134] Michael Kass, Andrew Witkin, and Demetri Terzopoulos. Snakes: Active contour models. *International journal of computer vision*, 1(4):321–331, 1988.
- [135] ALEC D Keith, ERNEST C Pollard, and W Snipes. Inositol-less death in yeast results in a simultaneous increase in intracellular viscosity. *Biophysical journal*, 17(3):205–212, 1977.
- [136] Dhruv Khatri, Thibault Brugière, Chaitanya A Athale, and Marie Delattre. Evolutionary divergence of anaphase spindle mechanics in nematode embryos constrained by antagonistic pulling and viscous forces. *Molecular Biology of the Cell*, 33(6):ar61, 2022.
- [137] Neha Khetan and Chaitanya A Athale. A motor-gradient and clustering model of the centripetal motility of mtocs in meiosis i of mouse oocytes. *PLoS Computational Biology*, 12(10):e1005102, 2016.

- [138] Neha Khetan and Chaitanya A Athale. Aster swarming by symmetry breaking of cortical dynein transport and coupling kinesins. *Soft Matter*, 16(37):8554–8564, 2020.
- [139] Alexey Khodjakov, Sabrina La Terra, and Fred Chang. Laser microsurgery in fission yeast: role of the mitotic spindle midzone in anaphase b. *Current Biology*, 14(15):1330–1340, 2004.
- [140] Tobias R Kießling, Roland Stange, Josef A Käs, and Anatol W Fritsch. Thermorheology of living cells—impact of temperature variations on cell mechanics. *New Journal of Physics*, 15(4):045026, 2013.
- [141] Akatsuki Kimura and Shuichi Onami. Computer simulations and image processing reveal length-dependent pulling force as the primary mechanism for c. elegans male pronuclear migration. *Developmental cell*, 8(5):765–775, 2005.
- [142] Hui Kong, Hatice Cinar Akakin, and Sanjay E Sarma. A generalized laplacian of gaussian filter for blob detection and its applications. *IEEE transactions on cybernetics*, 43(6):1719–1733, 2013.
- [143] Karsten König. Multiphoton microscopy in life sciences. *Journal of microscopy*, 200(2):83–104, 2000.
- [144] Kyung-Mo Koo and Eui-Young Cha. Image recognition performance enhancements using image normalization. *Human-centric Computing and Information Sciences*, 7:1–11, 2017.
- [145] Sachin Kotak. Mechanisms of spindle positioning: lessons from worms and mammalian cells. *Biomolecules*, 9(2):80, 2019.
- [146] Cleopatra Kozlowski, Martin Srayko, and Francois Nedelec. Cortical microtubule contacts position the spindle in c. elegans embryos. *Cell*, 129(3):499–510, 2007.

- [147] Cleopatra Kozlowski, Martin Srayko, and Francois Nedelec. Cortical Microtubule Contacts Position the Spindle in *C. elegans* Embryos. *Cell*, 129(3):499–510, 2007.
- [148] Irvin M Krieger and Thomas J Dougherty. A mechanism for non-newtonian flow in suspensions of rigid spheres. *Trans. Soc. Rheol*, 3(1):137–152, 1959.
- [149] Scott Krig and Scott Krig. Image pre-processing. *Computer Vision Metrics: Textbook Edition*, pages 35–74, 2016.
- [150] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6):84–90, 2017.
- [151] B. K. Shreyamsha Kumar. Image denoising based on gaussian/bilateral filter and its method noise thresholding. *Signal Image Video Process.*, 7:1159–1172, 2013.
- [152] Jean-Claude Labbé, Erin K McCarthy, and Bob Goldstein. The forces that position a mitotic spindle asymmetrically are tethered until after the time of spindle assembly. *The Journal of cell biology*, 167(2):245–256, 2004.
- [153] Ta-Chih Lee, Rangasami L Kashyap, and Chong-Nam Chu. Building skeleton models via 3-d medial surface axis thinning algorithms. *CVGIP: graphical models and image processing*, 56(6):462–478, 1994.
- [154] Leon Lenchik, Laura Heacock, Ashley A Weaver, Robert D Boutin, Tessa S Cook, Jason Itri, Christopher G Filippi, Rao P Gullapalli, James Lee, Marianna Zagurovskaya, et al. Automated segmentation of tissues using ct and mri: a systematic review. *Academic radiology*, 26(12):1695–1706, 2019.
- [155] Teng Li. A mechanics model of microtubule buckling in living cells. *Journal of biomechanics*, 41(8):1722–1729, 2008.

- [156] Gang Lin, Umesh Adiga, Kathy Olson, John F Guzowski, Carol A Barnes, and Badrinath Roysam. A hybrid 3d watershed algorithm incorporating gradient cues and object models for automatic segmentation of nuclei in confocal image stacks. *Cytometry Part A: the journal of the International Society for Analytical Cytology*, 56(1):23–36, 2003.
- [157] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision*, pages 2980–2988, 2017.
- [158] Lynn Liu, Erkan Tüzel, and Jennifer L Ross. Loop formation of microtubules during gliding at high density. *Journal of Physics: Condensed Matter*, 23(37):374104, 2011.
- [159] Yanli Liu, Yourong Wang, and Jian Zhang. New machine learning algorithm: Random forest. In *Information Computing and Applications: Third International Conference, ICICA 2012, Chengde, China, September 14-16, 2012. Proceedings 3*, pages 246–252. Springer, 2012.
- [160] Zhichao Liu, Luhong Jin, Jincheng Chen, Qiuyu Fang, Sergey Ablameyko, Zhaozheng Yin, and Yingke Xu. A survey on applications of deep learning in microscopy image analysis. *Computers in biology and medicine*, 134:104523, 2021.
- [161] Xinghua Lou, Minjung Kang, Panagiotis Xenopoulos, Silvia Munoz-Descalzo, and Anna-Katerina Hadjantonakis. A rapid and efficient 2d/3d nuclear segmentation method for analysis of early mouse embryo and stem cell image data. *Stem cell reports*, 2(3):382–397, 2014.
- [162] Sehla Loussaief and Afef Abdelkrim. Machine learning framework for image classification. In *2016 7th International Conference on Sciences of Electronics, Tech-*

- nologies of Information and Telecommunications (SETIT)*, pages 58–61. IEEE, 2016.
- [163] D. Lu and Qihao Weng. A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, 28:823 – 870, 2007.
- [164] Dengsheng Lu and Qihao Weng. A survey of image classification methods and techniques for improving classification performance. *International journal of Remote sensing*, 28(5):823–870, 2007.
- [165] Zhiyuan Lu, C Shan Xu, Kenneth J Hayworth, Song Pang, Stephen M Plaza, Louis K Scheffer, Gerald M Rubin, Harald F Hess, Patricia K Rivlin, and Ian A Meinertzhagen. En bloc preparation of drosophila brains enables high-throughput fib-sem connectomics. *Frontiers in Neural Circuits*, 16:917251, 2022.
- [166] Katherine Luby-Phelps. Physical properties of cytoplasm. *Current opinion in cell biology*, 6(1):3–9, 1994.
- [167] Katherine Luby-Phelps. Cytoarchitecture and physical properties of cytoplasm: volume, viscosity, diffusion, intracellular surface area. *International review of cytology*, 192:189–221, 1999.
- [168] Anthony L Luz, Cristina Lagido, Matthew D Hirschey, and Joel N Meyer. In vivo determination of mitochondrial function using luciferase-expressing caenorhabditis elegans: Contribution of oxidative phosphorylation, glycolysis, and fatty acid oxidation to toxicant-induced dysfunction. *Current protocols in toxicology*, 69(1):25–8, 2016.
- [169] Michael Lynch. The frailty of adaptive hypotheses for the origins of organismal complexity. *Proceedings of the National Academy of Sciences*, 104(suppl_1):8597–8604, 2007.

- [170] Michael Lynch, Mark C Field, Holly V Goodson, Harmit S Malik, José B Pereira-Leal, David S Roos, Aaron P Turkewitz, and Shelley Sazer. Evolutionary cell biology: two origins, one objective. *Proceedings of the National Academy of Sciences*, 111(48):16990–16994, 2014.
- [171] Rui Ma, Liedewij Laan, Marileen Dogterom, Nenad Pavin, and Frank Jülicher. General theory for the mechanics of confined microtubule asters. *New Journal of Physics*, 16(1):013018, 2014.
- [172] FC MacKintosh and CF Schmidt. Microrheology. *Current opinion in colloid & interface science*, 4(4):300–307, 1999.
- [173] Fred C MacKintosh and Alex J Levine. Nonequilibrium mechanics and dynamics of motor-activated gels. *Physical review letters*, 100(1):018104, 2008.
- [174] Sébastien Marcel and Yann Rodriguez. Torchvision the machine-vision package of torch. In *Proceedings of the 18th ACM international conference on Multimedia*, pages 1485–1488, 2010.
- [175] Lynn Margulis. Origin of eukaryotic cells: evidence and research implications for a theory of the origin and evolution of microbial, plant, and animal cells on the precambrian earth. *(No Title)*, 1970.
- [176] Martin Maška, Vladimír Ulman, David Svoboda, Pavel Matula, Petr Matula, Cristina Ederra, Ainhoa Urbiola, Tomás España, Subramanian Venkatesan, Deepak MW Balak, et al. A benchmark for comparison of cell tracking algorithms. *Bioinformatics*, 30(11):1609–1617, 2014.
- [177] Attila Mate. The frenet-serret formulas. *Brooklyn Collage Of The City University Of New York, izdano*, 19, 2017.

- [178] Manos Mavrakakis, Richa Rikhy, Mary Lilly, and Jennifer Lippincott-Schwartz. Fluorescence imaging techniques for studying drosophila embryo development. *Current protocols in cell biology*, 39(1):4–18, 2008.
- [179] Magdalena Mazur-Milecka and Jacek Ruminski. Deep learning based thermal image segmentation for laboratory animals tracking. *Quantitative InfraRed Thermography Journal*, 18(3):159–176, 2021.
- [180] Leland McInnes, John Healy, and James Melville. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018.
- [181] Timothy J McIntyre, Christian Maurer, Stephanie Fassel, Saranjam Khan, Stefan Bernet, and Monika Ritsch-Marte. Quantitative slm-based differential interference contrast imaging. *Opt Express*, 18(13):14063–78, Jun 2010.
- [182] Erik Meijering. Cell segmentation: 50 years down the road [life sciences]. *IEEE signal processing magazine*, 29(5):140–145, 2012.
- [183] Erik Meijering, Oleh Dzyubachyk, Ihor Smal, and Wiggert A van Cappellen. Tracking in cell and developmental biology. In *Seminars in cell & developmental biology*, volume 20, pages 894–902. Elsevier, 2009.
- [184] Nan Meng, Edmund Y Lam, Kevin K Tsia, and Hayden Kwok-Hay So. Large-scale multi-class image-based cell classification with deep learning. *IEEE journal of biomedical and health informatics*, 23(5):2091–2098, 2018.
- [185] B. Misselwitz, G. Strittmatter, Balamurugan Periaswamy, M. Schlumberger, Samuel Rout, Péter Horváth, K. Kozak, and W. Hardt. Enhanced cellclassifier: a multi-class classification tool for microscopy images. *BMC Bioinformatics*, 11:30 – 30, 2010.

- [186] Kota Miura. Tracking movement in cell biology. *Microscopy Techniques: -/-*, pages 267–295, 2005.
- [187] Daisuke Mizuno, Catherine Tardin, Christoph F Schmidt, and Frederik C MacKintosh. Nonequilibrium mechanics of active cytoskeletal networks. *Science*, 315(5810):370–373, 2007.
- [188] Phillip A Mlsna and Jeffrey J Rodriguez. Gradient and laplacian edge detection. In *The essential guide to image processing*, pages 495–524. Elsevier, 2009.
- [189] Erick Moen, Dylan Bannon, Takamasa Kudo, William Graf, Markus Covert, and David Van Valen. Deep learning for cellular image analysis. *Nature methods*, 16(12):1233–1246, 2019.
- [190] Alex Mogilner and Angelika Manhart. Intracellular fluid mechanics: Coupling cytoplasmic flow with active cytoskeletal gel. *Annual Review of Fluid Mechanics*, 50:347–370, 2018.
- [191] Cleve B Moler. *Numerical computing with MATLAB*. SIAM, 2004.
- [192] Viola Mönkemöller, Cristina Øie, Wolfgang Hübner, Thomas Huser, and Peter McCourt. Multimodal super-resolution optical microscopy visualizes the close connection between membrane and the cytoskeleton in liver sinusoidal endothelial cell fenestrations. *Scientific reports*, 5(1):16279, 2015.
- [193] Lewis S Mosby, Marco Polin, and Darius Vasco Köster. A python based automated tracking routine for myosin ii filaments. *Journal of Physics D: Applied Physics*, 53(30):304002, 2020.
- [194] Khushboo Munir, Hassan Elahi, Afsheen Ayub, Fabrizio Frezza, and Antonello Rizzi. Cancer diagnosis using deep learning: a bibliographic review. *Cancers*, 11(9):1235, 2019.

- [195] Gene Myers. Why bioimage informatics matters. *Nature methods*, 9(7):659–660, 2012.
- [196] Yukiko Nagao, Mika Sakamoto, Takumi Chinen, Yasushi Okada, and Daisuke Takao. Robust classification of cell cycle phase and biological feature extraction by image-based deep learning. *Molecular biology of the cell*, 31(13):1346–1354, 2020.
- [197] Javad Najafi, Serge Dmitrieff, and Nicolas Minc. Size-and position-dependent cytoplasm viscoelasticity through hydrodynamic interactions with the cell surface. *Proceedings of the National Academy of Sciences*, 120(9):e2216839120, 2023.
- [198] Kaustav Nandy, Rama Chellappa, Amit Kumar, and Stephen J Lockett. Segmentation of nuclei from 3d microscopy images of tissue via graphcut optimization. *IEEE Journal of Selected Topics in Signal Processing*, 10(1):140–150, 2015.
- [199] Francois Nedelec and Dietrich Foethke. Collective langevin dynamics of flexible cytoskeletal fibers. *New Journal of Physics*, 9(11):427, 2007.
- [200] Karl J Niklas. The cell walls that bind the tree of life. *BioScience*, 54(9):831–841, 2004.
- [201] Thomas A Nketia, Heba Sailem, Gustavo Rohde, Raghu Machiraju, and Jens Rittscher. Analysis of live cell images: Methods, tools and opportunities. *Methods*, 115:65–79, 2017.
- [202] Karen Oegema, Arshad Desai, Sonja Rybina, Matthew Kirkham, and Anthony A Hyman. Functional analysis of kinetochore assembly in *caenorhabditis elegans*. *The Journal of cell biology*, 153(6):1209–1226, 2001.
- [203] Kazuo Ohmi and Hang-Yu Li. Particle-tracking velocimetry with new algorithms. *Measurement Science and Technology*, 11(6):603, 2000.

- [204] Luana S Ortolan, Marion Avril, Jun Xue, Karl B Seydel, Ying Zheng, and Joseph D Smith. Plasmodium falciparum parasite lines expressing dc8 and group a pfemp1 bind to brain, intestinal, and kidney endothelial cells. *Frontiers in Cellular and Infection Microbiology*, 12:813011, 2022.
- [205] Matteo Osella, Eileen Nugent, and Marco Cosentino Lagomarsino. Concerted control of escherichia coli cell division. *Proceedings of the National Academy of Sciences*, 111(9):3431–3435, 2014.
- [206] Stanley Osher and James A Sethian. Fronts propagating with curvature-dependent speed: Algorithms based on hamilton-jacobi formulations. *Journal of computational physics*, 79(1):12–49, 1988.
- [207] Nobuyuki Otsu. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, 9(1):62–66, 1979.
- [208] Bugra Özdemir and Ralf Reski. Automated and semi-automated enhancement, segmentation and tracing of cytoskeletal networks in microscopic images: A review. *Computational and structural biotechnology journal*, 19:2106–2120, 2021.
- [209] Niall O’Mahony, Sean Campbell, Anderson Carvalho, Suman Harapanahalli, Gustavo Velasco Hernandez, Lenka Krpalkova, Daniel Riordan, and Joseph Walsh. Deep learning vs. traditional computer vision. In *Advances in Computer Vision: Proceedings of the 2019 Computer Vision Conference (CVC), Volume 1*, pages 128–144. Springer, 2020.
- [210] Marius Pachitariu and Carsen Stringer. Cellpose 2.0: how to train your own model. *Nature methods*, 19(12):1634–1641, 2022.
- [211] Kuldeep K Paliwal, Anant Agarwal, and Sarvajit S Sinha. A modification over sakoe and chiba’s dynamic time warping algorithm for isolated word recognition. *Signal Processing*, 4(4):329–333, 1982.

- [212] Sang-Hyun Park and Ronald T Raines. Green fluorescent protein as a signal for protein–protein interactions. *Protein science*, 6(11):2344–2349, 1997.
- [213] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32, 2019.
- [214] Pawan Kumar Patidar, Lalit Baldev Singh, Gaurav Bagaria, T. K. Thivakaran, Charu Khare, Kapil Kumar Nagwanshi, Mukesh C. Motwani, Mukesh C. Gadiya, Rakhi C. Motwani, Frederick C. Harris, Mozammel Hoque Chowdhury, Md. Ezharul Islam, Nasima Begum, James C. Church, Yixin Chen, and Alan Conrad Bovik. Image filtering using linear and non linear filter for gaussian noise. *International Journal of Computer Applications*, 93:29–34, 2014.
- [215] Jacques Pecreaux, Jens Christian Röper, Karsten Kruse, Frank Jülicher, Anthony A. Hyman, Stephan W. Grill, and Jonathon Howard. Spindle Oscillations during Asymmetric Cell Division Require a Threshold Number of Active Cortical Force Generators. *Current Biology*, 16(21):2111–2122, 2006.
- [216] Jacques Pecreaux, Jens-Christian Röper, Karsten Kruse, Frank Jülicher, Anthony A Hyman, Stephan W Grill, and Jonathon Howard. Spindle oscillations during asymmetric cell division require a threshold number of active cortical force generators. *Current Biology*, 16(21):2111–2122, 2006.
- [217] Hanchuan Peng, Jie Zhou, Zhi Zhou, Alessandro Bria, Yujie Li, Dean Mark Kleissas, Nathan G Drenkow, Brian Long, Xiaoxiao Liu, and Hanbo Chen. Bioimage informatics for big data. *Focus on Bio-Image Informatics*, pages 263–272, 2016.
- [218] Laura B Persson, Vardhaan S Ambati, and Onn Brandman. Cellular con-

- trol of viscosity counters changes in temperature and energy availability. *Cell*, 183(6):1572–1585, 2020.
- [219] Fabio Piano, Aaron J Schetter, Marco Mangone, Lincoln Stein, and Kenneth J Kemphues. Rnai analysis of genes expressed in the ovary of caenorhabditis elegans. *Current Biology*, 10(24):1619–1622, 2000.
- [220] Maksymilian Pluta. Nomarski’s dic microscopy: a review. *Phase contrast and differential interference contrast imaging techniques and applications*, 1846:10–25, 1994.
- [221] Ting-Chuen Pong, Linda G Shapiro, Layne T Watson, and Robert M Haralick. Experiments in segmentation using a facet model region grower. *Computer vision, graphics, and image processing*, 25(1):1–23, 1984.
- [222] Clémentine Puech, Sophie Hodin, Valérie Forest, Zhiguo He, Patrick Mismetti, Xavier Delavenne, and Nathalie Perek. Assessment of hbec-5i endothelial cell line cultivated in astrocyte conditioned medium as a human blood-brain barrier model for abc drug transport studies. *International journal of pharmaceutics*, 551(1-2):281–289, 2018.
- [223] D Quemada. Rheology of concentrated disperse systems and minimum energy dissipation principle: I. viscosity-concentration relationship. *Rheologica Acta*, 16(1):82–94, 1977.
- [224] Isabela Ramos, Ednildo Machado, Hatisaburo Masuda, and Fabio Gomes. Open questions on the functional biology of the yolk granules during embryo development. *Molecular Reproduction and Development*, 89(2):86–94, 2022.
- [225] Daniel H Rapoport, Tim Becker, Amir Madany Mamlouk, Simone Schicktanz, and Charli Kruse. A novel validation algorithm allows for automated cell tracking

- and the extraction of biologically meaningful parameters. *PloS one*, 6(11):e27315, 2011.
- [226] Luca Rappez, Alexander Rakhlin, Angelos Rigopoulos, Sergey Nikolenko, and Theodore Alexandrov. Deepcycle reconstructs a cyclic cell cycle trajectory from unsegmented cell images using convolutional neural networks. *Molecular systems biology*, 16(10):e9474, 2020.
- [227] Luca Rappez, Alexander Rakhlin, Angelos Rigopoulos, Sergey Nikolenko, and Theodore Alexandrov. DeepCycle reconstructs a cyclic cell cycle trajectory from unsegmented cell images using convolutional neural networks. *Molecular Systems Biology*, 16(10), oct 2020.
- [228] T Riklin Raviv, Vebjorn Ljosa, Annie L Conery, Frederick M Ausubel, Anne E Carpenter, Polina Golland, and Carolina Wählby. Morphology-guided graph search for untangling objects: C. elegans analysis. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2010: 13th International Conference, Beijing, China, September 20-24, 2010, Proceedings, Part III 13*, pages 634–641. Springer, 2010.
- [229] Simone Reber and Nathan W Goehring. Intracellular scaling mechanisms. *Cold Spring Harbor Perspectives in Biology*, 7(12):a019067, 2015.
- [230] Elisa Maria Rieckhoff, Frederic Berndt, Maria Elsner, Stefan Golfier, Franziska Decker, Keisuke Ishihara, and Jan Brugués. Spindle scaling is governed by cell boundary regulation of microtubule nucleation. *Current Biology*, 30(24):4973–4983, 2020.
- [231] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Lecture Notes in Computer Science*

- (including subseries *Lecture Notes in Artificial Intelligence* and *Lecture Notes in Bioinformatics*), volume 9351, pages 234–241. Springer Verlag, 2015.
- [232] Felix Ruhnnow, David Zwicker, and Stefan Diez. Tracking single particles and elongated filaments with nanometer precision. *Biophysical journal*, 100(11):2820–2828, 2011.
 - [233] Tim Sainburg, Leland McInnes, and Timothy Q Gentner. Parametric umap embeddings for representation and semisupervised learning. *Neural Computation*, 33(11):2881–2907, 2021.
 - [234] H. Sakoe and S. Chiba. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26(1):43–49, 1978.
 - [235] Abraham Savitzky and Marcel JE Golay. Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry*, 36(8):1627–1639, 1964.
 - [236] Johannes Schindelin, Ignacio Arganda-Carreras, Erwin Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Stephan Preibisch, Curtis Rueden, Stephan Saalfeld, Benjamin Schmid, et al. Fiji: an open-source platform for biological-image analysis. *Nature methods*, 9(7):676–682, 2012.
 - [237] Philip Sedgwick. Pearson’s correlation coefficient. *Bmj*, 345, 2012.
 - [238] Ramprasaath R Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE international conference on computer vision*, pages 618–626, 2017.

- [239] Ali Sharif Razavian, Hossein Azizpour, Josephine Sullivan, and Stefan Carlsson. Cnn features off-the-shelf: an astounding baseline for recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 806–813, 2014.
- [240] M. Shifat-E.-Rabbi, Xuwang Yin, Cailey E. Fitzgerald, and G. Rohde. Cell image classification: A comparative overview. *Cytometry Part A*, 97, 2019.
- [241] Connor Shorten and Taghi M Khoshgoftaar. A survey on image data augmentation for deep learning. *Journal of big data*, 6(1):1–48, 2019.
- [242] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
- [243] Michael Sinensky. Homeoviscous adaptation—a homeostatic process that regulates the viscosity of membrane lipids in escherichia coli. *Proceedings of the National Academy of Sciences*, 71(2):522–525, 1974.
- [244] GV Sivanarayana, K Naveen Kumar, Y Srinivas, and GVS Raj Kumar. Review on the methodologies for image segmentation based on cnn. In *Communication Software and Networks: Proceedings of INDIA 2019*, pages 165–175. Springer, 2021.
- [245] Mohammad Soheilypour, Mohaddeseh Peyro, Stephen J Peter, and Mohammad RK Mofrad. Buckling behavior of individual and bundled microtubules. *Biophysical journal*, 108(7):1718–1726, 2015.
- [246] Pierre Soille et al. *Morphological image analysis: principles and applications*, volume 2 of 3. Springer, 1999.
- [247] Ralf J Sommer. *Pristionchus pacificus: a nematode model for comparative and evolutionary biology*, volume 11. Brill, 2015.

- [248] M. Sonka, Václav Hlaváč, and R. Boyle. Image pre-processing. pages 56–111, 1993.
- [249] B Sönnichsen, LB Koski, A Walsh, P Marschall, B Neumann, M Brehm, A-M Alleaume, J Artelt, P Bettencourt, E Cassin, et al. Full-genome rnai profiling of early embryogenesis in *caenorhabditis elegans*. *Nature*, 434(7032):462–469, 2005.
- [250] Martin Straub, Peter Lodemann, Phillip Holroyd, Reinhard Jahn, and Stefan W Hell. Live cell imaging by multifocal multiphoton microscopy. *European journal of cell biology*, 79(10):726–734, 2000.
- [251] Carsen Stringer, Tim Wang, Michalis Michaelos, and Marius Pachitariu. Cellpose: a generalist algorithm for cellular segmentation. *Nature methods*, 18(1):100–106, 2021.
- [252] Angughali Sumi, Peran Hayes, Arturo D’Angelo, Julien Colombelli, Guillaume Salbreux, Kai Dierkes, and Jérôme Solon. Adherens junction length during tissue contraction is controlled by the mechanosensitive activity of actomyosin and junctional recycling. *Developmental cell*, 47(4):453–463, 2018.
- [253] Zehang Sun, George Bebis, and Ronald Miller. Object detection using feature subset selection. *Pattern recognition*, 37(11):2165–2176, 2004.
- [254] R Sunyer, X Trepát, JJ Fredberg, R Farre, and D Navajas. The temperature dependence of cell mechanics measured by atomic force microscopy. *Physical biology*, 6(2):025009, 2009.
- [255] Shan Suthaharan and Shan Suthaharan. Support vector machine. *Machine learning models and algorithms for big data classification: thinking with examples for effective learning*, pages 207–235, 2016.

- [256] William Sutherland. Lxxv. a dynamical theory of diffusion for non-electrolytes and the molecular mass of albumin. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 9(54):781–785, 1905.
- [257] S Sutradhar and R Paul. Tug-of-war between opposing molecular motors explains chromosomal oscillation during mitosis. *J Theor Biol*, 344:56–69, Mar 2014.
- [258] Jason R Swedlow, Ilya Goldberg, Erik Brauner, and Peter K Sorger. Informatics and quantitative analysis in biological imaging. *Science*, 300(5616):100–102, 2003.
- [259] Jason R Swedlow, Ilya G Goldberg, Kevin W Eliceiri, and OME consortium. Bioimage informatics for experimental biology. *Annual review of biophysics*, 38:327–346, 2009.
- [260] Sattar Taheri-Araghi, Serena Bradde, John T Sauls, Norbert S Hill, Petra Anne Levin, Johan Paulsson, Massimo Vergassola, and Suckjoon Jun. Cell-size control and homeostasis in bacteria. *Current biology*, 25(3):385–391, 2015.
- [261] Shin-ya Takemura, Kenneth J Hayworth, Gary B Huang, Michal Januszewski, Zhiyuan Lu, Elizabeth C Marin, Stephan Preibisch, C Shan Xu, John Bogovic, Andrew S Champion, et al. A connectome of the male drosophila ventral nerve cord. *bioRxiv*, pages 2023–06, 2023.
- [262] Mingxing Tan and Quoc Le. Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*, pages 6105–6114. PMLR, 2019.
- [263] Tian Tian, Yuanyuan Wang, Haitao Wang, Zhaoqi Zhu, and Zhongdang Xiao. Visualizing of the cellular uptake and intracellular trafficking of exosomes by live-cell microscopy. *Journal of cellular biochemistry*, 111(2):488–496, 2010.

- [264] Yonglong Tian, Yue Wang, Dilip Krishnan, Joshua B Tenenbaum, and Phillip Isola. Rethinking few-shot image classification: a good embedding is all you need? In *Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XIV 16*, pages 266–282. Springer, 2020.
- [265] 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.
- [266] Iva M Tolić. Mitotic spindle: kinetochore fibers hold on tight to interpolar bundles. *European Biophysics Journal*, 47(3):191–203, 2018.
- [267] Iva Marija Tolić-Nørrelykke, Emilia-Laura Munteanu, Genevieve Thon, Lene Oddershede, and Kirstine Berg-Sørensen. Anomalous diffusion in living yeast cells. *Physical Review Letters*, 93(7):078102, 2004.
- [268] Sébastien Tosi and Kyra Campbell. 3d tracking of migrating cells from live microscopy time-lapses. *Computer Optimized Microscopy: Methods and Protocols*, pages 385–395, 2019.
- [269] Yu Toyoshima, Terumasa Tokunaga, Osamu Hirose, Manami Kanamori, Takayuki Teramoto, Moon Sun Jang, Sayuri Kuge, Takeshi Ishihara, Ryo Yoshida, and Yuichi Iino. Accurate automatic detection of densely distributed cell nuclei in 3d space. *PLoS computational biology*, 12(6):e1004970, 2016.
- [270] William Y Tsang and Bernard D Lemire. The role of mitochondria in the life of the nematode, *caenorhabditis elegans*. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease*, 1638(2):91–105, 2003.

- [271] Ivan Usov and Raffaele Mezzenga. Fiberapp: an open-source software for tracking and analyzing polymers, filaments, biomacromolecules, and fibrous objects. *Macromolecules*, 48(5):1269–1280, 2015.
- [272] Aurore-Cécile Valfort, Caroline Launay, Marie Sémon, and Marie Delattre. Evolution of mitotic spindle behavior during the first asymmetric embryonic division of nematodes. *PLoS Biology*, 16(1):e2005099, 2018.
- [273] Aurore-Cécile Valfort, Caroline Launay, Marie Sémon, and Marie Delattre. Evolution of mitotic spindle behavior during the first asymmetric embryonic division of nematodes. *PLoS Biol*, 16(1):e2005099, 01 2018.
- [274] Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(11), 2008.
- [275] Giuseppe Vicidomini, Paolo Bianchini, and Alberto Diaspro. Sted super-resolved microscopy. *Nature methods*, 15(3):173–182, 2018.
- [276] A. Vilfan, S. Subramani, E. Bodenschatz, R. Golestanian, and I. Guido. Flagella-like Beating of a Single Microtubule. *Nano Lett*, 19(5):3359–3363, May 2019.
- [277] Pauli Virtanen, Ralf Gommers, Travis E Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, et al. Scipy 1.0: fundamental algorithms for scientific computing in python. *Nature methods*, 17(3):261–272, 2020.
- [278] Arne H Voie, DH Burns, and FA Spelman. Orthogonal-plane fluorescence optical sectioning: Three-dimensional imaging of macroscopic biological specimens. *Journal of microscopy*, 170(3):229–236, 1993.
- [279] Carolina Wählby, I-M Sintorn, Fredrik Erlandsson, Gunilla Borgefors, and Ewert Bengtsson. Combining intensity, edge and shape information for 2d and 3d seg-

- mentation of cell nuclei in tissue sections. *Journal of microscopy*, 215(1):67–76, 2004.
- [280] Gang Wang, Carlos Lopez-Molina, and Bernard De Baets. Automated blob detection using iterative laplacian of gaussian filtering and unilateral second-order gaussian kernels. *Digital Signal Processing*, 96:102592, 2020.
- [281] Pinxian Wang, En Fan, and Peng Wang. Comparative analysis of image classification algorithms based on traditional machine learning and deep learning. *Pattern Recognit. Lett.*, 141:61–67, 2020.
- [282] Shang Wang, Irina V Larina, and Kirill V Larin. Label-free optical imaging in developmental biology. *Biomedical Optics Express*, 11(4):2017–2040, 2020.
- [283] Tae-su Wang, Gi Tae Kim, Minyoung Kim, and Jongwook Jang. Contrast enhancement-based preprocessing process to improve deep learning object task performance and results. *Applied Sciences*, 13(19):10760, 2023.
- [284] Michael L Waskom. Seaborn: statistical data visualization. *Journal of Open Source Software*, 6(60):3021, 2021.
- [285] Greg Welch, Gary Bishop, et al. An introduction to the kalman filter. 1995.
- [286] Kenong Wu, David Gauthier, and Martin D Levine. Live cell image segmentation. *IEEE Transactions on biomedical engineering*, 42(1):1–12, 1995.
- [287] Christian A Wurm, Daniel Neumann, Marcel A Lauterbach, Benjamin Harke, Alexander Egner, Stefan W Hell, and Stefan Jakobs. Nanoscale distribution of mitochondrial import receptor tom20 is adjusted to cellular conditions and exhibits an inner-cellular gradient. *Proceedings of the National Academy of Sciences*, 108(33):13546–13551, 2011.

- [288] Jing Xie, Javad Najafi, Rémi Le Borgne, Jean-Marc Verbavatz, Catherine Durieu, Jeremy Sallé, and Nicolas Minc. Contribution of cytoplasm viscoelastic properties to mitotic spindle positioning. *Proceedings of the National Academy of Sciences*, 119(8):e2115593119, 2022.
- [289] Qizhe Xie, Minh-Thang Luong, Eduard Hovy, and Quoc V Le. Self-training with noisy student improves imagenet classification. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10687–10698, 2020.
- [290] C Shan Xu, Kenneth J Hayworth, Zhiyuan Lu, Patricia Grob, Ahmed M Hassan, José G García-Cerdán, Krishna K Niyogi, Eva Nogales, Richard J Weinberg, and Harald F Hess. Enhanced fib-sem systems for large-volume 3d imaging. *elife*, 6:e25916, 2017.
- [291] Hongming Xu, Cheng Lu, Richard Berendt, Naresh Jha, and Mrinal Mandal. Automatic nuclei detection based on generalized laplacian of gaussian filters. *IEEE journal of biomedical and health informatics*, 21(3):826–837, 2016.
- [292] Ting Xu, Christos Langouras, Maral Adeli Koudehi, Bart E Vos, Ning Wang, Gijsje H Koenderink, Xiaolei Huang, and Dimitrios Vavylonis. Automated tracking of biopolymer growth and network deformation with tsoax. *Scientific reports*, 9(1):1717, 2019.
- [293] Ting Xu, Dimitrios Vavylonis, Feng-Ching Tsai, Gijsje H Koenderink, Wei Nie, Eddy Yusuf, I-Ju Lee, Jian-Qiu Wu, and Xiaolei Huang. Soax: a software for quantification of 3d biopolymer networks. *Scientific reports*, 5(1):9081, 2015.
- [294] Shivani A Yadav and Chaitanya A Athale. Flagella-like beating of microtubules reconstituted in vitro with dynein and pinned filaments. *Biophysical Journal*, 121(3):164a, 2022.

- [295] Shivani A Yadav, Neha Khetan, Dhruv Khatri, and Chaitanya A Athale. Emergence of flagella-like oscillations in single microtubules driven by collective dynein transport. *bioRxiv*, pages 2022–09, 2022.
- [296] Soichiro Yamada, Denis Wirtz, and Scot C Kuo. Mechanics of living cells measured by laser tracking microrheology. *Biophysical journal*, 78(4):1736–1747, 2000.
- [297] Kai Yao, Nash D Rochman, and Sean X Sun. Cell type classification and unsupervised morphological phenotyping from low-resolution images using deep learning. *Scientific reports*, 9(1):13467, 2019.
- [298] Reza Yazdi and Hassan Khotanlou. A survey on automated cell tracking: challenges and solutions. *Multimedia Tools and Applications*, pages 1–37, 2024.
- [299] Kevin D Young. The selective value of bacterial shape. *Microbiology and molecular biology reviews*, 70(3):660–703, 2006.
- [300] Xiaobo Zhou and Stephen TC Wong. Informatics challenges of high-throughput microscopy. *IEEE Signal Processing Magazine*, 23(3):63–72, 2006.
- [301] Roseanna N Zia. Active and passive microrheology: Theory and simulation. *Annual Review of Fluid Mechanics*, 50:371–405, 2018.
- [302] Christophe Zimmer, Bo Zhang, Alexandre Dufour, Ayméric Thébaud, Sylvain Berlemont, Vannary Meas-Yedid, and J-CO Marin. On the digital trail of mobile cells. *IEEE Signal Processing Magazine*, 23(3):54–62, 2006.
- [303] Djemel Ziou and Salvatore Tabbone. Edge detection techniques-an overview. *Pattern Recognition and Image Analysis: Advances in Mathematical Theory and Applications*, 8(4):537–559, 1998.

- [304] Peder Zipperlen, Andrew G Fraser, Ravi S Kamath, Maruxa Martinez-Campos, and Julie Ahringer. Roles for 147 embryonic lethal genes on c. elegans chromosome i identified by rna interference and video microscopy. *The EMBO journal*, 20(15):3984–3992, 2001.